Aus dem Institut für
medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE)
der Ludwig-Maximilians-Universität München

Dissertation
zum Erwerb des Doctor of Philosophy (Ph.D.)
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

Integrating evidence, norms and values into guideline and other health decisions: the WHO-INTEGRATE evidence to decision framework version 1.0

vorgelegt von:

Jan M Stratil

aus:

Herrenberg, Deutschland

Jahr:

2021

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

First evaluator (1. TAC member): Prof. Dr. Eva A Rehfuess
Second evaluator (2. TAC member): Prof. Dr. Manfred Wildner
Third evaluator: Prof. Dr. Katja Radon
Fourth evaluator: PD Dr. Alberer

Dean: **Prof. Dr. med. Thomas Gudermann**

date of the defense:

01.10.2021

Affidavit

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Promotionsbüro
Medizinische Fakultät

MMRS

Affidavit

Stratil, Jan Marcel

Surname, first name

Elisabeth-Winterhalter-Weg 6

Street

81541, Munich, Germany

Zip code, town, country

I hereby declare, that the submitted thesis entitled:

Integrating evidence, norms and values into guideline and other health decisions: the WHO-INTEGRATE evidence to decision framework version 1.0

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Berlin, 27.01.2022

place, date

Jan M Stratil

Signature doctoral candidate

Confirmation of congruency

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Promotionsbüro
Medizinische Fakultät

MMRS

Confirmation of congruency between printed and electronic version of the doctoral thesis

Stratil, Jan Marcel

Surname, first name

Elisabeth-Winterhalter-Weg 6

Street

81541, Munich, Germany

Zip code, town, country

I hereby declare, that the submitted thesis entitled:

Integrating evidence, norms and values into guideline and other health decisions: the WHO-INTEGRATE evidence to decision framework version 1.0

is congruent with the printed version both in content and format.

Berlin, 27.01.2022

place, date

Jan M Stratil

Signature doctoral candidate

Table of content

1. List of abbreviations

A4R	Accountability for Reasonableness
CERQual	Confidence in the Evidence from Reviews of Qualitative Research
EIDM	Evidence-Informed Decision-Making
EtD framework,	evidence-to-decision framework
EUnetHTA	EUropean network for Health Technology Assessment
EVIDEM	Evidence and Value: Impact on DEcisionMaking
FGD	Focus Group Discussion
GDG	Guideline Development Group
GIS	geographical information system
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HIV	Human Immunodeficiency Virus
HFCS	High Fructose Corn Syrup
HTA	Health Technology Assessment
KII	Key-Informant Interview
MCDA	Multi-Criteria Decision Analysis
Q-SEA	Quality Standards for Ethics Analyses in health technology assessment
RCT	Randomized Controlled Trial
SDG	Sustainable Development Goal
WHO	World Health Organization
WHO-INTEGRATE	World Health Organization INTEGRATe Evidence
WICID	WHO-INTEGRATE COVID-19 adaptation

2. List of publications for the dissertation

Paper I: Rehfuess EA*, **Stratil JM***, Scheel IB, et al (2019): The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. BMJ Global Health 2019;4:e000844. doi: 10.1136/bmjgh-2018-000844 [1]
**Both authors contributed equally to this publication*

Paper II: **Stratil JM**, Paudel D, Setty KE, et al (2020): Advancing the WHO-INTEGRATE framework as a tool for evidence-informed, deliberative decision-making processes: exploring the views of developers and users of who guidelines. Int J Health Policy Manag. doi: 10.34172/ijhpm.2020.193[2]

Paper III: Stratil JM, Baltussen R, Scheel IB, *et al.* (2020) Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making. *Cost Eff Resour Alloc* **18,** 8. doi: 10.1186/s12962-020-0203-6[3]

3. Contribution of the PhD candidate to the publications

3.1 Contribution to paper I: The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective

Paper I is a shared first authorship between the PhD candidate (JMS) and Prof. Eva Rehfuess (EAR), as both authors contributed equally to the publication and to the overall research project which lead to the development of the WHO-INTEGRATE framework presented in it.

Prof. Rehfuess conceived and coordinated the overall research project, undertook the work of what is referred to in the publication as step 2a (assessing complexity and individual vs. population- and system level features as part of the assessment of the framework value and undertaking the work) and step 3 (conducting a literature review and expert consultation to suggest methods of evidence gathering and synthesis for the criteria in the form of the "evidence toolbox"), and wrote most parts of the manuscript.

The PhD candidate provided substantial input in the conception of the overall research project.
JMS undertook most of the work in what is referred to in the publication as step 1: development of the framework. This includes the literature and document review of key WHO documents reflecting WHO norms and values, of other evidence-to-decision frameworks, and of public health ethics framework (referred to as Step 1a in the publication). This furthermore includes conducting most of the work on the overview of systematic reviews of criteria, which is referred to as step 1b in the publication and further detailed in the contribution to paper III. In Step 1, EAR and one other coauthor duplicated relevant elements in the process and JMS, EAR, and one other coauthor analyzed the findings regarding the development of the framework. Step 1 resulted in a preliminary WHO-INTEGRATE framework with criteria, sub-criteria, and example questions.
JMS designed what is in the publication referred to as step 2b: Key informant interviews with developers of WHO guidelines and 2c: Focus group discussions with users of WHO guidelines in Brazil, Germany, Nepal, and Uganda with input from EAR. JMS collected the data and jointly analyzed them with another co-author. This is further detailed in the contribution to paper II.
JMS and all authors discussed different versions of the framework, revising criteria, definitions, sub-critria and example questions in an iterative manner.
JMS drafted individual sections of the manuscript and all authors critically reviewed different versions of the manuscript, suggested revisions and approved the version to be published.

3.2 Contribution to paper II: Advancing the WHO-INTEGRATE Framework as a Tool for Evidence-Informed, Deliberative Decision-Making Processes: Exploring the Views of Developers and Users of WHO Guidelines

JMS and EAR designed and directed the project with contribution from another coauthor.

For the part on the key-informant interviews (KIIs): JMS developed the interview guide with input from EAR and another coauthor. JMS and EAR recruited the participants with the help from WHO staff members. JMS conducted the KIIs. KII-transcripts were analyzed by JMS with contribution from a coauthor and two research assistants.

For the part on the focus group discussions (FGDs): The coauthors in the role of local researchers suggested topics and recruited participants for the FGDs in the four countries as well as prepared and facilitated focus group discussions. JMS drafted a discussion guide, which was revised and adapted by the local researchers. The FGD in Germany was conducted by JMS while the other interviews were conducted by the local researchers with JMS being available in presence or via video call for questions. The local researchers transcribed, translated, and contextualized the respective FGD-audio records. JMS analyzed the FGD transcripts with contribution from two research assistants, as well as input from all authors on the findings.

JMS took the lead in writing the manuscript with input from all authors. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

3.3 Contribution to paper III: Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making (Appendix)

JMS with support from a coauthor conducted the literature searches as well as screening of the identified publications.

JMS conducted the data extraction and analysis with support from a research assistant. JMS developed the category system and approach to organize the criteria with supported from EAR and the other coauthors.

JMS developed different versions of the table of criteria and the content with the different versions being discussed with EAR and the other coauthors. JMS wrote the manuscript with input from EAR.

All authors critically reviewed different versions of the manuscript, suggested revisions and approved the version to be published.

4. Introductory summary

4.1 Introduction

4.1.1 Evidence-informed public health decision making

Evidence-informed decision-making in the context of public health and health policy describes processes of integrating research evidence with public health expertise, community preferences, and additional factors in program planning and policy decision making to improve the health of populations[4-7]. It is often applied by a group of experts and advisory groups, who systematically gather, distil, and disseminate the best available knowledge form research, context, and experience to support evidence-informed decisions or provide evidence-informed recommendations to health policy and public health decision makers[4-6].

An example for the practical application of evidence-informed decision making in public health are public health guidelines: Here, a group of content and methods experts as well as stakeholder representatives collect and synthesize evidence (e.g., of the effect of an intervention), which is balanced with other relevant factors (e.g., cost) to develop a recommendation intended to support informed decisions on matters of public health and health policy[2 8 9]. Other applications of evidence-informed public health decision-making can include conducting health technology assessments[10-13], setting (research) priorities[3 14 15], or advising policy makers[3 16 17].

4.1.1.1 Challenges in making evidence-informed public health decisions

While sharing several similarities with evidence-informed decision-making in the context of clinical medicine (e.g., in the development of clinical practice guidelines), additional challenges arise when making evidence-informed decisions in the context of public health and health policy[18-22]. Which is therefore often a complex task[18-22]:

First, the demand for clear and actionable evidence-informed recommendations is often at odds with the lack of high quality and reliable scientific evidence or evidence of questionable transferability and applicability to the context at hand[23 24]. This is a particular challenge for public health and health policy interventions which can be described as complex interventions, e.g. due to them consisting of multiple active and interacting components, targeting a range of different organizational levels to change outcomes along often long and non-linear causal pathways[25]. Furthermore, rather than producing effects independent of the context they are implemented in, these measures are often best characterized as "events in a systems"[26] where the effects of a measure arise from components of the intervention interacting with components of a complex (social) system[19 20 26 27]. As a result, evidence of effectiveness of public health and health policy interventions often is no clear guide for the direction of a decisions or recommendations. And for this reason, additional factors beyond evidence of effectiveness tend to be of greater relative importance in the decision-making process.

A second challenge is that even if reliable and applicable evidence on the effectiveness of a public health intervention were available, it would be insufficient to make sound recommendations[1-3], as evidence-informed public health and health policy decision-making is a deeply value-laden process[4 22 28 29]. This process requires the balancing of numerous and often conflicting normative and technical considerations beyond effectiveness[1-3 30-32], ranging from ethical considerations[33-39] to utility assessments or legal conformity[15 30 32 40-42]. For example, assuming that strong and applicable evidence on the effectiveness of school closures as an infection control measure in the SARS-CoV-2 pandemic was available. This beneficial

impact on SARS-CoV-2 related outcomes would still need to be balanced against the adverse effects on the mental health of children and parents, the loss of education, the economic losses resulting from the workforce engaging in child care, the risk of increased social inequity over the short and long term, and the fundamental right to education[22].

4.1.2 The role of selecting and weighting decision criteria

As a result, factors beyond effectiveness, which are operationalized as decision criteria, are of particular importance and have to be taken into account: the criteria which are considered in a decision-making process, as well as how these factors are weighted and balanced against each other, will all affect the final decision[1-3]. For example, where no criteria beyond the direct implications of school closures for the course of the SARS-CoV-2 pandemic are considered, this is likely to lead to a recommendation favoring school closures as an anti-pandemic measure. In contrast, when additionally considering criteria such as educational outcomes, workplace participation, and child mental health and assigning a large weight to these in the decision-making process, this is likely to influence the decision in the opposite direction.

The importance ascribed to criteria in a given decision-making process will vary widely, depending on e.g., the type of measure, decision-making process, or the broader physical and societal context, including value systems[43 44]. Within and across societies, value systems of stakeholders can vary considerably[1-3]. And as there are multiple reasonable positions regarding which values should guide a decision-making process, reasonable disagreement about the *right* or *best* factors to be considered and – as a result – the, *right* or *best* recommendation or decision is likely in pluralist societies[1-3 45-48].

4.1.3 Selecting the "right" criteria and implementing fairness in decision-making

Various approaches have been proposed to help address the challenge of selecting and weighting criteria in evidence-informed decision-making[4 49], including among others[33 34 36 50 51] the Accountability for Reasonableness (A4R) framework by Daniels and Sabin[45 46 48 52]. These approaches tend to focus on a set of decision-making principles, such as transparency and openness of the process, participation and representation of relevant stakeholders, or allowing for the possibility of appeals and revision of the decisions[1 2]. Following these procedural principles can increase the perceived legitimacy and acceptability of a recommendation or decision[1 2 48 53 54].

One key principle in the A4R framework[45], is the condition of relevance. It states that recommendations or decisions must be based on reasons and principles (including the evidence informing them) that fair-minded parties can agree on to be relevant to balancing the diverse needs of affected stakeholders under the constraints of limited resources[1 2 45]. To meet the principle of relevance in identifying, selecting, and weighting the decision-making criteria serving as the foundation in a decision-making process, it is considered ideal to involve representatives of all stakeholder groups in the decision-making process[1 2 4 9 10 22 55 56]. However, this can be at odds with the great number of individual or groups affected by a public health or health policy decision and difficult to achieve when decisions need to be made within a limited time frame and under the constraint of finite resources available for a participatory process[1 2]. However, limiting the range of participating stakeholders risks that relevant factors are overlooked and, as a result, the decision can lose legitimacy and the support of the public[1 2].

4.1.4 Evidence-to-Decision (EtD) frameworks as a second-best option

An additional approach to addressing the challenge of selecting criteria of relevance under the constraints of real-world public health and health policy decision-making are so called Evidence-to-Decision (EtD) frameworks[1 2]. EtD frameworks tend to comprise sets of decision-making criteria as well as procedural guidance on how to apply and adapt them in a given decision-making process[1 2 57 58]. Their purpose is to ensure that factors of relevance are considered in the decision-making criteria, that these criteria are informed by the best available evidence, and that the rationale underlying the decision is made transparent and explicit[1 2 59]. While neither intended nor able to replace involving stakeholders in the process of making decisions or developing recommendations within expert groups, EtD frameworks can help ensure that all factors of relevance are considered – in the form of decision criteria – even if not all voices of affected stakeholders could be heard[1 2 22 57 58]. Through increasing relevance and transparency of the decision-making process, well developed and applied EtD frameworks could thereby serve as a tool to increase the relevance and acceptability of public health and health policy decisions[1 3 22].

4.1.5 Developing an EtD framework for WHO guideline development

Against this background, the World Health Organization (WHO) initiated a project intended to strengthen its processes and methods for developing guidelines on complex health interventions and interventions delivered in complex systems[8].

WHO guidelines are systematically developed by group of content and methods experts as well as representatives of national governments and stakeholder groups, the so-called guideline development group (GDG)[2 8 9]. The GDG systematically develops research and guideline questions, gathers, appraises and synthesis relevant evidence, and balances the evidence with other relevant factors in the so-called evidence-to-decision process to develop a recommendation intended to support policy makers and public health practitioners in making informed decisions on matters of public health and health policy[2 8 9].
WHO guidelines are of particular importance for policymakers and program managers in low- and middle-income countries, as these often have limited resources for conducting comprehensive processes of evidence gathering and analysis by themselves[2].

The project, which was initiated in 2016, explored different aspects of complexity in guideline development. Among others the implications of a complexity perspective for systematic reviews and guideline development in health decision making[19], how to integrate considerations of context from a complexity perspective[60], and methodological challenges when synthesizing quantitative or qualitative data for complex interventions[61-63].
One part of the project was to address the previous derived challenge in evidence-based decision-making in the process of developing recommendations: balancing evidence of effectiveness, which is often uncertain or indirect, with a multiplicity of other factors to provide recommendations in a timely manner and often under considerable resource constraints[1-3].

4.2 Research question

The above leads to the following research question: What criteria should be considered in an evidence-to-decision framework that has a normative foundation in WHO norms and values, that is reflective of real-world public health and health policy decision making, and that is suitable for reflecting complex interventions implemented in complex systems[1 8]?

The objectives of the doctoral thesis, therefore, were to

- develop an initial framework that is firmly rooted in WHO norms and values and reflective of the changing global public health landscape[1], based on the analysis of key documents on WHO norms and values as well as key public health ethics frameworks.
- create a set of criteria used in real-world public health and health policy decision-making through an overview of systematic reviews[3] and matching them against the initial EtD framework
- assess comprehensiveness, relevance, and usefulness of the resulting framework – the WHO-INTEGRATE framework – and its criteria among those developing WHO guidelines and identify opportunities for revising the framework, and
- assess the comprehensiveness, relevance, and usefulness of the WHO-INTEGRATE framework and its criteria among those using WHO guidelines and identify opportunities for revising the framework.

The focus of this doctoral thesis was the development of a framework providing substantive decision-making criteria (e.g., “what should the decision be based on?”, such as acceptability or health impact). Procedural criteria or aspects (e.g., “how should the decision be made?”, such as transparent or participatory processes) were beyond the scope of this project[1-3].

4.3 Overview of the Doctoral Thesis

Part I of this thesis will outline the development process of the WHO-INTEGRATE framework with part II briefly describing the framework. More details was provided in the respective publications[1-3].

4.3.1 Part I. Development of the WHO-INTEGRATE framework

The approach for developing the framework is illustrated in **Figure 1**.

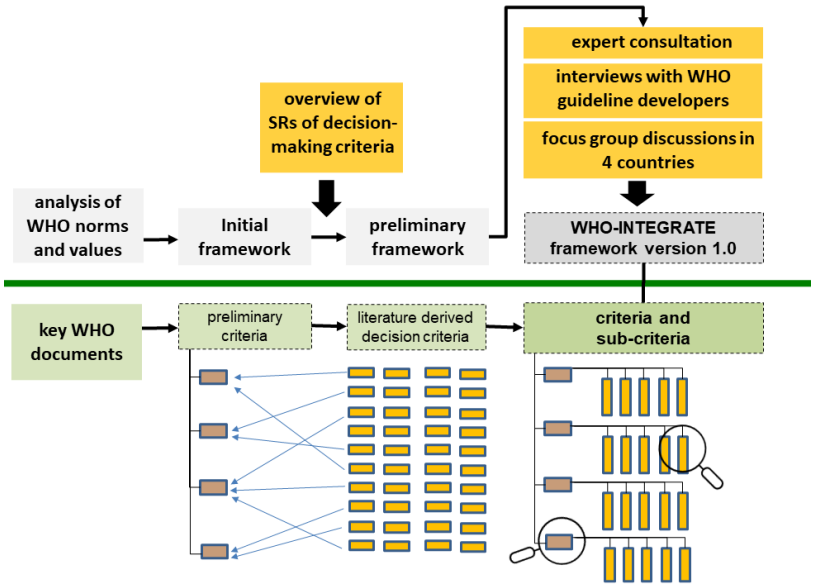

***Figure 1**: Outline of the steps of developing the WHO-INTEGRATE framework.*
WHO: World Health Organization; SRs: Systematic Reviews.

4.3.1.1 Development of the initial framework

To increase legitimacy and acceptability of the framework, as well to ensure comprehensiveness of the criteria included in it, we intended to base the framework on a firm theoretical or conceptual normative foundation: normative concepts and principles rooted in WHO norms and values[1]. This foundation was chosen, as these norms and values are rooted in the universally acknowledged concept of human rights and as these were legitimized globally as a fundamental canon of values due to having been agreed upon and approved by all 194 member states of the WHO[1 64].

In order to identify normative concepts and principles of relevance, which are rooted in WHO norms and values, we conducted a grey literature search of key WHO documents[1]. Furthermore, in order to ensure comprehensiveness, we conducted a scoping review of the public health ethics literature to identify key public health ethics frameworks and assess the principles defined within them[34 36-39 50 51 65-69].

Methods box 1: Approach to explore WHO norms and values

As a starting point served the Constitution of the WHO[64] as well as Chapter *"Incorporating equity, human rights, gender and social determinants into guidelines"* of the *WHO guideline handbook for guideline development*[1 70]. Among others, these documents emphasized human rights (in particular the right to health), equity considerations, and the principle of non-discrimination[1 64 70]. They furthermore emphasized the importance of social determinants of health and of well-functioning health systems as underlying preconditions for human health and well-being [1 64 70]. Therefore, additional documents were retrieved and analyzed which focused on these topics and aspects[1 71-77], for example the comment of the Committee on Economic Social and Cultural Rights on the right to health[1 71]. We furthermore reviewed the sustainable development goals (SDGs) and related documents[78]. This was

done to account for the current activities in the WHO which is mainstreaming the SDGs throughout the work of the organization[79-81], the likely impact of SDGs on national and sub-national level political decision making, as well as the legitimacy the SDGs as a normative document received through general agreement by member states of the united nations[1 78].

Based on principles and concepts derived from these sources, an initial version of the WHO-INTEGRATE framework was developed in an iterative process through multiple rounds of exploring these principles and concepts, formulating criteria based on them, grouping and/or merging criteria where possible, and assessing the framework for overlap and redundancies. This process was conducted in close collaboration with all co-authors and at several stages WHO staff who were members of the WHO Guidelines Review Committee or experts on selected principles and concepts were consulted[1]. This is illustrated in the first and second column of figure 1.

4.3.1.2 Advancing the initial framework with real-world decision-making criteria

Next, we sought to expand the initial framework based on real world decision-making criteria to ensure comprehensiveness of criteria and sub-criteria (e.g., as the WHO norms and values documents provided only a limited basis to develop criteria or sub-criteria related to feasibility, resources, and cost). We defined real-world decision-making criteria as criteria which are being used by governmental or non-governmental organizations on a supranational, national, or program level in decision-making process regarding public health or health policy interventions. To identify those, we conducted an overview of systematic reviews (see methods box 2 for details).

Methods box 2: Methods for the overview of reviews on real-world decision-making criteria

We used a combination of systematic database searches followed by extensive citation searches of included publications, to identify systematic reviews which provided sets (≥3) of decision making criteria[3]. The eligibility was assessed independently by two researchers and the quality of included reviews were assessed using the tool AMSTAR 2[82]. Key characteristics of the included systematic reviews, as well as the criteria included in them, were extracted, de-duplicated, and sorted into first-level (i.e. criteria), second-level (i.e. sub-criteria), third-level (i.e. decision aspects), and fourth-level (i.e. sub-aspects) categories[3]. The first-level categories were based on the initial WHO-INTEGRATE framework[1 3], the second-, third-, and fourth-level categories were developed inductively in an iterative process[3].

Out of the 4,448 unique records identified through database searches and the 88 additional references identified through citation searches, we included 36 systematic reviews[3]. The included publications focused on a range of topic in the field of public health and health policy decision-making, including priority setting (n=17), health technology assessments (n=9), criteria used in multi-criteria decision analysis (n=6), or investment or disinvestment decisions in the field of public health (n=3)[3]. Across the reviews, we identified more than 2,400 individual decision-making criteria, which were subsumed into 8 criteria (first-level categories), 45 sub-criteria (second-level) and 200 decision aspects (third level)[3]. The tables listing and categorizing all criteria are provided in the publication Stratil et al. 2020 published in the journal of Cost Effectiveness and Resource Allocation[3].

Based on this set of real-world decision-making criteria, we expanded the initial framework[3]. This was done by matching the identified criteria (first level categories) and sub-criteria (second-level categories) against the initial framework, while keeping the subsumed decision aspects and sub-aspects (third and fourth level categories) in mind[1]. This process did not lead to the identification of additional criteria beyond the 6+1 criteria in the initial framework[1]. However, the process led to many additional sub-criteria[1]. This is illustrated in the second and third column of figure 1.

In a next step, definitions and descriptions for each criterion as well as sample questions for each sub-criterion of the first preliminary version of the EtD framework were developed[1]. This was done by drawing on the documents used as sources for the WHO norms and values, existing EtD frameworks, the overview of systematic reviews of decision-making criteria, as well as other documents of relevance for a given criterion or the principles underlying it (e.g. Scott et al. 2017[83] for the definition of a*cceptability*)[1 3].

4.3.1.3 Finalization of the WHO-INTEGRATE framework version 1.0 based on feedback from key-informants and friendly peer review

A next step in the development was focused on examining the usefulness, understandability, comprehensiveness, and relevance of the WHO-INTEGRATE framework and its criteria among those intended to apply the framework[1 2]. We therefore conducted key-informant interviews (KIIs) with experts involved in developing WHO guidelines[1 2] (see methods box 3 for details). For this process, we selected a purposive sample of three WHO guidelines which were identified in consultation with members of the Secretariat of the WHO Guideline Review Committee (GRC)[1 2]. The aim was to cover distinct types of public health and health policy interventions as well as to capture positive as well as challenging experiences with applying the EtD framework previously suggested by the WHO[9], the GRADE EtD framework[1 2 9].

Methods box 3: Approach for conducting KIIs to assess the WHO-INTEGRATE framework

The semi-structured KIIs were conducted either at the WHO Headquarters in Geneva in the form of face-to-face interviews or by telephone/video by the PhD candidate[1 2]. The KIIs were conducted with content or methods experts who were involved in guideline development processes in the role of methodologist supporting the process, coordinating WHO staff, or GDG chair[1 2]. These experts were selected based on their contribution to the following three guidelines: the WHO guideline for emergency risk communication policy and practice[84], the consolidated guideline on sexual and reproductive health and rights of women living with HIV[85], and the WHO recommendations on antenatal care for a positive pregnancy experience[1 2 86]. The pre-tested interview guide focused on challenges and general experiences with the evidence-to-decision process in the respective guideline, followed by a second part of the interview aimed at exploring the perspectives of the interviewees on a preliminary version of the WHO-INTEGRATE framework they were presented with[1 2]. A team of two researchers analyzed the pseudonymized transcripts through qualitative content analysis[1 2 87] using the software MAXQDA12 (VERBI Software GmbH, Berlin)[1 2]. The analysis followed a mixed deductive and inductive approach[1 2].

In total, nine KIIs with a median duration of 62 minutes were conducted with methodologists (n=3), coordinating WHO staff (n=2), and GDG chairpersons (n=4)[1 2]. In general, the framework as a whole, its underlying conceptualization in WHO norms and values and, its comprehensive nature received positive feedback[1 2]. As did the level of detail regarding criteria and sub-criteria[1 2]. Several specific remarks were made, among others regarding (i) the order and hierarchy of criteria, (ii) potential redundancies and overlap of criteria and sub-criteria, and (iii) the terminology, definitions, and descriptions of criteria and sub-criteria[1 2]. Some interviewees expressed concerns regarding the implications of applying the framework in everyday guideline development processes, given the additional burden placed on the guideline development group[1 2]. A comprehensive overview of the findings from this qualitative study is presented in the publication by Stratil et al (2020), published in the International Journal of Health Policy and Management[2].

In response to the concerns and suggestions by the interviewees, several modifications were made to the framework to improve clarity and reduce overlap and redundancies[1]. These included changes to the terminology and definition of several criteria and sub-criteria as well as changes to the order and grouping of sub-criteria[1].

Throughout the development process, we sought and received feedback from several content and methods experts from within and outside of the WHO, to integrate their expertise and experience with working on complex public health and health policy interventions[1] (see acknowledgement section). This included feedback on the framework, as well as on specific components and criteria of the framework, which was integrated through the course of the framework development process[1]. The framework, resulting from this process, is described in the section *Part II. Overview of the WHO-INTEGRATE framework*.

4.3.1.4 Assessing the WHO-INTEGRATE framework version 1.0 through focus group discussions with potential users of WHO guidelines

In a final step, we sought to capture the perspective of stakeholders intended to use guidelines and other products having applied the framework[1 2]. To do so, we conducted focus group discussions (FGDs) with national and sub-national level stakeholders working in the field of public health and health policy decision makers[1 2]. With the FGDs, we intended to capture the perspective of these stakeholders regarding the comprehensiveness, relevance, and usefulness of the WHO-INTEGRATE framework version 1.0, and identify potential for improvement towards a version 2.0[1 2].

To capture the diversity of views, we conducted FGDs with decision makers from four countries and continents (i.e., Brazil, Germany, Nepal, and Uganda)[1 2]. In the FGDs, we sought to maximize heterogeneity among decision-making levels (i.e., national level, regional level), countries (i.e. country income group, world region), and topics of discussion (i.e. type of health problem and intervention)[1 2]. The approach taken is outlined in methods box 4.

Methods box 4: Approach for conducting FGDs with potential users of WHO guidelines

The FGDs were conducted in close collaboration between local researchers and the developers of the WHO-INTEGRATE framework[2]. The local researchers suggested a public health or health policy topic of current importance, reached out to a purposive sample of local experts, and conducted the semi-structured FGDs based on an interview guide jointly developed by the team who developed the WHO-INTEGRATE framework and the local researchers[2].

All FGDs were conceived in two phases: In a first phase, the local content experts discussed criteria of relevance for decision making – which were later assessed against criteria and sub-criteria of the WHO-INTEGRATE framework[2]. In a second phase of the FGDs, the participating experts were presented either with the WHO-INTEGRATE framework version 1.0 or a preliminary version[2]. They were then asked to review whether the framework covered all decision-making criteria discussed in the first phase of the FGD, whether criteria or considerations regarded as relevant for decision-making were missing, whether there is a need to revise terminology, and whether there might be other suggestions for improvement[2].

Next, the audio-records of the discussions were transcribed, pseudonymized, and reviewed after translation the local researchers[2]. The transcripts were translated into English in the case of the FGD from Nepal and Brazil, while the transcript of the FGD conducted in Germany were analyzed in German[2]. A team of two researchers (including the PhD candidate) analyzed the transcripts using the software MAXQDA 12 (VERBI Software GmbH, Berlin)[2]. The analysis was conducted through qualitative content analysis in a mixed inductive and deductive approach, following the approach defined by Mayring[87]. This included the coding of text passages of phase I of the FGDs against the criteria and sub-criteria of the WHO-INTEGRATE framework, to assess whether they were covered by the framework. The findings were provided to the local researchers for review, clarification, feedback, and discussion[2].

The four FGDs were conducted with 7 to 17 participants and had a duration of between 95 and 130 minutes[2]. The analysis of the first phase of the FGDs showed, that all criteria and sub-criteria of the WHO-INTEGRATE framework version 1.0 were mentioned as relevant or were discussed in at least one of the four FGDs[2]. This aligns with general remarks on the framework, where the comprehensiveness of the framework, as well as the inclusion of specific criteria (e.g., human rights, acceptability) were applauded[2].

While the framework was well received as a whole, several suggestions were either explicitly made by the FGD participants or derived from their statements by the research team (e.g., a misunderstanding of criteria indicated a need to revise descriptions)[2]. These included suggestions regarding terminology, missing aspects, grouping of criteria and sub-criteria, as well as delineation of criteria and sub-criteria[2]. Overall, the FGDs yielded a number of valuable suggestions to serve as a point of departure for further advancing the framework into a version 2.0[2]. A more detailed description of the study, their findings, and the conclusions drawn from them are presented in the publication by Stratil et al (2020), published in the International Journal of Health Policy and Management[2].

4.3.2 Part II. Overview of the WHO-INTEGRATE framework

This development process led to the version 1.0 of the WHO-INTEGRATE framework[1]. The framework consists of six substantive criteria, which are: *Balance of health benefits and harms, Human rights and sociocultural acceptability, Health equity, equality and non-discrimination, Societal implications, Financial and economic considerations,* and *Feasibility and health system considerations*[1 2]. According to the framework, all six substantive criteria are relevant to public health and health policy decision-making and should be reflected in the decision-making process[1 2]. The frameworks includes a definition for each criterion, and offers details and explanations on how the criterion can be applied and considered in formulating the recommendations[1].

Each of the substantive criteria comprises several sub-criteria, 26 in total[1]. Likely, not all sub-criteria may be of relevance for a given topic[1 2]. For example, the sub-criteria for the criterion *Balance of health benefits and harms are: (i) efficacy or effectiveness on health of individuals, (ii) effectiveness or impact on health of population, (iii) patients'/beneficiaries' values in relation to health outcomes, (iv) safety-risk-profile of intervention, (v) broader positive or negative health-related impacts*[1]. A set of guiding questions for each sub-criterion is intended to support users of the framework in assessing their relevance for a given decision-making process and in adapting it accordingly[1].

Depending on the perspective of the guideline or decision-making process, the criteria can be reflected from an individual, population, or system-level perspective[1].

The criterion *quality of evidence*, is a meta-criterion, which is intended to be considered across all substantive criteria for the respective body of evidence[1]. Depending on the criterion and decision-making process, evidence may comprise qualitative, quantitative, or mixed-methods data sources as well as forms of normative analyses[1]. As a result, a variety of approaches in identifying, developing and assessing the quality of the respective body of evidence is required[1]. Therefore, a methods "tool box" was developed which intends to support the users regarding the integration of the appropriate evidence for each criterion[1].

The framework is outlined in figure 2. A comprehensive description of the framework, the criteria and sub-criteria, as well as related content (e.g., the definitions), are provided in the publication by Rehfuess and Stratil et al (2019), published in the journal BMJ Global Health[1].

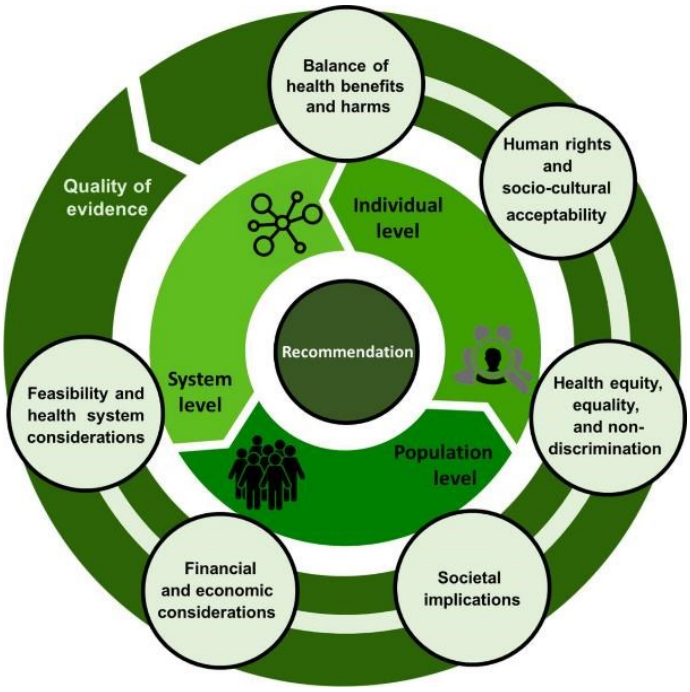

Figure 2: *Overview of the WHO-INTEGRATE (INTEGRATe Evidence) framework version 1.0. Retrieved from the publication by Rehfuess and Stratil et al.*[1]

4.4 Contribution of the doctoral thesis

The research conducted as part of this doctoral thesis led to the development of a new Evidence-to-Decision framework - the WHO-INTEGRATE framework version 1.0[1] – which has a firm conceptual and normative foundation rooted in WHO norms and values, that was further expanded based on a comprehensive overview of reviews on decision criteria used in real-world public health and health policy decision-making[3], and which was repeatedly reviewed by developers and users of WHO guidelines and revised based on their feedback[1 2]. Although developed with WHO guidelines in mind, the framework was from the start intended as a tool which could equally be used for other public health and health policy decision making processes at the global, national, and sub-national level[1].

In comparison with other EtD frameworks (e.g. the GRADE EtD framework as the one which was primarily used in WHO guidelines[9]), the WHO-INTEGRATE framework has a number of unique characteristics: (i) it offers a firm conceptual foundation for the criteria included in it, (ii) it was explicitly developed to be applicable to complex interventions in complex systems, (iii) it reflects a broad understanding of human health and its social, ecological, and economic determinants and therefore emphasizes the interconnectedness between health and other sectors, and (iv) it emphasizes the importance of considering criteria beyond the effectiveness of an intervention on health outcomes or cost.

To explore the added value of the framework to potential users of the framework and users of products whose development made use of the WHO-INTEGRATE framework, empirical qualitative research was conducted as part of this doctoral thesis. Overall, the framework, the underlying conceptualization of its criteria and sub-criteria, its comprehensive nature, as well as the detailed criteria and sub-criteria received positive feedback from experts involved in the development of WHO guidelines and potential national and sub-national level decision-makers across four continents.

Other indications of the added value of the framework are the more than 60 citations in the two years since first publication, endorsement by institutions such as the WHO guideline review committee[8 88], or

the practical usage of the framework in guidelines. Among others, the authors of a recent viewpoint article in the journal The Lancet Global Health applauded the WHO-INTEGRATE framework as an approach to develop guidelines which are more suitable to the needs of stakeholders in low resource settings[89]. Since the publication, the framework was used in a number of WHO guidelines, among them the Guidelines on Sanitation and Health[90], the interim guidance of COVID-19 mitigation in the aviation sector[91], or the WHO Guideline on parenting to prevent child maltreatment and promote positive development in children aged 0-17 years[92]. Furthermore, the WHO-INTEGRATE framework is referenced in an online expansion of the WHO Guideline handbook for guideline recommendations on how to address complex interventions in WHO guidelines[88 93].

The SARS-CoV-2 pandemic has highlighted the importance of taking a broad range of considerations in public health and health policy decision making into account[22]. This led, among others, to an adaption of the WHO-INTEGRATE framework to COVID-19 decision making in the form of the WICID framework (WHO-INTEGRATE COVID-19 adaptation), which was published in the journal BMJ Global Health[22]. Furthermore, it led to the use of the WHO-INTEGRATE framework (with its additional specifications for COVID) in the recent AWMF S3 guideline: "Measures for prevention and control of SARS-CoV-2 transmission in schools"[94], which was well received in the German media and by public health decision makers[95].

In addition, the importance of weighting health of public health and health policy interventions against broader environmental, economic, and social impacts and their potential interactions was acknowledged in a recent publication by the GRADE Public Health Group[96], referencing the WHO-INTEGRATE framework. The acknowledgement of the added value of the WHO-INTEGRATE framework led to the implementation of the WHO-INTEGRATE framework as a template in the software GRADEpro (https://gradepro.or).

From the start, the WHO-INTEGRATE framework version 1.0 was intended to be further developed. In working towards a version 2.0, the feedback collected through the focus group discussions with national and sub-national decision makers[2] and future research based on the experiences of those applying the framework (e.g., in the mentioned WHO guidelines) will provide a valuable basis for revisions.
Further, as pointed out in the key-informant interviews with WHO-guideline developers, improved guidance on how to apply the WHO-INTEGRATE is needed[2]. In particular, guidance which takes the resource constraints into account, under which real world decision-making processes are conducted[2]. To address the need for guidance and further facilitate the use of the WHO-INTEGRATE framework, the author had been commissioned by the WHO to write an additional chapter for the WHO handbook for guideline development[9] dedicated to providing guidance on the application of the WHO-INTEGRATE framework.

EtD frameworks such as the WHO-INTEGRATE framework and its criteria can be a valuable tool in supporting decision makers (which includes scientific expert committees and advisory groups) in identifying relevant criteria for public health and health policy decisions. However, the use of an EtD framework, does not supersede the need to address the value-laden nature of public health and health policy decision-making processes[1 2]. Efforts should be made to achieve a process that is characterized by the principles of openness, participation, representation, as well as the right to revisions and appeal[4 10 45 47].

Public health and health policy decision-making is a challenging and complex undertaking. However, this doctoral thesis aims to provide those involved in these decisions with a valuable tool to address the challenges. And thereby, hopefully, contributes to real-world public health and health policy decisions which improve health for all.

5. Paper I: The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective

Rehfuess EA*, Stratil JM*, Scheel IB, et al (2019): The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. BMJ Global Health 2019;4:e000844. doi: 10.1136/bmjgh-2018-000844 64

*Both authors contributed equally to this publication

Research

BMJ Global Health

The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective

Eva A Rehfuess,[1] Jan M Stratil,[1] Inger B Scheel,[2] Anayda Portela,[3] Susan L Norris,[4] Rob Baltussen[5]

To cite: Rehfuess EA, Stratil JM, Scheel IB, *et al*. The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. *BMJ Glob Health* 2019;**4**:e000844. doi:10.1136/bmjgh-2018-000844

Handling editor Seye Abimbola

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjgh-2018-000844).

EAR and JMS contributed equally.

Received 20 March 2018
Revised 5 July 2018
Accepted 20 July 2018

For numbered affiliations see end of article.

Correspondence to
Dr Eva A Rehfuess;
rehfuess@ibe.med.uni-muenchen.de

ABSTRACT

Introduction Evidence-to-decision (EtD) frameworks intend to ensure that all criteria of relevance to a health decision are systematically considered. This paper, part of a series commissioned by the WHO, reports on the development of an EtD framework that is rooted in WHO norms and values, reflective of the changing global health landscape, and suitable for a range of interventions and complexity features. We also sought to assess the value of this framework to decision-makers at global and national levels, and to facilitate uptake through suggestions on how to prioritise criteria and methods to collect evidence.

Methods In an iterative, principles-based approach, we developed the framework structure from WHO norms and values. Preliminary criteria were derived from key documents and supplemented with comprehensive subcriteria obtained through an overview of systematic reviews of criteria employed in health decision-making. We assessed to what extent the framework can accommodate features of complexity, and conducted key informant interviews among WHO guideline developers. Suggestions on methods were drawn from the literature and expert consultation.

Results The new WHO-INTEGRATE (INTEGRATe Evidence) framework comprises six substantive criteria—*balance of health benefits and harms*, *human rights and sociocultural acceptability*, *health equity*, *equality and non-discrimination*, *societal implications*, *financial and economic considerations*, and *feasibility and health system considerations*—and the meta-criterion *quality of evidence*. It is intended to facilitate a structured process of reflection and discussion in a problem-specific and context-specific manner from the start of a guideline development or other health decision-making process. For each criterion, the framework offers a definition, subcriteria and example questions; it also suggests relevant primary research and evidence synthesis methods and approaches to assessing quality of evidence.

Conclusion The framework is deliberately labelled version 1.0. We expect further modifications based on focus group discussions in four countries, example applications and input across concerned disciplines.

Key questions

What is already known?

- Evidence-to-decision (EtD) frameworks help to ensure that all criteria of relevance in a given guideline development or other health decision-making process are considered in a systematic way.

What are the new findings?

- The WHO-INTEGRATE (INTEGRATe Evidence) framework is a new EtD framework that is rooted in the norms and values of the WHO, which are agreed on by all WHO Member States.
- The framework was developed to be applicable to all health interventions, although it is particularly well suited for decisions about population-level and system-level interventions at the global as well as national levels.
- The WHO-INTEGRATE framework offers structured definitions for each of the six substantive criteria as well as the meta-criterion *quality of evidence*; example questions and suggested methods are provided to facilitate uptake.

What do the new findings imply?

- As part of a more holistic approach, the framework is devised as a tool to facilitate structured reflection and discussion from the beginning to the completion of a guideline development or other health decision-making process; this entails prioritisation among criteria and subcriteria to ensure appropriate evidence collection and appraisal.

BACKGROUND

Health decision-making at local, national, regional and global levels is complex,[1–3] and can be influenced by a broad range of factors.[4–9] Their importance varies depending on the type of health decision and the decision-making context,[10 11] where context can relate to the institutional context (eg, Ministry of Health vs municipality), as well as the broader physical and social context, including epidemiological, geographical, sociocultural,

political and other aspects.[12] Health decision or evidence-to-decision (EtD) frameworks are intended to ensure that all important factors—in the form of decision criteria—are considered in a systematic and transparent way.[13–19] They provide a structured approach for guideline panels or other decision-making bodies to consider the available evidence and to make informed judgements about the advantages and drawbacks of a given health decision; this approach can comprise substantive criteria as well as procedural aspects. Health decision frameworks have been applied in a variety of decision-making contexts.[20 21]

Guidelines by the WHO provide recommendations for clinical practice, public health and health system strengthening, and are intended to support health decision-makers in prioritising among or selecting suitable clinical, public health or health system interventions. When formulating recommendations, WHO generally uses an EtD framework which encompasses eight criteria: *quality of evidence* (in relation to intervention benefits and harms), *values and preferences* (in relation to outcomes), *balance of benefits and harms*, *resource implications*, *priority of the problem*, *equity and human rights*, *acceptability*, and *feasibility* (table 10.1 of the *WHO Handbook for Guideline Development*).[22]

Chapter 10 in the *WHO Handbook for Guideline Development* [22] was written by one of the lead authors of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) and the GRADE EtD frameworks. The criteria in the current WHO EtD framework represent an advanced but—given their publication in 2014—not the final version of the GRADE EtD framework, which offers different versions for clinical recommendations from an individual or population perspective, coverage decisions, health system/public health decisions and recommendations about tests.[13 23] In a recent systematic review of frameworks concerned with generic health decision-making and resource allocation processes, health technology assessments, as well as very specific health decisions, the GRADE EtD framework emerged as the best fit-for-purpose framework (Stratil *et al*, forthcoming). In particular, this framework can be applied across diverse types of health decisions and was developed following an iterative and multipronged process, combining a literature review, brainstorming and feedback from stakeholders,[24] with application of the framework to examples and user-testing.[20 25]

However, a number of weaknesses were identified with the GRADE EtD frameworks (Stratil *et al*, forthcoming). First, the framework was developed using a pragmatic approach and lacks an explicit theoretical or conceptual basis. This makes it difficult to assess objectively whether the set of criteria is complete and organised in a meaningful way.

Second, while the frameworks are largely congruent with WHO norms and values, they do not sufficiently consider the central role of the social and economic determinants of health[26] and the implications of health sector or intersectoral interventions for society as a whole. This is particularly important given the significant shifts in the global health landscape and the objectives and values manifest in the Sustainable Development Goals (SDGs),[27] which are likely to shape health decision-making in the future.

A third concern is whether the decision criteria in the GRADE EtD framework are sufficiently complete and useful for decisions about complex interventions and/or the complex systems in which these are implemented, especially interventions aiming to bring about system-level changes.[28]

Fourth, the frameworks were originally developed in consultation with healthcare decision-makers in Europe, Canada and Africa, the majority of whom were physicians with significant clinical experience and research training.[24] As a result, the frameworks may not be entirely suitable to broader public health and health system decision-making contexts, particularly in low-income and middle-income countries of Asia and Latin America.

A final and important concern relates to consistency in the application of the GRADE EtD frameworks within the WHO guideline development processes. While there are exemplar guidelines, where the WHO EtD framework has been employed as intended,[29 30] many WHO guideline development groups focus extensively on the criterion *balance of benefits and harms* and apply the remaining criteria as a check box exercise rather than as a process that structures the development of guidelines from the start: from scoping a guideline and prioritising questions, to collecting, synthesising and appraising evidence, to formulating recommendations (SL Norris, 2017, personal communication). While there are many potential reasons for this, the current content and structure of the GRADE EtD framework may result in superficial use rather than indepth collection and assessment of evidence for the different criteria. In particular, guidance on how to frame questions for and collect evidence towards criteria beyond balance of health benefits and harms appears to be missing.

This paper, one of a series exploring the implications of complexity in systematic reviews and guideline development, reports on the development of a new EtD framework that is rooted in WHO norms and values and suitable for a broad range of health interventions, including complex interventions and interventions delivered in complex systems.

The paper addresses the following three objectives:

1. Develop an EtD framework that (a) is firmly rooted in WHO norms and values and reflective of the changing global health landscape, and (b) encompasses a comprehensive set of criteria suitable for clinical practice, public health and health system interventions.
2. Explore the value of this framework in relation to (a) complexity in individual-level as well as population-level and system-level interventions, (b) the views of developers of the WHO guidelines (global level), and

Rehfuess EA, *et al*. *BMJ Glob Health* 2019;**4**:e000844. doi:10.1136/bmjgh-2018-000844

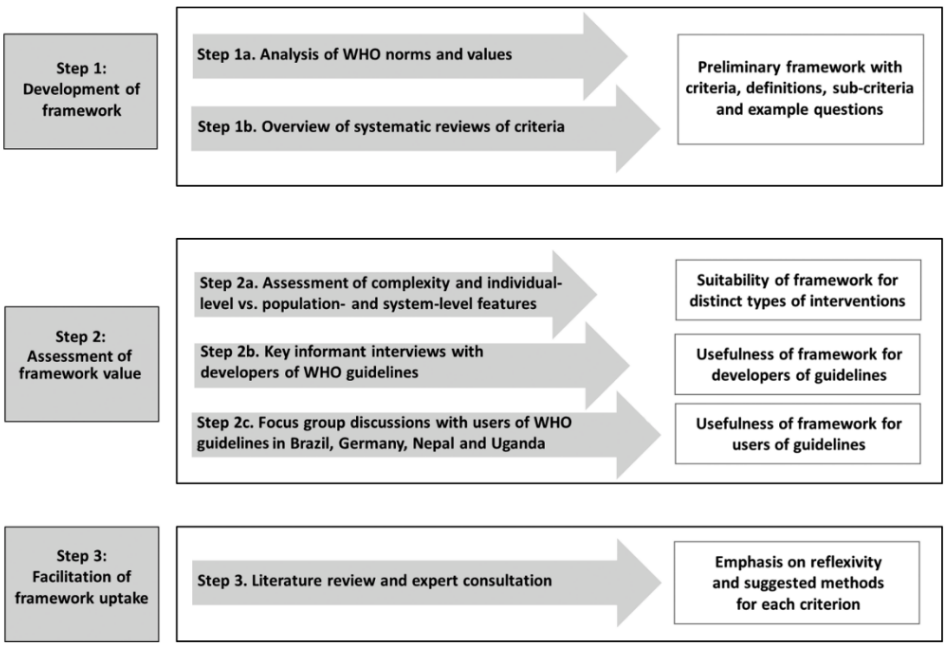

Figure 1 Towards a useful and operational WHO-INTEGRATE (INTEGRATe Evidence) framework.

(c) the views of users of the WHO guidelines (national level).

3. Facilitate uptake of the framework by emphasising the need for structured, evidence-based reflection and suggesting methods to populate the criteria with evidence in the context of decision-making under uncertainty.

The EtD framework developed out of this process is referred to as the WHO-INTEGRATE (INTEGRATe Evidence) framework version 1.0. It is proposed for use in the WHO guideline development as well as in other guideline development or health decision-making processes at the global or national level. It is intended to be used holistically—from the beginning of a health decision-making process to formulating recommendations or making a decision at the end of this process.

METHODS

In addressing these objectives, we followed a three-step approach, as illustrated in figure 1.

This paper provides an overview of the research project with all of its constituent components. It presents the current version of the framework (WHO-INTEGRATE framework version 1.0) and its intended use. It also reports the detailed methods and findings for steps 1a, 2a and 3, as well as an overview of the methods and findings for steps 1b and 2b. A full account of the methods and findings of step 1b is currently in preparation (Stratil *et al*, forthcoming). An integrated analysis of the views of those developing (step 2b) and using WHO guidelines (step 2c) with respect to the WHO-INTEGRATE framework will also be published separately (Stratil *et al*, forthcoming).

Step 1: Development of the framework

In step 1a we analysed WHO norms and values and, rooted in these norms and values, proposed a structure for the WHO-INTEGRATE framework and derived preliminary criteria. A universally agreed normative theory for health does not exist, but most rivalling theories converge on a set of principles.[31] As the use of these principles is less restrictive than the choice of one theory over another, we pursued a principles-based approach,[31 32] and used WHO norms and values as the guiding principles for developing a new EtD framework. Given the complexities of normative orientation in modern pluralistic and globalised societies, we believe that WHO norms and values represent a useful foundation: they are rooted in the universally recognised concept of human rights and receive their legitimacy from having been agreed on by all 194 Member States of the WHO. To identify WHO norms and values of relevance to the process of guideline development and implementation, we used the WHO Constitution[33] and chapter 5 *'Incorporating equity, human rights, gender and social determinants into guidelines'* of the *WHO Handbook for Guideline Development*[22] as a starting point. Given the emphasis in these two documents on human rights, equity and non-discrimination, social determinants of health and the role of health systems, we retrieved and analysed relevant related documents,[34–40] including several public health ethics frameworks.[16 18 41–50] We

also reviewed the SDGs[51] in view of their likely impact at global and national levels and as WHO is mainstreaming these throughout the organisation's work.[52–54]

From these documents and sources, we derived principles and concepts. The structure of the WHO-INTEGRATE framework was developed via an iterative process among coauthors. We explored the meaning of different principles and concepts and assessed overlap and redundancies, making rearrangements to derive preliminary criteria. In doing so, we used a structure and wording as close as possible to the existing GRADE EtD framework to build on its strengths and to maximise potential synergies. On several occasions, we also consulted with members of the WHO Guidelines Review Committee as well as other WHO staff considered experts on selected principles or concepts (see Acknowledgements).

During the development process, we focused on substantive criteria or what decisions are based on (eg, cost, acceptability) rather than procedural criteria or how the decision-making process is conducted (eg, composition of guideline panels, participation, transparency). This is consistent with the approach promoted by the *WHO Handbook for Guideline Development*,[22] whose overall purpose is to specify procedural rules for an objective, transparent and acceptable guideline development process. Embedded in these procedural rules, the current WHO EtD framework (table 10.1 of the *WHO Handbook for Guideline Development*)[22]—and, by extension, the WHO-INTEGRATE framework presented here—is concerned with how to facilitate the use of evidence in decision-making in a structured and comprehensive manner. It is important to note that a distinction between structural and procedural aspects is widely practised in guideline development and several other health decision-making processes,[5 6 19 55] but is not commonly seen in the public health ethics literature.[43 44 46 56]

In step 1b we refined the preliminary criteria derived from WHO and other related documents and supplemented them with a comprehensive set of subcriteria; we also developed definitions for criteria and example questions relating to each of the subcriteria. We conducted an overview of systematic reviews of criteria used in decision-making, priority setting and resource allocation processes for health to derive a comprehensive set of health-relevant criteria (Stratil *et al*, forthcoming). We then compared the preliminary criteria developed in step 1a against this comprehensive set of criteria and subcriteria. To do so, one author (JMS), in a discussion with a second author (EAR), allotted the subcriteria obtained from the overview of systematic reviews to the preliminary criteria within the WHO-INTEGRATE framework. Subcriteria that did not fit were kept in a separate category. Any uncertainties were resolved in discussion with a third author (RB).

We then prepared definitions for each of the criteria using the above-described source documents for health norms and values, existing health decision frameworks (Stratil *et al*, forthcoming), and any definitions or descriptions provided in the publications included in our overview of systematic reviews of criteria (Stratil *et al*, forthcoming). Where appropriate, we also drew on additional key documents (eg, Scott *et al*[57] for the definition of *acceptability*, Hultcrantz *et al*[58] for the conceptualisation of *quality of evidence*, and Maeckelberghe and Schröder-Bäck[59] for details on the subcriteria for *human rights and sociocultural acceptability* and *health equity, equality and non-discrimination*). Each definition (1) provides an overall definition of the criterion, (2) offers details and explanations regarding the subcriteria, and (3) gives guidance on how the criterion in question influences the recommendation.

As we prepared definitions, we also examined the extent to which the criteria and subcriteria relate to the intervention itself *versus* the health system and the broader context, in which an intervention is implemented. For example, the same label (eg, equity) may be employed to describe different underlying concepts, relating to *process versus outcome* (an intervention can either be implemented taking equity principles into account, or it can increase or decrease equity in health outcomes) and the *point in time* when these criteria apply (eg, equity before, during or after intervention implementation). To enable better access to sometimes abstract constructs, we also developed example questions for each of the subcriteria, drawing on the same set of documents as above.

Step 2: assessment of framework value

In step 2a we explored whether the WHO-INTEGRATE framework would be able to accommodate different types of health interventions and different features of complexity. We assessed to what extent the WHO-INTEGRATE framework would be able to accommodate features of distinct types of health interventions.[60] We broadly distinguished between interventions targeting individuals (eg, diagnosis, treatment or preventative measures addressed at individuals), interventions targeting populations, and interventions targeting the health system or other systems. Population-level interventions encompass those concerned with whole populations or population groups as defined by their age, sex, risk factor profile or other characteristics; they are often implemented in specific settings or organisations (eg, school health programmes). System-level interventions specifically redesign the context in which health-relevant behaviours occur; they are often implemented through geographical jurisdictions from national to local levels (eg, laws and regulations regarding the taxation, sale and use of tobacco products). Health system interventions represent a specific type of system-level intervention and often result in complex rearrangements across multiple health system building blocks (eg, task shifting as a process of delegating specific health service tasks from medical doctors or nurses to less specialised health workers). Interventions implemented at any of these levels can be conceptualised and analysed from a complexity perspective. To do so, we mapped core and additional components of complex interventions as defined in the iCAT_SR tool[61] and sources of complexity

in systems reported in another paper in this series[28] against the included criteria.

In step 2b we examined the usefulness and relevance of the WHO-INTEGRATE framework and its criteria among those developing WHO guidelines. We conducted key informant interviews with individuals who had recently participated in a WHO guideline development process. In consultation with the Secretariat of the WHO Guidelines Review Committee, we purposively selected three ongoing or completed guidelines that had applied the current WHO EtD framework,[29 62 63] seeking to cover distinct types of health interventions and positive as well as more difficult experiences with the application of the framework. For each guideline, we interviewed the WHO staff coordinating the guideline, the Chair of the guideline development group and the methodologist. The interviews were semistructured and used a pretested interview guide concerned with practical considerations (eg, understandability, operationalisability), as well as an assessment of missing and redundant criteria of the WHO-INTEGRATE framework. Interviewees were also asked to reflect on the implications of the WHO-INTEGRATE framework for evidence collection and guideline formulation. Interviews were held between June and November 2017 either face-to-face at the WHO Headquarters in Geneva or by telephone (JMS). Interviews were audiotaped and transcribed; data were then analysed by two researchers (JMS and IBS) using qualitative content analysis.[64] We employed a combination of deductive (based on the guiding research questions) and inductive approaches using the software MAXQDA (VERBI Software, Berlin).

Step 3: facilitation of framework uptake

We critically examined how to enable use of the WHO-INTEGRATE framework as intended, and generated a table linking the criteria with suggested methods for primary research, evidence synthesis and assessing quality of evidence. The current WHO EtD framework is intended to be used right from the planning stages of a guideline, to help derive relevant questions and structure the process, but in practice it is usually used at the end of a guideline process to help decide on the recommendations. To determine how the new framework could be used more holistically, we reflected on the literature reviewed in the context of developing the WHO-INTEGRATE framework and sought feedback from a large number of experts (see Acknowledgements). We specifically sought suggestions on how to use the framework during the early stages of the guideline development process and in a context-specific manner.

To make it easier for guideline panels to populate the criteria in the framework with evidence, we identified types of primary research, evidence synthesis methods and methods for assessing evidence quality that could inform each criterion. To accomplish this, two researchers (AP and EAR) reviewed the research questions and methods described or mentioned in the systematic review of health decision frameworks and the overview of systematic reviews (Stratil *et al*, forthcoming). We also consulted a broad range of experts comprising other authors of papers in this series, selected guideline development organisations (eg, Guidelines International Network, UK National Institute for Health and Care Excellence) and researchers with an interest in evaluating complex health technologies (see Acknowledgements).

RESULTS

Developing the preliminary framework

Using the review of the WHO Constitution,[33] chapter 5 of the *WHO Handbook for Guideline Development*[22] and other source documents, we identified six major, partly overlapping concepts. Further sorting of these yielded four sets of principles and concepts (human rights principles, ethical principles, sustainability elements and health system goals and building blocks). Figure 2 illustrates how we derived preliminary criteria from WHO norms and values.

- *Human rights principles*, for the purposes of this framework, were primarily derived from international human rights law and its interpretation by the Committee on Economic, Social and Cultural Rights' General Comment on the Right to the Highest Attainable Standard of Health (Art 12).[34] These contain the interrelated concepts of availability and accessibility of public health and healthcare facilities, goods and services, which are required to be of appropriate quality and acceptable to users. They also include the general human rights principles of equity and non-discrimination, accountability and participation.
- Given the large number of biomedical and public health ethics frameworks,[44 46 56] in consultation with WHO, we structured the *ethical principles* primarily according to the public health ethics framework of Childress and colleagues. This framework inter alia includes the aspects of producing benefits, avoiding harms, maximising the balance between benefits and harms, as well as distributive justice and autonomy.[41] Based on analytical tools by the Nuffield Council of Bioethics,[45] we also added the principle of low intrusiveness, which is related to privacy and dignity.
- Acknowledging the importance of the social determinants of health and the SDGs, we derived *sustainability elements* to capture the wide range of factors that promote conditions in which people can lead a healthy life and allow societies and individuals to develop and flourish; these sustainability elements also reflect the societal impact that interventions can have beyond health outcomes. Importantly, good health is both a precondition for achieving sustainable development and an outcome of sustainable development.[65]
- To capture the importance of feasibility of implementation as well as the impact of interventions on the health system, we used the WHO health systems framework with its four *goals* (ie, improved health,

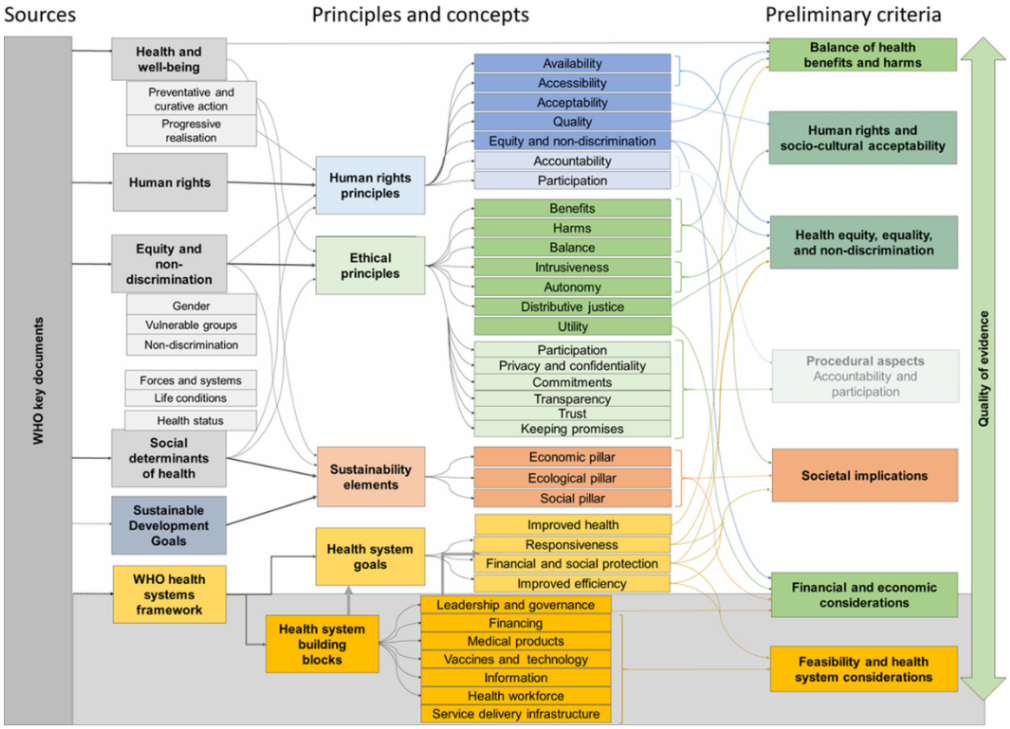

Figure 2 Sources and concepts for deriving principles-based preliminary criteria rooted in WHO norms and values.

responsiveness, social and financial protection, improved efficiency) and six *building blocks* (ie, leadership and governance, financing, medical products, vaccines and technologies, information, health workforce, service delivery infrastructure).[19 38–40]

Figure 3 presents the WHO-INTEGRATE framework with its six criteria: *balance of health benefits and harms*, *human rights and sociocultural acceptability*, *health equity, equality and non-discrimination*, *societal implications*, *financial and economic considerations*, and *feasibility and health system considerations*. A seventh criterion, *quality of evidence*, represents a metacriterion that applies to each of the six substantive criteria. All seven criteria are relevant to health decision-making and the formulation of recommendations as part of the guideline development process. Each criterion may apply to interventions targeting individuals, populations or systems, or any combination of these levels.

While priority of the problem featured in both the health decision frameworks included in our systematic review (eg, Alonso-Coello *et al*[13]) and in the overview of systematic reviews of criteria (eg, Guindo *et al*[6]), we did not include this as a stand-alone substantive criterion for two reasons: First, many of the aspects included, for example, political will or public concern, are used to inform the decision to develop a guideline (or make another health decision) and thus apply before the start of the guideline development process. Second, selected aspects are captured under the other six substantive criteria, for example, burden of disease features under *balance of health benefits and harms*, and large cost of disease

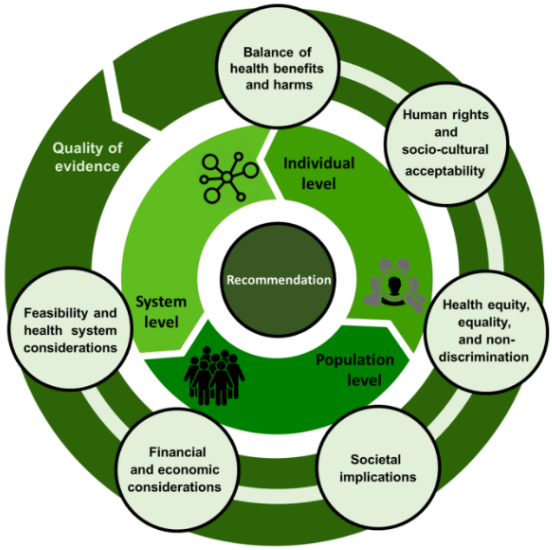

Figure 3 The WHO-INTEGRATE (INTEGRATe Evidence) framework version 1.0.

to health system features under *financial and economic considerations*.

Defining criteria, subcriteria and example questions

Our overview of systematic reviews yielded more than 30 systematic reviews that contained several thousand criteria and subcriteria currently used in decision-making (Stratil *et al*, forthcoming). Recurrent aspects addressed by the subcriteria focused on the health outcomes and benefits of the intervention, health benefit for individuals and the benefit for society as a whole, the societal importance of the disease, economic considerations, quality or uncertainty of evidence, as well as population priorities, priorities within the health system and stakeholders' interests and pressures. Feasibility criteria were concerned with the available budget, the capacities within the health system, technological complexity and acceptability of the intervention within society. Some systematic reviews were primarily concerned with interventions that would benefit vulnerable or marginalised populations (eg, children, mothers, people with lower socioeconomic status). In many reviews, normative criteria such as ethics, justice or fairness were mentioned without clear definitions or contextualisation. This comprehensive list did not yield any further criteria beyond the seven presented in figure 2. It did, however, provide many subcriteria as well as elements used in the development of detailed definitions and example questions for each criterion.

Suitability of the framework for decisions about complex health interventions

An earlier paper in this series[28] emphasises the importance and added value of reviewing evidence from a complex systems perspective. In developing the new EtD framework, we wanted to ensure that it would be fit for purpose when making decisions about complex interventions implemented in complex systems. We first explored to what extent different features of intervention and system complexity apply to two broad categories of interventions, that is, individual-level versus population-level and system-level interventions (table 1). Notably, even population-level and system-level interventions (eg, regulations and programmes to increase access to improved sanitation) eventually bring about changes in individual behaviour (eg, use and maintenance of toilets or latrines). Some criteria apply to a greater extent with population-level and system-level interventions (eg, *societal implications*) than individual-level interventions. Some subcriteria may take on a different meaning when applied to individual-level versus population-level and system-level interventions (eg, *autonomy*). Broadly speaking, most features of complex interventions apply to both individual-level and population-level/system-level interventions but are more salient for the latter. In contrast, many features of complex systems only apply to population-level and system-level interventions.

The last column of table 1 illustrates that distinct features of complexity do not neatly map onto specific criteria. Instead, distinct features of complexity usually affect multiple, sometimes all, criteria in the WHO-INTEGRATE framework. For example, the worked example of childhood obesity, introduced in an earlier paper in this series,[28] discussed adaptivity of the system in response to raised taxes on soft drinks (eg, creation of lower-sugar alternatives by the soft drinks industry). This adaptivity can thus influence the *balance of health benefits and harms* (eg, consumption patterns of soft drinks change but in less pronounced ways, thereby dampening the expected effect on childhood obesity), and it may even have unwanted social consequences by stigmatising those unable to afford soft drinks (*social impact*). Raising taxes on only one sugar-sweetened product may lead to increasing the sugar content of other sugar-sweetened products (*impact on economy, broader positive or negative health-related impacts*) or have implications on agricultural production patterns nationally and internationally (*impact on economy and environmental impact*), illustrating the complexity of downstream implications of a 'simple' intervention. Drawing on the same worked example, box 1 illustrates how a simple linear perspective on the effect of an intervention will place the emphasis on one or a few criteria for decision-making, whereas a complexity perspective may take all criteria into account when making a recommendation.

Usefulness of the framework from the perspective of WHO guideline developers

The key informants we interviewed had been involved in developing three very different guidelines—the WHO recommendations on antenatal care,[29] the WHO consolidated guideline on sexual and reproductive health and rights of women living with HIV,[62] and the WHO guideline on risk communication (online supplementary table S1).[63] Each of these guidelines faced different challenges in terms of scope, availability of evidence and ability to incorporate multiple perspectives. All three had used the current WHO EtD framework with varying success. The diverse experiences and viewpoints of the key informants on the practical application of these criteria in guideline development were helpful in refining the framework. Further detail on and complete findings from the key informant interviews will be reported separately (Stratil *et al*, forthcoming).

Most participants commented positively on the WHO-INTEGRATE framework and highlighted the value of a criterion assessing societal implications, as well as the broader and more detailed specification of the criteria *human rights and sociocultural acceptability* and *health equity, equality and non-discrimination*. Two participants questioned the added value of the new EtD framework, since any guideline development process led by an experienced methodologist would automatically address the details covered in the subcriteria. Several participants were concerned about the workload that the use of the WHO-INTEGRATE framework might add to the guideline development process.

Table 1 Features of complex interventions (adapted from Lewin *et al*[61]) and complex systems (adapted from Petticrew *et al*[28]) and their impact on individual-level versus population-level and system-level interventions, as well as criteria in the WHO-INTEGRATE framework

Population-level and system-level interventions	Individual-level interventions		Complexity-relevant differences between individual-level and population-level /system-level interventions	WHO-INTEGRATE framework criteria that are typically relevant*
Features of complex interventions				
Number of active components in the intervention[61]; interactions between components of complex interventions.[28]	+	++	Both types of interventions can comprise multiple components entailing synergistic or dissynergistic interactions among them. For population-level and system-level interventions, these interactions tend to occur among a greater number of more diverse components located at one or several organisational levels.	► Balance of health benefits and harms. ► Human rights and sociocultural acceptability. ► Health inequity, equality and non-discrimination. ► Societal implications. ► Financial and economic considerations. ► Feasibility and health system considerations.
Number of behaviours of recipients to which the intervention is directed.	+	++	Both types of interventions can require behaviour change among recipients. For curative and preventative interventions at the individual level, these mostly relate to treatment adherence or tightly defined health-relevant behaviours, often among an 'activated' population seeking care or willing to engage in other ways. Population-level and system-level interventions tend to be concerned with a larger set of behaviours directly or indirectly linked to health, often in healthy general or at-risk populations.	► Feasibility and health system considerations.
Range and number of organisational levels targeted by the intervention.	–	++	Individual-level interventions tend to target their recipients in a defined setting, for example, in a household or healthcare setting. Many population-level and system-level interventions target multiple levels, for example individuals living in households located in communities and influenced by community-level or national-level interventions; importantly, they often concern sectors beyond health.	► Balance of health benefits and harms. ► Financial and economic considerations. ► Feasibility and health system considerations.
Level of skill required by those delivering the intervention.	++	++	The skills required for effective intervention delivery vary greatly depending on the nature of an intervention, and can be equally high for individual-level and population-level/system-level interventions. For population-level and system-level interventions, there may be a greater number of distinct implementation agents with a more diverse set of necessary skills.	► Human rights and sociocultural acceptability. ► Feasibility and health system considerations.
Level of skill required by those receiving the intervention.	++	++	Both types of interventions can require a high level of skill among recipients, where skill can refer to specific (technical) abilities, as well as broader resources and characteristics, such as motivation and capacity (time, money, physical and mental energy). Interventions directed at individuals tend to require greater recipient skills and resources than many population-level and system-level interventions. Population-level and system-level interventions, on the other hand, often impact multiple behaviours related to diverse aspects of life and thus potentially rely on a more diverse set of skills and resources.	► Human rights and sociocultural acceptability. ► Health inequity, equality and non-discrimination. ► Feasibility and health system considerations.

Continued

Table 1 Continued

Population-level and system-level interventions			Complexity-relevant differences between individual-level and population-level /system-level interventions	WHO-INTEGRATE framework criteria that are typically relevant*
Individual-level interventions				
Features of complex systems				
Interactions of interventions with context and adaptation[28]/ degree of tailoring intended or flexibility permitted across sites or individuals in applying or implementing the intervention.[61]	+	++	Individual-level interventions tend to involve a small degree of tailoring, typically revolving around the health professional–patient relationship. In contrast, many population-level and system-level interventions are highly context-dependent and, in order to be effective, their design and delivery strategies must be tailored to the setting or context in which they are to be implemented.	▶ Balance of health benefits and harms. ▶ Human rights and sociocultural acceptability. ▶ Health inequity, equality and non-discrimination. ▶ Societal implications. ▶ Feasibility and health system considerations.
System adaptivity (how does the system change).	–	++	Some population-level and system-level interventions may directly attempt to change or indirectly influence the context in which they are implemented. The system thus reacts and adapts in expected or unexpected ways to the intervention.	▶ Balance of health benefits and harms. ▶ Societal implications. ▶ Feasibility and health system considerations.
Emergent properties.	–	++	Population-level and system-level interventions tend to impact diverse aspects of life and may produce emergent features in relation to one or several of these (eg, changes in social norms). Some individual-level interventions, when implemented and viewed at the population/system level, can yield emergent features (eg, herd immunity as a result of vaccination).	▶ Balance of health benefits and harms. ▶ Human rights and sociocultural acceptability. ▶ Health inequity, equality and non-discrimination. ▶ Societal implications.
Non-linearity and phase changes.	–	++	Some population-level interventions may only begin to deliver meaningful outcomes once they have reached a certain scale (phase changes at a threshold); they may be highly effective at particular levels of coverage and less effective at others.	▶ Balance of health benefits and harms. ▶ Human rights and sociocultural acceptability. ▶ Societal implications.
Negative and positive feedback loops.	–	++	Population-level and system-level interventions with their specific components or the set of interacting components can produce negative feedback loops and thus reduce the overall intervention effect (damping); similarly, positive feedback loops may result in an overall intervention effect that is greater than expected.	▶ Balance of health benefits and harms. ▶ Human rights and sociocultural acceptability. ▶ Health inequity, equality and non-discrimination. ▶ Financial and economic considerations. ▶ Feasibility and health system considerations.

Continued

Table 1 Continued

Population-level and system-level interventions		Complexity-relevant differences between individual-level and population-level /system-level interventions	WHO-INTEGRATE framework criteria that are typically relevant*
Individual-level interventions			
Multiple (health and non-health) outcomes and long complex causal pathways.	+ ++	Both types of interventions can be characterised by multiple outcomes and long, complex causal pathways. Given their large number of components impacting health as well as non-health outcomes, this feature of complex systems is particularly prevalent among population-level and system-level interventions and complicated by often long lag periods. An individual-level intervention has to be sufficiently popular and impactful to diffuse through families, peers and among the broader community or nation to eventually have population-relevant impacts, whereas a population-level or system-level intervention tends to have more immediate impacts (intended and unintended).	▲ Balance of health benefits and harms. ▲ ▲ Human rights and sociocultural acceptability. ▲ Societal implications. ▲ ▲ Financial and economic considerations.

–, indicates not relevant; +, indicates somewhat relevant; ++, indicates highly relevant.
*Each feature of a complex system tends to influence most or all criteria; here we highlight those criteria that may be of greatest relevance.
INTEGRATE, INTEGRATe Evidence.

Box 1 Thinking through the criteria in relation to raised taxes on soft drinks and their implications

A simple perspective on raising taxes on soft drinks would emphasise the linear impacts of this intervention on consumption of sugar-sweetened beverages (intermediate outcome) and different measures of childhood obesity (ultimate outcome of interest); with this perspective, the criterion *balance of health benefits and harms* would warrant the most attention. A complexity perspective on the same intervention would not start off with a preconception about a single criterion being most influential but carefully examine all criteria. For illustration purposes, this complexity perspective would examine acceptability among and likely reactions from different groups of stakeholders (eg, children, their parents), and pay specific attention to the response from vendors and producers of soft drinks (eg, potential sugar reduction in drinks with implications for the prices of these drinks), which may dampen the expected effect of the tax intervention, in terms of changes in consumption patterns, perceptions of the intervention and changes in social norms.

This complexity perspective would also encompass potential negative impacts on health equity, equality and non-discrimination (eg, expected or unexpected changes in consumption patterns across different socioeconomic or other population groups), explore positive or negative social, environmental or economic impacts (eg, changes in social norms in relation to sugar-sweetened beverages or their alternatives being more or less desirable among different population groups, changes in acceptability of further interventions to reduce sugar consumption), adopt a societal perspective in estimating the financial and economic impacts of the intervention (eg, including how costs and benefits of the raised taxes are distributed among different stakeholders and sectors), and pay attention to feasibility and health system considerations (eg, implications for human resources involved with other ongoing efforts to reduce consumption of sugar-sweetened beverages and childhood obesity).

Specific remarks were made in regard to (1) missing criteria and subcriteria; (2) the hierarchy and order of criteria and subcriteria; (3) overlap and redundancies between criteria and subcriteria; (4) the precise wording and definitions of criteria; (5) the need for (more) guidance on how to use and interpret criteria and subcriteria; (6) the challenges of identifying and synthesising the required evidence; (7) resource, time and skill implications for the guideline development process; as well as (8) procedural aspects for using the framework in the guideline development process.

In response to these concerns and suggestions, we made several modifications, including changing the name and definition of several criteria and subcriteria to improve clarity and reduce overlap. We also expanded the example questions for the subcriteria to improve understandability and facilitate the development of specific questions for a given guideline. Moreover, we added suggestions on how to prioritise among criteria and subcriteria in a problem-specific and context-specific manner. Finally, we emphasised the importance of incorporating the voices of those directly affected by

the recommendations into the guideline development process.

Table 2 presents the WHO-INTEGRATE framework version 1.0 criteria with abbreviated definitions and lists subcriteria. Online supplementary table S2 provides detailed definitions of the criteria as well as example questions for each of the sub criteria.

Facilitating uptake: using the framework holistically and populating the criteria with evidence

The WHO-INTEGRATE framework is intended to improve transparency in health decision-making by supporting a structured process of reflection and discussion in a problem-specific and context-specific manner. To be most effective, this process must begin at the start of a guideline or other health decision-making process and must take evidence into account. The WHO-INTEGRATE framework is *not* intended as a 'tick-box exercise'; there must be prioritisation of the most relevant criteria and subcriteria depending on the questions addressed by a given guideline, and the time and resources at disposition. It would be impossible and probably unnecessary for every guideline development or health decision-making process to examine all subcriteria. This flexibility can, however, lead to misuse, as stakeholders may disproportionately (eg, academics from high-income countries) or unduly (eg, participants with substantial declared or undeclared financial or other conflicts of interest) influence the decision-making process. Safeguards can be put in place through explicit procedures, in particular in relation to the composition of guideline panels or other decision-making groups. The WHO-INTEGRATE framework is also *not* an algorithm for integrating evidence across different criteria: making decisions under uncertainty and agreeing on trade-offs across criteria and subcriteria and among (and within) diverse stakeholder groups remain a core task for a guideline panel.

All criteria are important and should be reflected on, but their relevance varies depending on the type of health decision and the decision-making context. In contrast, not all subcriteria are always relevant. At the start of a guideline or other decision-making process, an appropriately composed guideline panel or other decision-making group needs to discuss which of the subcriteria are applicable and useful in relation to the nature and specific characteristics of the intervention (see table 1); this group will also need to consider the specific information needed to populate criteria or subcriteria (see table 3). Complexity in the intervention and complexity in the system into which this intervention is implemented can usually be detected; the critical question is whether it is of value to examine this complexity in depth (see box 1 in this paper and box 2 in an earlier paper in this series[28]). This prioritisation process should take the views of relevant stakeholder groups into account; which stakeholder groups are relevant depends on the nature of the problem and the institutional as well as broader physical and social context. In principle, these should include those directly affected by the intervention (eg, patients, beneficiaries), those financing (eg, health insurance providers, ministries of health, other ministries) or implementing the intervention (eg, healthcare providers, public health professionals, professionals outside of the health sector), as well as the general public.

A systematic weakness in many guideline development and other health decision-making processes is that consumer participation is obviated and guideline panels often substitute their own values and views for those of patients/beneficiaries. The voices of patients/beneficiaries and other relevant stakeholder groups can be incorporated through direct participation or representative surveys[66] as well as qualitative research (see table 3).

The guideline panel will also need to decide how best to populate the criteria with evidence and whether a formal evidence synthesis or a more pragmatic approach is warranted for each. This decision will be influenced by the relevance of criteria and subcriteria in relation to a specific intervention or decision, and by the likely types and quantity of evidence available, as well as time and resource constraints. At the end of the process, the guideline panel will need to reassess the criteria and relevant subcriteria in light of the assembled evidence and make a judgement regarding each criterion.

Table 3 suggests relevant types of primary research, evidence synthesis or mapping methods, streamlined or pragmatic approaches, as well as methods to assess the quality of evidence for each of the six substantive criteria. We provide a collection of suitable primary research and synthesis approaches, but make no firm distinction between more or less suitable methods. We note that the approach to gathering evidence may depend on the criterion: for some criteria a systematic review will be most appropriate, while for others a representative survey or other single primary study may be more suitable. Surprisingly, the majority of the health decision frameworks included in our systematic review (Stratil *et al*, forthcoming) did not offer insights for operationalising frameworks, for example by specifying research questions or suggesting methods for primary research or evidence synthesis. The GRADE EtD framework[13 67] and the EUnetHTA (EUropean network for Health Technology Assessment) core model[68] provided some methods. We also identified relevant information in the following sources: the EVIDEM (Evidence and Value: Impact on DEcisionMaking) framework,[14] Marckmann and colleagues,[16] the health systems framework[19] and publications included in our overview of systematic reviews of criteria.[69] Expert consultation played a critical role in identifying methods for inclusion in table 3.

DISCUSSION

Added value of the WHO-INTEGRATE framework

The WHO-INTEGRATE framework represents a new comprehensive EtD framework that is rooted in WHO norms and values. It offers an explicit conceptualisation

Table 2 WHO-INTEGRATE framework version 1.0: criteria with abbreviated definitions, subcriteria and implications for a recommendation. All criteria are relevant for all interventions in health decision or guideline development processes. For subcriteria there should be a discussion as to which are most relevant and if or how evidence should be collected to inform these. Online supplementary table S2 provides detailed definitions of the criteria and example questions for each of the subcriteria.

Criteria and abbreviated definitions	Subcriteria	Implications for a recommendation
Balance of health benefits and harms The balance of health benefits and harms reflects the magnitude and types of health impact of an intervention on individuals or populations, taking into account how those affected value different health outcomes.	▲ Efficacy or effectiveness on health of individuals. ▲ Effectiveness or impact on health of population. ▲ Patients'/beneficiaries' values in relation to health outcomes. ▲ Safety risk profile of intervention. ▲ Broader positive or negative health-related impacts.	The greater the net health benefit associated with an intervention, the greater the likelihood of a general recommendation in favour of this intervention.
Human rights and sociocultural acceptability This criterion encompasses two distinct constructs: The first refers to an intervention's compliance with universal human rights standards and other considerations laid out in international human rights law beyond the right to health (as the right to health provides the basis of other criteria and subcriteria in this framework). The second, sociocultural acceptability, is highly time-specific and context-specific and reflects the extent to which those implementing or benefiting from an intervention as well as other relevant stakeholder groups consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention.	▲ Accordance with universal human rights standards. ▲ Sociocultural acceptability of intervention to patients/beneficiaries and those implementing the intervention. ▲ Sociocultural acceptability of intervention to the public and other relevant stakeholder groups. ▲ Impact on autonomy of concerned stakeholders. ▲ Intrusiveness of intervention.	All recommendations should be in accordance with universal human rights standards and principles. The greater the sociocultural acceptability of an intervention to all or most relevant stakeholders, the greater the likelihood of a general recommendation in favour of this intervention.
Health equity, equality and non-discrimination Health equity and equality reflect a concerted and sustained effort to improve health for individuals across all populations, and to reduce avoidable systematic differences in how health and its determinants are distributed. Equality is linked to the legal principle of non-discrimination, which is designed to ensure that individuals or population groups do not experience discrimination on the basis of their sex, age, ethnicity, culture or language, sexual orientation or gender identity, disability status, education, socioeconomic status, place of residence, or any other characteristics.	▲ Impact on health equality and/or health equity. ▲ Distribution of benefits and harms of intervention. ▲ Affordability of intervention. ▲ Accessibility of intervention. ▲ Severity and/or rarity of the condition. ▲ Lack of a suitable alternative.	The greater the likelihood that the intervention increases health equity and/or equality and that it reduces discrimination against any particular group, the greater the likelihood of a general recommendation in favour of this intervention.
Societal implications Societal implications recognise that health interventions do not take place in isolation and may enhance or inhibit broader social, environmental or economic goals in the short or long term. It also reflects the fact that many regulatory, environmental or other population-level health interventions are directly aimed at system-level rather than individual-level changes.	▲ Social impact. ▲ Environmental impact.	The greater the net societal benefit associated with an intervention, the greater the likelihood of a general recommendation in favour of this intervention.
Financial and economic considerations Financial and economic considerations acknowledge that available financial (budgetary) resources are constrained and take into account the economic impact of an intervention on the health system, government or society as a whole.	▲ Financial impact. ▲ Impact on economy. ▲ Ratio of costs and benefits.	The more advantageous the financial and economic implications of an intervention, the greater the likelihood of a general recommendation in favour of this intervention.

Continued

Table 2 Continued

Criteria and abbreviated definitions	Subcriteria	Implications for a recommendation
Feasibility and health system considerations Feasibility and health system considerations recognise that the most appropriate and feasible interventions may vary significantly across different contexts, both across countries and across jurisdictions within countries. Legislation and governance, the structure of the health system and existing programmes, as well as human resources and infrastructure, should be taken into account.	▲ Legislation. ▲ Leadership and governance. ▲ Interaction with and impact on health system. ▲ Need for, usage of and impact on health workforce and human resources. ▲ Need for, usage of and impact on infrastructure.	The greater the feasibility of an option from the perspective of all or most stakeholders, the greater the likelihood of a general recommendation in favour of the intervention. The more advantageous the implications for the health system as a whole, the greater the likelihood of a general recommendation in favour of the intervention.
Quality of evidence Quality of evidence, also referred to as certainty of evidence or strength of evidence, reflects the confidence that the available evidence is adequate to support a recommendation. In principle, quality of evidence can be applied across all criteria in the WHO-INTEGRATE framework. As a large number of criteria are integrated in the decision-making process, evidence is interpreted in the broadest sense and allows for relevant contributions from a variety of disciplinary approaches. Moreover, decision-making under uncertainty often involves stakeholder experience and judgement, when stronger evidence is unavailable.	–	The greater the quality of the evidence across different criteria in the WHO-INTEGRATE framework, the greater the likelihood of a general recommendation.

INTEGRATE, INTEGRATe Evidence.

of each criterion and a rationale for including relevant concepts as criteria or subcriteria. The WHO norms and values apply across all WHO Member States and settings, and the new framework should, in principle, be relevant for health decision-making at global, national and subnational levels. It reflects a broad understanding of health and its determinants and takes account of complex interventions and complex systems perspectives. It emphasises sustainability and the interconnectedness between health and other sectors, inherent in the SDGs. While the framework is conceived for individual-level, population-level and system-level interventions, it is likely to be particularly well suited for public health and health system interventions characterised by complexity and/or approached from a complexity perspective. The WHO-INTEGRATE framework is intended as a tool to facilitate structured reflection and discussions from the beginning of a guideline development or other health decision-making process. This has ramifications in terms of the need to prioritise among criteria and subcriteria and the need to collect evidence for each. The framework supports this process by offering structured definitions for each criterion and example questions for each subcriterion, and by suggesting methods for primary research, evidence synthesis and assessing the quality of the evidence.

There are many similarities between the WHO-INTEGRATE framework and the widely used GRADE EtD framework. As stated in our methods, we deliberately attempted to stay as close as possible to the GRADE EtD framework, thus building on established terms and concepts (eg, *balance of health benefits and harms*). In contrast, criteria with a strong normative foundation (eg, *health equity, equality and non-discrimination)* were much less developed in the GRADE EtD framework; notably, the criterion *societal implications*, which has its roots in the recognition of the multisectoral determinants of health, is absent from the GRADE EtD framework. There are also more fundamental differences. While the GRADE EtD framework emphasises the efficacy/effectiveness of interventions and their potential harmful impacts, there is no inherent weighting of criteria in the WHO-INTEGRATE framework: guideline panels must decide in a context-specific and problem-specific manner which criteria and subcriteria are most relevant. Moreover, in contrast to the narrower certainty of evidence concept in the GRADE EtD framework, the WHO-INTEGRATE framework has deliberately adopted a broad quality of evidence concept that applies across all criteria and is not linked to a prespecified grading system. For several criteria (and/or subcriteria) GRADE[70 71] and GRADE CERQual (Confidence in the Evidence from Reviews of Qualitative Research[72]) are the most appropriate approaches to examining quality of evidence, and we would encourage users of the WHO-INTEGRATE framework to adopt these. In fact, another paper in this series explores how complexity can be considered when assessing the certainty of evidence on intervention effectiveness.[73] For other criteria (and/or subcriteria), these

Table 3 WHO-INTEGRATE framework version 1.0: criteria and suggested types of primary studies, evidence synthesis methods and approaches to assessing quality of evidence

Criteria	Types of primary studies*	Evidence synthesis or mapping methods	Pragmatic approaches	Approaches to assessing quality of evidence
Balance of health benefits and harms.	▲ Efficacy or effectiveness on health of individuals/populations: RCTs, pragmatic trials, quasi-experimental studies, comparative observational studies; longer term observational studies, modelling (eg, transmission modelling for infectious diseases). ▲ Patients'/beneficiaries' values in relation to health outcomes: qualitative studies (eg, semistructured interviews, focus groups), cross-sectional studies. ▲ Safety risk profile of intervention: RCTs, quasi-experimental studies, comparative observational studies for anticipated harms; registry studies, longer term observational studies, case series, case reports for unanticipated effects. ▲ Broader positive or negative health-related impacts: RCTs, quasi-experimental studies, observational studies, qualitative studies.	▲ Systematic reviews of efficacy/effectiveness[83] for anticipated effects. ▲ Qualitative evidence syntheses[84 85] and mixed-method reviews[86] or cross-sectional studies[66] for patients'/beneficiaries' values in relation to health outcomes. ▲ Scoping reviews[87 88] for unanticipated effects.	▲ Rapid reviews of efficacy/effectiveness.[89–91] ▲ Overviews of systematic reviews.[83 92]	▲ GRADE.[70 71 73]

Continued

Table 3 Continued

Criteria	Types of primary studies*	Evidence synthesis or mapping methods	Pragmatic approaches	Approaches to assessing quality of evidence
Human rights and sociocultural acceptability.	▶ Accordance with universal human rights standards: mapping of relevant aspects, pro et contra analysis,[93] ethical analysis (eg, casuistry, coherence analysis, wide reflective equilibrium),[94] power analyses, human rights impact assessment.[95] ▶ Sociocultural acceptability of intervention, impact on autonomy of concerned stakeholders, intrusiveness of intervention: mapping of relevant aspects, pro et contra analysis,[93] discourse analysis, qualitative studies (ideally longitudinal to examine changes over time), discrete choice experiments, cross-sectional studies,[66] longitudinal quantitative studies (to examine changes over time), mixed-method studies.	▶ Ethics syntheses[96 97] for accordance with universal human rights standards. ▶ Qualitative evidence syntheses[84 85 98] and mixed-method reviews[86] for sociocultural acceptability and impact on autonomy of concerned stakeholders and intrusiveness of interventions.	▶ Purposively selected studies from different contexts (to illustrate broad spectrum of issues).	▶ GRADE CERQual[72 99] (where applicable). ▶ Q-SEA for ethics analyses.[57]
Societal implications.	▶ Social impacts: RCTs, quasi-experimental studies, comparative observational studies, longitudinal implementation studies, qualitative studies, case studies, power analyses. ▶ Environmental impacts: RCTs, quasi-experimental studies, comparative observational studies, longitudinal implementation studies, qualitative studies, case studies, environmental impact assessments, modelling studies. ▶ Combined social, environmental and economic impacts: health impact assessments, modelling studies (eg, decision-analytical modelling).	▶ Systematic reviews of effectiveness.[83] ▶ Qualitative evidence syntheses.[11 84 85] ▶ Mixed-method reviews.[86] ▶ Health technology assessments.[68]	▶ Purposively selected studies from different contexts (to illustrate broad spectrum of issues).	▶ No standardised approach. ▶ GRADE[70 71] (where applicable).

Continued

Table 3 Continued

Criteria	Types of primary studies*	Evidence synthesis or mapping methods	Pragmatic approaches	Approaches to assessing quality of evidence
Health equity, equality and non-discrimination.	► Impact on health equality and/or health equity, distribution of benefits and harms of intervention: human rights impact assessment,[95] disaggregated RCTs, quasi-experimental or comparative observational studies, RCTs and quasi-experimental or comparative observational studies conducted in disadvantaged groups,[100] power analyses, GIS-based studies, qualitative studies, ethical analysis. ► Affordability of intervention: cross-sectional or longitudinal observational studies, discrete choice experiments, qualitative studies, catastrophic health expenditure studies. ► Accessibility of intervention: health system barrier studies, cross-sectional or longitudinal observational studies, discrete choice experiments, qualitative studies, ethical analysis, GIS-based studies. ► Severity and/or rarity of the condition: health state valuations, cross-sectional studies for severity of condition; observational studies for frequency (incidence, prevalence) of condition. ► Lack of a suitable alternative: situation analysis of intervention options; quantitative or qualitative studies of adverse effects of existing options.	► Quantitative systematic reviews[83] using PROGRESS[101] or PROGRESS PLUS,[102] where possible using prespecified subgroup analyses. ► Quantitative systematic reviews targeting disadvantaged groups. ► Equity weights and social welfare functions in economic analyses (*see Financial and economic considerations*). ► Qualitative evidence syntheses[11 84 85] and mixed-method reviews.[86] ► Ethics syntheses.[96 97]	► Purposively selected studies from different contexts (to illustrate broad spectrum of issues). ► Scoping reviews.[87 88] ► Overviews of systematic reviews.[83 92]	► No standardised approach. ► GRADE[70 71] for subgroup analyses (where applicable). ► Relevant considerations, such as including health equity as an outcome, in Welch *et al*.[103]

Continued

Table 3 Continued

Criteria	Types of primary studies*	Evidence synthesis or mapping methods	Pragmatic approaches	Approaches to assessing quality of evidence
Financial and economic considerations.	► Financial impact: prices and price justifications for unit costs per beneficiary/population according to relevant perspectives, budget impact analysis.[104] ► Impact on economy: economic burden of disease studies,[105] quasi-experimental studies, comparative observational studies, longitudinal implementation studies, qualitative studies, case studies, modelling studies. ► Ratio of costs and benefits: economic analyses as a comparative analysis of alternative courses of action in terms of their costs and consequences (eg, cost-minimisation analysis, cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis).	► Comprehensive or representative cost or budget impact data at the appropriate level (global, regional, national, subnational). ► Economic burden of disease studies undertaken at the appropriate level (global, regional, national, subnational). ► Economic analyses undertaken at the appropriate level[106 107] or economic analysis reviews.[108–111]	► Cost or budget impact data for purposively selected contexts. ► Economic analyses undertaken for selected contexts.	► No standardised approach. ► Relevant considerations in Drummond *et al*[106] (chapter 3 and box 3.1) and Brunetti *et al*.[112]
Feasibility and health system considerations.	► Legislation, leadership and governance, interaction with and impact on health system, need for, usage of and impact on health workforce, human resources and infrastructure: health systems research,[113] including mapping of relevant aspects, situation analysis, cross-sectional studies, qualitative studies, case studies.	► Qualitative evidence syntheses,[11 84 85] mixed-method reviews.[86]	► Formal consultation of content experts.	► No standardised approach. ► GRADE CERQual[72] (where applicable).

*This table offers a collection of suitable methods rather than guidance on the most appropriate method, which depends on the specific research question. Where appropriate, the order in which the methods are presented implies a hierarchy of evidence (eg, RCTs are more suited to assessing questions of efficacy than modelling).
GIS, geographical information system; GRADE, Grading of Recommendations Assessment, Development and Evaluation; GRADE CERQual, Confidence in the Evidence from Reviews of Qualitative Research; HTA, health technology assessment; INTEGRATE, INTEGRATe Evidence; Q-SEA, Quality Standards for Ethics Analyses in HTA; RCT, randomised controlled trial.

existing tools are not well suited, and we hope that more appropriate approaches will become available—whether through further developments within the GRADE Working Group or independent efforts.

The GRADE EtD framework allows for tailoring of criteria, for example by considering a detailed judgement as a stand-alone criterion or by removing a criterion from the GRADE EtD framework and considering it prior to the start of the decision-making process[13]; in fact, refinement of the GRADE EtD framework continues and has already resulted in suggestions towards more detailed specifications of selected criteria.[74] Similarly, we expect various developments towards a version 2.0 of the WHO-INTEGRATE framework (see below). We thus envisage specific innovations to be adopted across these evolving frameworks and, potentially, convergence over time.

Strengths and limitations of the development process

In developing the WHO-INTEGRATE framework, we combined a principles-based approach with an overview of systematic reviews of decision criteria and thus ensured a solid, comprehensive normative foundation. We were explicit and transparent as to how criteria (see figure 1) and subcriteria (Stratil *et al*, forthcoming) were derived. While there is some conceptual overlap at the level of the criteria (eg, *societal implications* and *financial and economic considerations*), there are no significant redundancies among the subcriteria (Stratil *et al*, forthcoming). Cross-linkages among the criteria are emphasised in the definitions and example questions.

Solely adapting the substantive criteria may be insufficient to overcome limitations in guideline development or other decision-making processes.[22] The WHO-INTEGRATE framework is concerned with substantive criteria; it does *not* comprise procedural criteria but is intended to be embedded in a clearly specified health decision-making process as described, for example, in the *WHO Handbook for Guideline Development*.[22] We recognise that transparent and inclusive procedures are essential to achieve legitimate health decisions and to resolve reasonable disagreement based on competing criteria and the various individual, social, cultural and political values affecting their interpretation and the explicit or implicit weight assigned to them. In this context legitimacy refers to the reasonableness, or acceptability, of decisions as perceived by the population.[75 76] Compromised legitimacy may hinder the effective implementation of guidelines or other health-relevant decisions. Transparent and inclusive procedures require, among other considerations, the involvement of relevant stakeholders in the decision-making process, the public announcement of forthcoming decisions including their underlying argumentation, and the instalment of mechanisms for appeal.[75 76] This is relevant for the development of WHO guidelines at the global level, as well as their adaptation at the national or subnational levels, where a wide array of stakeholders with diverse sets of values should be involved.[77 78] In our overview of systematic reviews, we distilled procedural criteria (Stratil *et al*, forthcoming) and suggest that these be reviewed separately to inform guideline development and other health decision processes.[18 41 42 45 75 76 79] We also refer to evidence-informed deliberative processes, which explicitly integrate the use of substantive criteria with procedural criteria to set priorities at national and subnational levels.[80–82]

The WHO-INTEGRATE framework is a highly interdisciplinary framework: each criterion, especially those criteria that are less developed in current EtD frameworks (eg, *human rights and sociocultural acceptability*) or absent from the literature (eg, *societal implications*), merits research to unpack them and, where applicable, provide a more detailed normative justification. We anticipate constructive input from and exchange with relevant disciplines, in particular public health ethics but also sociology, environmental sciences, economics and many others. Future collaborative research is expected to lead to a WHO-INTEGRATE framework version 2.0. This may advance the criteria and subcriteria and their normative foundations, as well as methodological approaches to populate these criteria with evidence.

To examine the value of the WHO-INTEGRATE framework to potential users, we conducted empirical qualitative research. Insights from interviews with key informants in relation to their recent experiences with developing WHO guidelines led to several refinements in the wording of the criteria and subcriteria and highlighted the importance of providing example questions as well as suggested methods. We expect that the second empirical qualitative research component, focus group discussions in Nepal, Uganda, Germany and Brazil, will yield additional insights from different perspectives and possibly further modifications to the framework. An integrated analysis of the views of WHO guideline developers and users will be published separately (Stratil *et al*, forthcoming).

Several of our key informants expressed concern about the potential workload resulting from collecting evidence for each of the criteria and, in particular, for the many subcriteria in the WHO-INTEGRATE framework. Both the process of prioritisation and the process of collecting evidence—through high-quality evidence synthesis or more pragmatic approaches—need to be tested in practice. We anticipate sharing worked examples and developing additional guidance on how to implement the framework in practice.

CONCLUSIONS

The WHO-INTEGRATE framework represents a comprehensive EtD framework rooted in WHO norms and values that is, in principle, suitable for individual-level, population-level and system-level health interventions that may or may not be characterised by complexity. It offers structured definitions for each of the six substantive criteria as well as the meta-criterion quality of evidence; example

questions and suggested methods are provided to facilitate uptake. Importantly, this framework is intended to be used from the beginning and throughout a guideline or other health decision-making process, whether this process takes place at the global, national or subnational level. In working towards version 2.0, we welcome learning from the experiences of those applying the framework, as well as from researchers in disciplines concerned with the included criteria or subcriteria.

Author affiliations
[1]Institute for Medical Information Processing, Biometry and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Munich, Germany
[2]Department of Global Health, Norwegian Institute of Public Health, Oslo, Norway
[3]Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland
[4]Department of Information, Evidence and Research, World Health Organization, Geneva, Switzerland
[5]Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands

Acknowledgements In our search for a suitable name that would transport the ideas contained within the new EtD framework, we adopted a name previously coined for the EU-funded INTEGRATE-HTA project, which two authors on this paper, EAR and RB, were part of and influenced by. We would like to thank INTEGRATE-HTA colleagues who were happy for us to use the project name in a different but related context. We are grateful to all participants in the key informant interviews. We are indebted to the many individuals who gave freely of their time and who thought 'out of the box' to raise important questions about and provide constructive criticism on previous versions of this paper: Saskia den Boon, Randy Elder, Nathan Ford, Elizabeth Fox, Davina Ghersi, Claire Glenton, David Gough, Sean Grant, Thomas Harder, Mike Kelly, Etienne Langlois, Fergus Macbeth, Paul Montgomery, Ani Movsisyan, Pierre Ongolo-Zogo, Deepak Paudel, Jennifer Petcovic, Mark Petticrew, Kevin Pottie, Gerard Schmets, Elham Shakibazadeh, James Thomas, Rebekah Thomas-Bosco, Andrea Tricco, Vivian Welch, Verina Wild and Teun Zuiderent-Jerak. We would also like to thank those colleagues who specifically provided input towards the table on how decision criteria may be populated with evidence: Kristin Bakke-Lysdahl, Björn Hofmann, Stephanie Levy, Graham Moore, Jane Noyes, Wija Oortwijn, Pietro Refolo and Dario Sacchini.

Contributors EAR conceived the overall research project with substantial input from JMS. RB, IBS, AP and SLN all provided important contributions to the overall direction of the project and to the draft manuscript. JMS undertook most of the work in step 1, with EAR and RB duplicating relevant elements in the process and all three analysing the findings. EAR implemented step 2a. JMS designed step 2b with input from EAR, collected the data and jointly analysed them with IBS. EAR, with input from AP, undertook the work in step 3. All authors discussed different versions of the framework, making revisions to criteria, definitions, subcriteria and example questions in an iterative manner. EAR wrote the manuscript, with JMS drafting selected sections; all authors critically reviewed the different versions of the manuscript, suggested revisions and approved the version to be published.

Funding Funding provided by the World Health Organization Department of Maternal, Newborn, Child and Adolescent Health through grants received from the United States Agency for International Development and the Norwegian Agency for Development Cooperation. The development of this framework would not have been possible without the financial and, even more so, intellectual and procedural, support from the WHO, in particular from AP and SLN. We also gratefully acknowledge that JMS's position was funded by the Bavarian Health and Food Safety Authority, and that IBS's input was funded by the Norwegian Agency for Development Cooperation (NORAD).

Disclaimer The USAID, the Bavarian Health and Food Safety Authority, and NORAD had no influence on the research process or content of this manuscript. SLN and AP are staff members of the WHO. The authors alone are responsible for the views expressed in this publication, which do not necessarily represent the decisions or policies of the WHO.

Competing interests SLN helps oversee the quality of WHO guidelines and reports being a member of the RIGHT Working Group and the GRADE Working Group. EAR is a member of the GRADE Working Group.

Patient consent Not required.

Ethics approval Ethical approval was obtained from both the WHO Ethics Review Committee and the Ethics Committee of the Ludwig-Maximilians-Universität (LMU Munich), Germany.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Smith KE, Katikireddi SV. A glossary of theories for understanding policymaking. *J Epidemiol Community Health* 2013;67:198–202.
2. Sanderson I. Intelligent policy making for a complex world: pragmatism, evidence and learning. *Polit Stud* 2009;57:699–719.
3. Rajan D, Adam T, El Husseiny D, *et al*. *Briefing Note - Policy Dialogue: What it is and how it can contribute to evidence-informed decision-making*. Geneva: WHO, 2015.
4. Burchett HE, Mounier-Jack S, Griffiths UK, *et al*. National decision-making on adopting new vaccines: a systematic review. *Health Policy Plan* 2012;27:ii62–ii76.
5. Cromwell I, Peacock SJ, Mitton C. 'Real-world' health care priority setting using explicit decision criteria: a systematic review of the literature. *BMC Health Serv Res* 2015;15:164.
6. Guindo LA, Wagner M, Baltussen R, *et al*. From efficacy to equity: Literature review of decision criteria for resource allocation and healthcare decisionmaking. *Cost Eff Resour Alloc* 2012;10:9.
7. MacLeod TE, Harris AH, Mahal A. Stated and revealed preferences for funding new high-cost cancer drugs: a critical review of the evidence from patients, the public and payers. *Patient* 2016;9:201–22.
8. Noorani HZ, Husereau DR, Boudreau R, *et al*. Priority setting for health technology assessments: a systematic review of current practical approaches. *Int J Technol Assess Health Care* 2007;23:310–5.
9. Youngkong S, Kapiriri L, Baltussen R. Setting priorities for health interventions in developing countries: a review of empirical studies. *Trop Med Int Health* 2009;14:930–9.
10. Majone G. *Evidence, Argument, and Persuasion in the Policy Process*. New Haven: Yale University Press, 1989.
11. Booth A, Moore G, Flemming K, *et al*. Taking account of context in systematic reviews and guidelines considering a complexity perspective. *BMJ Glob Health* 2019;0:e000840. doi:10.1136/bmjgh-2018-000840.
12. Pfadenhauer LM, Gerhardus A, Mozygemba K, *et al*. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. *Implement Sci* 2017;12:21.
13. Alonso-Coello P, Schünemann HJ, Moberg J, *et al*. GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016;353:i2016.
14. Goetghebeur MM, Wagner M, Khoury H, *et al*. Evidence and Value: Impact on DEcisionMaking--the EVIDEM framework and potential applications. *BMC Health Serv Res* 2008;8:270.
15. Heintz E, Lintamo L, Hultcrantz M, *et al*. Framework for systematic identification of ethical aspects of healthcare technologies: the SBU approach. *Int J Technol Assess Health Care* 2015;31:124–30.
16. Marckmann G, Schmidt H, Sofaer N, *et al*. Putting public health ethics into practice: a systematic framework. *Front Public Health* 2015;3:23.
17. Norheim OF, Baltussen R, Johri M, *et al*. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Eff Resour Alloc* 2014;12:18.
18. Tannahill A. Beyond evidence--to ethics: a decision-making framework for health promotion, public health and health improvement. *Health Promot Int* 2008;23:380–90.
19. Tromp N, Baltussen R. Mapping of multiple criteria for priority setting of health interventions: an aid for decision makers. *BMC Health Serv Res* 2012;12:454.

20. Neumann I, Brignardello-Petersen R, Wiercioch W, *et al*. The GRADE evidence-to-decision framework: a report of its testing and application in 15 international guideline panels. *Implement Sci* 2016;11:93.
21. Goetghebeur MM, Wagner M, Khoury H, *et al*. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. *Med Decis Making* 2012;32:376–88.
22. World Health Organization (WHO). *WHO Handbook for Guideline Development*. 2nd ed. Geneva: World Health Organization, 2014.
23. Alonso-Coello P, Oxman AD, Moberg J, *et al*. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ* 2016;353:i2089.
24. Vogel JP, Oxman AD, Glenton C, *et al*. Policymakers' and other stakeholders' perceptions of key considerations for health system decisions and the presentation of evidence to inform those considerations: an international survey. *Health Res Policy Syst* 2013;11:19.
25. Treweek S, Oxman AD, Alderson P, *et al*. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. *Implement Sci* 2013;8:6.
26. Kenny NP, Melnychuk RM, Asada Y. The promise of public health: ethical reflections. *Can J Public Health* 2006;97:402–4.
27. Ki-Moon B. *The road to dignity 2030: ending poverty, transforming all lives and protecting the planet. Synthesis report of the Secretary-General on the post 2015 sustainable development agenda*. New York: United Nations, 2014.
28. Petticrew M, Knai C, Thomas J, *et al*. Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Glob Health* 2019;0:e000899. doi:10.1136/ bmjgh-2018-000899.
29. *World Health Organization (WHO). WHO recommendations on antenatal care for a positive pregnancy experience*. Geneva: World Health Organization, 2014.
30. *World Health Organization (WHO). WHO recommendations: optimizing health worker roles to improve access to key maternal and newborn health interventions through task shifting*. Geneva: World Health Organization, 2012.
31. Beauchamp T, Childress J. *Principles of biomedical ethics*. Oxford University Press: Oxford, 1994.
32. Keeling M, Bellefleur O. *'Principlism' and Frameworks in Public Health Ethics*. National Collaborating Centre for Healthy Public Policy: Montréal, 2016.
33. World Health Organization (WHO). *Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference*. New York: signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force, 1946. (on 7 Apr 1948).
34. Committee on Economic Social and Cultural Rights (CESCR). *CESCR General Comment No. 14: The Right to the Highest Attainable Standard of Health (Art. 12) Adopted at the Twenty-second Session of the Committee on Economic, Social and Cultural Rights*: (Contained in Document E/C.12/2000/4) Office of the High Comissioner for Human Rights, 2000. (on 11 August 2000).
35. World Health Organization (WHO). *Rio Political Declaration on Social Determinants of Health - Rio de Janeiro*. Brazil: Rio de Janeiro: World Health Organization, 2011. (21 Oct 2011).
36. Bustreo F, Women's HP. *Children's Health: Evidence of Impact of Human Rights*. Geneva: World Health Organization, 2013.
37. *Comission on Social Determinants of Health (CSDH). Closing the gap in a generation: Health equity through action on the social determinants of health. Final Report of the Comission on Social Determinants of Health*. Geneva, Switzerland: World Health Organization, 2008.
38. World Health Organization (WHO). *The world health report 2000 - Health systems: improving performance*. Geneva: World Health Organization, 2000.
39. World Health Organization (WHO). *The world health report 2008 - Primary health care: now more than ever*. Geneva: World Health Organization, 2008.
40. World Health Organization (WHO). *Everybody's Business: Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action*. Geneva: World Health Organization, 2007.
41. Childress JF, Faden RR, Gaare RD, *et al*. Public health ethics: mapping the terrain. *J Law Med Ethics* 2002;30:170–8.
42. Kass NE. An ethics framework for public health. *Am J Public Health* 2001;91:1776–82.
43. ten Have M, de Beaufort ID, Mackenbach JP, *et al*. An overview of ethical frameworks in public health: can they be supportive in the evaluation of programs to prevent overweight? *BMC Public Health* 2010;10:638.
44. Lee LM. Public health ethics theory: review and path to convergence. *J Law Med Ethics* 2012;40:85–98.
45. Krebs K, Urwin J. *Public health: Ethical issues*. London: Nuffield Council on Bioethics, 2009.
46. Petrini C. Theoretical models and operational frameworks in public health ethics. *Int J Environ Res Public Health* 2010;7:189–202.
47. Thompson AK, Faith K, Gibson JL, *et al*. Pandemic influenza preparedness: an ethical framework to guide decision-making. *BMC Med Ethics* 2006;7:12.
48. Upshur RE. Principles for the justification of public health intervention. *Can J Public Health* 2002;93:101–3.
49. Baum NM, Gollust SE, Goold SD, *et al*. Looking ahead: addressing ethical challenges in public health practice. *J Law Med Ethics* 2007;35:657–67.
50. Swain GR, Burns KA, Etkind P. Preparedness: medical ethics versus public health ethics. *J Public Health Manag Pract* 2008;14:354–7.
51. Nations U. *United Nations General Assembly A/69/L.6 Draft outcome document of the United Nations summit for the adoption of the post-2015 development*. agenda, 2015.
52. World Health Organization (WHO). *A69/15. Health in the 2030 Agenda for Sustainable Development. 08*. Geneva: World Health Organization, 2016.
53. WHO. *WHO A69/11 Health in the 2030 Agenda for Sustainable Development*, 2016.
54. World Health Organization (WHO). *EB138.R5. Strengthening essential public health functions in support of the achievement of universal health coverage*. Geneva: World Health Organization, 2016.
55. Mobinizadeh M, Raeissi P, Nasiripour AA, *et al*. The health systems' priority setting criteria for selecting health technologies: A systematic review of the current evidence. *Med J Islam Repub Iran* 2016;30:329–29.
56. Abbasi M, Majdzadeh R, Zali A, *et al*. The evolution of public health ethics frameworks: systematic review of moral values and norms in public health policy. *Med Health Care Philos* 2017.
57. Scott AM, Hofmann B, Gutiérrez-Ibarluzea I, *et al*. Q-SEA - a tool for quality assessment of ethics analyses conducted as part of health technology assessments. *GMS Health Technol Assess* 2017;13.
58. Hultcrantz M, Rind D, Akl EA, *et al*. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 2017;87:4–13.
59. Maeckelberghe ELM, Schröder-Bäck P. Ethics in public health: call for shared moral public health literacy. *Eur J Public Health* 2017;27:49–51.
60. Rehfuess EA, Akl EA. Current experience with applying the GRADE approach to public health interventions: an empirical study. *BMC Public Health* 2013;13:13.
61. Lewin S, Hendry M, Chandler J, *et al*. Assessing the complexity of interventions within systematic reviews: development, content and use of a new tool (iCAT_SR). *BMC Med Res Methodol* 2017;17:76.
62. World Health Organization (WHO). *Consolidated guideline on sexual and reproductive health and rights of women living with HIV*. Geneva: World Health Organization, 2017.
63. World Health Organization (WHO). *Communicating Risk in Public Health Emergencies - A WHO Guideline for Emergency Risk Communication (ERC) policy and practice*. Geneva: World Health Organization, 2017.
64. Mayring P. *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Lagenfurt: SSOAR, 2014.
65. Kieny MP, Bekedam H, Dovlo D, *et al*. Strengthening health systems for universal health coverage and sustainable development. *Bull World Health Organ* 2017;95:537–9.
66. Siegfried N, Narasimhan M, Kennedy CE, *et al*. Using GRADE as a framework to guide research on the sexual and reproductive health and rights (SRHR) of women living with HIV - methodological opportunities and challenges. *AIDS Care* 2017;29:1088–93.
67. Alonso-Coello P, Schünemann HJ, Moberg J, *et al*. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016;353:i2016.
68. EUnetHTA Joint Action 2 WP. *HTA Core Model ® version 3.0*, 2016.

69. Johri M, Norheim OF. Can cost-effectiveness analysis integrate concerns for equity? Systematic review. *Int J Technol Assess Health Care* 2012;28:125–32.
70. Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
71. GRADE Working Group. GRADE Working Group Webpage: GRADE;. 2017 http://www.gradeworkinggroup.org/ (accessed 20 Sep 2017).
72. Lewin S, Glenton C, Munthe-Kaas H, *et al*. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med* 2015;12:e1001895.
73. Montgomery P, Movsisyan A, Grant SP, *et al*. Considerations of complexity in rating certainty of evidence in systematic reviews: a primer on using the GRADE approach in global health. *BMJ Glob Health* 2019;0:e000848. doi:10.1136/ bmjgh-2018-000848.
74. Morgan RL, Kelley L, Guyatt GH, *et al*. Decision-making frameworks and considerations for informing coverage decisions for healthcare interventions: a critical interpretive synthesis. *J Clin Epidemiol* 2018;94:143–50.
75. Daniels N. Accountability for reasonableness Establishing a fair process for priority setting is easier than agreeing on principles. *BMJ* 2000;321:1300–1.
76. Daniels N, Sabin J. Limits to health care: fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philos Public Aff* 1997;26:303–50.
77. Tantivess S, Tangcharoensathien V. Coverage decisions and the court: a public health perspective on glucosamine reimbursement in Thailand. *Health Syst Reform* 2016.
78. Baltussen R, Jansen MPM, Bijlmakers L, *et al*. Value assessment frameworks for hta agencies: the organization of evidence-informed deliberative processes. *Value Health* 2017;20:256–60.
79. Krebs K, Unwin J. *Public health: ethical issues*. London: Nuffield Council on Bioethics, 2007.
80. Baltussen R, Jansen MP, Mikkelsen E, *et al*. Priority setting for universal health coverage: we need evidence-informed deliberative processes, not just more evidence on cost-effectiveness. *Int J Health Policy Manag* 2016;5:615–8.
81. Baltussen R, Jansen MP, Bijlmakers L, *et al*. Progressive realisation of universal health coverage: what are the required processes and evidence? *BMJ Glob Health* 2017;2:e000342.
82. Tromp N, Prawiranegara R, Siregar A, *et al*. Translating international HIV treatment guidelines into local priorities in Indonesia. *Trop Med Int Health* 2018;23:279–94.
83. In: Higgins J, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. Chichester: The Cochrane Collaboration and John Wiley & Sons Ltd, 2008.
84. Flemming K, Booth A, Garside R, *et al*. Qualitative evidence synthesis for complex interventions and guideline development: clarification of the purpose, designs and relevant methods. *BMJ Glob Health* 2019;0:e000882. doi:10.1136/ bmjgh-2018-000882.
85. Booth A, Noyes J, Flemming K, *et al*. Seven RETREAT criteria inform selection from 19 qualitative evidence synthesis approaches. *J ClinEpidemiol* 2018;99:41–52.
86. Noyes J, Booth A, Moore G, *et al*. Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Glob Health* 2019;0:e000893. doi:10.1136/ bmjgh-2018-000893.
87. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8:19–32.
88. Gough D, Oliver S, Thomas J. *An introduction to systematic reviews*: SAGE Publishing, 2017.
89. Tricco AC LE, Straus SE, eds. *Rapid reviews to strengthen health policy and systems: a practical guide*. Geneva: World Health Organization, 2017.
90. Khangura S, Konnyu K, Cushman R, *et al*. Evidence summaries: the evolution of a rapid review approach. *Syst Rev* 2012;1:10.
91. Kaltenthaler E, Cooper K, Pandor A, *et al*. The use of rapid review methods in health technology assessments: 3 case studies. *BMC Med Res Methodol* 2016;16:108.
92. Pollock M, Fernandes RM, Becker LA, *et al*. What guidance is available for researchers conducting overviews of reviews of healthcare interventions? A scoping review and qualitative metasummary. *Syst Rev* 2016;5:190.
93. Hofmann B, Droste S, Oortwijn W, *et al*. Harmonization of ethics in health technology assessment: a revision of the Socratic approach. *Int J Technol Assess Health Care* 2014;30:3–9.
94. Lysdahl K, Mozygemba K, Burns J, *et al*. Guidance for assessing effectiveness, economic aspects, ethical aspects, socio-cultural aspects and legal aspects in complex technologies [Online]. 2016 http://www.integrate-hta.eu/downloads/ ([7th Nov 2017).
95. *The Danish Institute for Human Rights*. *Human rights impact assessment - guidance and toolbox*. Copenhagen: The Danish Institute for Human Rights, 2016.
96. Mertz M, Kahrass H, Strech D. Current state of ethics literature synthesis: a systematic review of reviews. *BMC Med* 2016;14:152.
97. Droste S, Dintsios CM, Gerber A. Information on ethical issues in health technology assessment: how and where to find them. *Int J Technol Assess Health Care* 2010;26:441–9.
98. Booth A, Noyes J, Flemming K, *et al*. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Glob Health* 2019;0:e001107. doi:10.1136/ bmjgh-2018-001107.
99. Lewin S, Booth A, Glenton C, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implement Sci* 2018;13:2.
100. Pottie K, Welch V, Morton R, *et al*. GRADE equity guidelines 4: considering health equity in GRADE guideline development: evidence to decision process. *J Clin Epidemiol* 2017;90:84–91.
101. O'Neill J, Tabish H, Welch V, *et al*. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol* 2014;67:56–64.
102. Campbell and Cochrane Equity Methods Group. PROGRESS-PLUS. 2017 http://methods.cochrane.org/equity/projects/evidence-equity/progress-plus (accessed 7th Nov 2017).
103. Welch V, Akl E, Pottie K, *et al*. GRADE equity guidelines 3: health equity considerations in rating the certainty of synthesized evidence. *Journal of Clinical Epidemiology* 2017.
104. Mauskopf JA, Sullivan SD, Annemans L, *et al*. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis. *Value Health* 2007;10:336–47.
105. World Health Organization (WHO). *WHO guide to identifying the economic consequences of disease and injury*. Geneva: World Health Organization, 2009.
106. Drummond M, Sculpher M, Claxton K, *et al*. *Methods for the economic evaluation of health care programmes*. 4th ed. Oxford: Oxford University Press, 2015.
107. Tan-Torres Edejer T, Baltussen R, Adam T, *et al*. *Making Choices in Health. WHO Guide to Cost-Effectiveness Analysis*. Geneva: World Health Organization, 2003.
108. van Mastrigt GA, Hiligsmann M, Arts JJ, *et al*. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3). *Expert Rev Pharmacoecon Outcomes Res* 2016;16:689–704.
109. Thielen FW, Van Mastrigt G, Burgers LT, *et al*. How to prepare a systematic review of economic evaluations for clinical practice guidelines: database selection and search strategy development (part 2/3). *Expert Rev Pharmacoecon Outcomes Res* 2016;16:705–21.
110. Wijnen B, Van Mastrigt G, Redekop WK, *et al*. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: data extraction, risk of bias, and transferability (part 3/3). *Expert Rev Pharmacoecon Outcomes Res* 2016;16:723–32.
111. Shemilt I, McDaid D, Marsh K, *et al*. Issues in the incorporation of economic perspectives and evidence into Cochrane reviews. *Syst Rev* 2013;2:83.
112. Brunetti M, Shemilt I, Pregno S, *et al*. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol* 2013;66:140–50.
113. World Health Organization (WHO). *Systems thinking for health systems strengthening*. Geneva: World Health Organization, 2009.

6. Paper II: Advancing the WHO-INTEGRATE Framework as a Tool for Evidence-Informed, Deliberative Decision-Making Processes: Exploring the Views of Developers and Users of WHO Guidelines

Stratil JM, Paudel D, Setty KE, et al (2020): Advancing the WHO-INTEGRATE Framework as a Tool for Evidence-Informed, Deliberative Decision-Making Processes: Exploring the Views of Developers and Users of WHO Guidelines. Int J Health Policy Manag. doi: 10.34172/ijhpm.2020.19359

http://ijhpm.com
Int J Health Policy Manag 2020, x(x), 1–13

doi 10.34172/ijhpm.2020.193

IJHPM International Journal of Health Policy and Management

Original Article

Advancing the WHO-INTEGRATE Framework as a Tool for Evidence-Informed, Deliberative Decision-Making Processes: Exploring the Views of Developers and Users of WHO Guidelines

Jan M. Stratil[1,2*], Deepak Paudel[3], Karen E. Setty[4], Carlos E. Menezes de Rezende[5,6], Aline A. Monroe[6], Jimmy Osuret[7], Inger B. Scheel[8], Manfred Wildner[9,2], Eva A. Rehfuess[1,2]

Abstract

Background: Decision-making on matters of public health and health policy is a deeply value-laden process. The World Health Organization (WHO)-INTEGRATE framework was proposed as a new evidence-to-decision (EtD) framework to support guideline development from a complexity perspective, notably in relation to public health and health system interventions, and with a foundation in WHO norms and values. This study was conducted as part of the development of the framework to assess its comprehensiveness and usefulness for public health and health policy decision-making.

Methods: We conducted a qualitative study comprising nine key informant interviews (KIIs) with experts involved in WHO guideline development and four focus group discussions (FGDs) with a total of forty health decision-makers from Brazil, Germany, Nepal and Uganda. Transcripts were analyzed using MAXQDA12 and qualitative content analysis.

Results: Most key informants and participants in the FGDs appreciated the framework for its relevance to real-world decision-making on four widely differing health topics. They praised its broad perspective and comprehensiveness with respect to new or expanded criteria, notably regarding societal implications, equity considerations, and acceptability. Some guideline developers questioned the value of the framework beyond current practice and were concerned with the complexity of applying such a broad range of criteria in guideline development processes. Participants made concrete suggestions for improving the wording and definitions of criteria as well as their grouping, for covering missing aspects, and for addressing overlap between criteria.

Conclusion: The framework was well-received by health decision-makers as well as the developers of WHO guidelines and appears to capture all relevant considerations discussed in four distinct real-world decision processes that took place on four different continents. Guidance is needed on how to apply the framework in guideline processes that are both transparent and participatory. A set of suggestions for improvement provides a valuable starting point for advancing the framework towards version 2.0.

Keywords: Decision-Making, Priority Setting, Resource Allocation, Guideline Development, World Health Organization, Framework

Citation: Stratil JM, Paudel D, Setty KE, et al. Advancing the WHO-INTEGRATE framework as a tool for evidence-informed, deliberative decision-making processes: exploring the views of developers and users of WHO guidelines. *Int J Health Policy Manag.* 2020;x(x):x–x. doi:10.34172/ijhpm.2020.193

Article History:
Received: 30 March 2020
Accepted: 29 September 2020
ePublished: 27 October 2020

***Correspondence to:**
Jan M. Stratil
Email:
stratil@ibe.med.uni-muenchen.de

Background

Making evidence-informed decisions about public health and health system interventions and policies is complex.[1-3] On the one hand, producing and assessing evidence eg, on the effectiveness of public health and health policy interventions is challenging due to the complexity of the interventions themselves (eg, the number of components, or the pathway leading to multiple outcomes).[4] Furthermore, due to interactions with the system in which these are implemented (eg, system changes due to emergent properties, adaptivity, or feedback mechanisms) as well as due to the high context-dependency of the effects of the intervention.[2,5] On the other hand, simply producing more and stronger evidence eg, on the efficacy or cost-effectiveness of an intervention is in itself not sufficient to make better choices, as evidence-informed decision-making is a deeply value-laden process.[6-8] Decision-makers must balance numerous and often conflicting normative and technical aspects for a decision-making process,[9-11] which represents an additional source of complexity (eg, which criteria should be considered and how should these be weighed against each other)? This holds true for all forms of structured decision-making processes

Full list of authors' affiliations is available at the end of the article.

Key Messages

Implications for policy makers

- Public health and health policy processes are complex and deeply value-laden. This includes different types of decision-making processes at national or sub-national levels as well as the development of guidelines at a global level.
- The various affected stakeholder groups have their own reasons and principles guiding their decisions. Involving diverse stakeholders and taking their views into account in a structured way can ensure transparency, legitimacy, and acceptability of the decision, and increase the likelihood of implementation. Evidence-to-Decision (EtD) frameworks, such as the World Health Organization (WHO)-INTEGRATE framework, can serve as helpful tools.
- The WHO-INTEGRATE framework could be a valuable tool to support decision-making processes and, with regard to WHO guideline development, could enhance the relevance and applicability of WHO recommendations in public health and health policy. Suggestions provided will help to further advance the framework and to develop concrete guidance on how to apply it in practice.

Implications for the public

When making public health or health policy decisions, decision-makers should consider the best available scientific evidence and other factors (eg, cost, feasibility, or acceptability). They should also ensure that members in the committee preparing or making decisions is sufficiently diverse and represents all relevant viewpoints. This applies to political decisions at national or subnational levels and to more technical processes, eg, development of guidelines. When not adequately or transparently considered, decisions may not lead to the desired impacts or may not be considered acceptable and legitimate. Decision frameworks, such as the World Health Organization (WHO)-INTEGRATE framework, can support decision-makers and help ensure that all factors of relevance are considered. We discussed this framework with developers of guidelines at the WHO and with groups of decision-makers across four continents. They reported that the factors (called criteria) included in this framework are both comprehensive and relevant to real world public health and health policy decisions. This suggests that the WHO-INTEGRATE framework can be a valuable tool for application from global to local levels.

in health, notably priority setting,[6,12,13] health technology assessments,[14] and the development of guidelines.[15,16]

To support decision-makers in making informed decisions on matters of public health and health policy, the World Health Organization (WHO) provides systematically developed guidelines.[15] The recommendations set forth in these guidelines are particularly important for policy-makers and program managers in low- and middle-income countries who often have limited resources for conducting comprehensive processes of evidence gathering and analysis. Those responsible for developing WHO guidelines are challenged to balance the need for a comprehensive approach – which is indicated due to the complexity of the intervention, the challenges of evidence generation, and the multiplicity of values affected – with the necessity to provide recommendations in a timely manner and often under considerable resource constraints.

In guidelines and beyond, various approaches to integrate a range of specific considerations have been proposed. These approaches – the Accountability for Reasonableness (A4R) framework by Daniels and Sabin[17-21] among others[6,14,22,23] – emphasize the importance of transparency throughout the process, the inclusion of relevant stakeholders, the appropriate composition of the decision-making panel, and the identification and weighting of criteria to be considered, as well as the possibility of revisions and appeals. According to A4R,[17] one key aspect is the condition of relevance: the decision or recommendation must rest on evidence, reasons, and principles that all fair-minded parties can agree are relevant to deciding how to meet the diverse needs of affected stakeholders under the imposed resource constraints.[17] Such processes can increase the acceptability and perceived legitimacy of a decision[19,24,25] even if – given varying and sometimes contradictory interests – no consensus regarding the *right* selection and weighting of criteria can be achieved.[19]

Involving representatives of all relevant stakeholder groups, including community representatives (eg, citizens, patients) in the process of identifying these reasons and principles is considered ideal.[6,14,26,27] However, it is often difficult to meet this ideal due to time and resource constraints; this increases the risk of relevant criteria being overlooked. While not intended to nor able to replace stakeholder participation, Evidence-to-Decision (EtD) frameworks can support decision-makers and guideline developers in this balancing act.[28,29] EtD frameworks tend to comprise sets of criteria as well as procedural guidance. They are intended to ensure that all relevant criteria are considered, the best available evidence is assessed, and the underlying rationale is made explicit and transparent.[30] When applied well, these EtD frameworks can help identify and integrate the criteria of relevance for a given decision-making process, even if the voices of all relevant stakeholders were not heard. Their use should, however, not be misinterpreted as a justification for an unbalanced or incomplete composition of the committee preparing or making decisions. Furthermore, structured processes guided by EtD frameworks can lead to better, more rational decisions by counteracting inadequate (cognitive, emotional or social) heuristics, cognitive biases, or in-group dynamics.[31-35] Therefore, EtD frameworks should be as comprehensive as possible, which often is at odds with the constraints, and needs to be balanced against the resources and time available for developing an informed decision.

WHO uses EtD frameworks in their process to develop guidelines.[15,27] Given the reach and potential impact of the recommendations set forth in WHO guidelines, the nature of the guideline development process and the criteria used to inform recommendations set a benchmark for other uses. Both are described in the WHO guideline handbook for guideline development.[27] In formulating recommendations WHO Guideline Development Groups (GDG)[27] are asked to consider not only evidence of effectiveness but also a range

of other criteria (eg, resource implications, acceptability, feasibility[27]) derived from an early version of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) EtD framework.[27,30,35] Methods for guideline development were originally tailored to clinical interventions, and are still profoundly influenced by this field.[36,37] WHO recently commissioned a series of papers to make guideline development methods more applicable to complex public health and health system challenges.[15] In this context, a new EtD framework, the WHO-INTEGRATE framework version 1.0,[16] was developed with a strong conceptual and normative foundation[16] primarily based on an evaluation of WHO norms and values derived from key WHO documents (eg, the WHO constitution) and widely used public health ethics frameworks. To ensure the relevance of the framework, this normative approach was combined with a literature review of decision criteria used in real world decision-making,[38] and an assessment of complexity features.[16]

The WHO-INTEGRATE framework comprises six criteria – *balance of health benefits and harms*; *human rights and sociocultural acceptability*; *health equity, equality and non-discrimination*; *societal implications; financial and economic considerations*; and *feasibility and health system considerations* – as well as the meta-criterion *quality of evidence*. Each criterion encompasses a detailed definition, a set of sub-criteria, example questions to assess these sub-criteria and a methodological toolbox with suggested methods for collecting, synthesizing and appraising evidence. Applicability and benefit of the WHO-INTEGRATE framework to real-world decision-making situations remain to be tested.

The objective of this study was to assess the WHO-INTEGRATE framework version 1.0 through a wider participatory process with experts involved in developing WHO guidelines on an international level as well as decision-makers developing national guidelines and/or adapting and implementing WHO guidelines on a national level. Specifically, it served to review the framework in terms of its overall structure and specific criteria and to shed light on the comprehensiveness, relevance, and usefulness in real-world decision-making contexts.

Methods

We conducted nine key informant interviews (KIIs) with experts involved in WHO guideline development on an international level as well as four focus group discussions (FGDs) with health decision-makers in Brazil, Germany, Nepal, and Uganda.

Participants and Data Collection

Key Informant Interviews

The KIIs were conducted with experts who had recently participated in a WHO guideline development process, either as the coordinating WHO staff, GDG chair, or the methodologist supporting the process. In consultation with the Secretariat of the WHO Guideline Review Committee, three WHO guidelines – on sexual and reproductive health and rights of women living with HIV,[39] on communicating risk in public health emergencies[40] and, on antenatal care for a positive pregnancy experience[41] – were selected purposefully, with the aim to cover distinct types of interventions and to capture positive as well as challenging experiences with applying the GRADE EtD framework.[27]

In the face-to-face or telephone/video interviews carried out between June and October 2017, we used a semi-structured, pre-tested interview guide, developed based on the guiding research questions (Supplementary file 1). The first interview part focused on the experience of using the GRADE EtD framework to formulate recommendations and decide on their strength. In the second part, the interviewees received an interim version of the WHO-INTEGRATE framework and were asked to comment on practical considerations (eg, understandability), the framework content (eg, missing criteria), and the implications of using the framework in WHO guideline development processes.

The audio files of the recorded interviews were transcribed, reviewed, pseudonymized, and then deleted.

Focus Group Discussions

The FGDs were conducted with decision-makers across four countries and continents, as detailed in Table 1 and Supplementary file 2. To capture the diversity of views, we sought to maximize heterogeneity among countries (ie, country income group, region) and topics of discussion (ie, type of intervention/approach).

Local researchers undertook the FGDs in close collaboration with the developers of the framework between August 2017 and October 2018. A topic of current importance was suggested by the local researchers. Decision-makers were identified by the local contact, who reached out to a purposive sample of experts involved with making national recommendations, and potential WHO guideline users, ie, those responsible for adapting and implementing recommendations locally, regionally, or nationally. The recruitment strategy and composition of decision-makers varied across the four FGDs regarding the nature of the committee (providing advice vs. making recommendations), the level of decision-making (national vs. local) and exact composition (representative sub-set of committee vs. ad hoc assembly of participants) (see Table 1 for details).

We developed a preliminary interview guide based on the KII interview guide and adapted it to the setting and topic of the FGDs (Supplementary file 1). The FGDs were set up as a thought experiment: First, the participants conducted a guided brainstorming session on criteria of relevance for the decision-making process on their chosen topic, at which point they were unaware about the content of the WHO-INTEGRATE framework. This was one in order to not have the discussion of criteria and considerations of relevance be "contaminated" or framed by the content of the framework. Second, they were presented with an interim version of the framework and asked to review whether the framework covered the previously discussed criteria, whether aspects were missing, and whether there might be specific suggestions for improvement.

The files of the audio-recorded FGDs were transcribed and the transcripts reviewed by the local researchers. Transcripts

Table 1. Countries, Thematic Areas and Topics of FGDs

Country	Brazil	Germany	Nepal	Uganda
Thematic area	Infectious diseases; healthcare system	Public health nutrition; non-communicable diseases	Sexual and reproductive health; health services research	WaSH; infectious diseases
Topic of FGD	Tuberculosis guidelines and decentralized actions related to tuberculosis control	(Health) implications of an elimination of an EU quota system on isoglucose and considerations regarding countermeasures (eg, labelling, taxation and/or prohibition of products)	Health services related to sexual reproductive health and rights of adolescents	Management of untreated wastewater, including sewage from septic tanks and fecal sludge from pit latrines
Country income group	Middle-income country	High-income country	Low-income country	Low-income country
WHO region	Latin American Region	European Region	South East Asian Region	African Region
Researcher(s) conducting FGD	AAM, CEMR	JMS	DP	JO, KS
Date of FGD	June 2018	June 2018	October 2017	August 2017
Number of participants	n = 17	n = 7	n = 8	n = 8
Characteristics of participants and rationale	Multidisciplinary staff and invited members of National Coordination for Tuberculosis Control Program directly dealing with national policies implementation, decision-making and public health protocols design and adaptation	Staff of the Bavarian Health and Food Safety Authority across several departments advising on and preparing decisions regarding food safety and food regulation on the level of a German federal state	National level experts from governmental institutions (eg, divisions of ministry of health) and Nepali experts from national and international NGOs working on development and implementation of programs of and rights ASRH	Members of a national workgroup on water and sanitation from diverse organizational sectors, including national and sub-national government, nongovernmental/civil society, NGO network, and private actors with expertise in water and sanitation guideline implementation; including Ugandan representative of the Sanitation and Water for All network and WHO sanitation guideline development; (note: foreign aid and multilateral organizations were excluded)
Recruitment approach	Direct contact with implementation science expert and researchers working on the topic to identify a diverse set of key experts; invitation of identified experts by the local researchers directly via email.	Direct contact (personal, via telephone or email) of staff members involved with analyzing the implications of food safety or providing guidance on such countermeasures within the Bavarian Health and Food Safety Authority	Direct contact (in person, telephone) to individual experts from the federal ministry of health and individual national and international NGOs; snowballing recruitment of additional experts in the field through the directly contacted experts	In-person recruitment of key contacts at professional conferences followed by more extensive email recruitment within national water and sanitation work group; snowballing recruitment through referral to other potential participants in their networks of professionals doing work on national WaSH issues
Duration FGD	130 minutes	95 minutes	100 minutes	120 minutes
Setting of data collection	Brasília, Brazil Conference room in the building of the National Coordination for the National Tuberculosis Control Program	Munich, Germany Conference room in the building of the Bavarian Health and Food Safety Authority (1 participant via video link)	Kathmandu, Nepal Conference room at centrally located hotel	Kampala, Uganda Conference room at WHO offices
Language of FGD/ analysis	Portuguese/English	German/German	Nepali/English	English/English

Abbreviations: WaSH, Water, sanitation, and hygiene; FGD, focus group discussion; EU, European Union; WHO, World Health Organization; ASRH, adolescent sexual and reproductive health and rights; NGO, non-governmental organization.

were subsequently pseudonymized, translated into English (Brazil and Nepal) and reviewed after translation in the light of the audio-records. Audio-records were then deleted.

Data Analysis

A two-person team (JMS and IS, ST or KK) analyzed the pseudonymized transcripts (KIIs and FGDs) through qualitative content analysis, following the approach by Mayring.[42] Findings for the FGDs were provided to local contacts for feedback, clarification, and discussion. The analysis of the transcript of the FGD conducted in Germany was not translated but rather analyzed in German by two native speakers (JMS and KK). We followed mixed deductive and inductive approaches to develop the coding frames (Supplementary file 3) using the software MAXQDA12 (VERBI Software GmbH, Berlin). Furthermore, one researcher (JMS) assessed coded passages of the FGD transcripts with respect to whether aspects considered relevant for decision-making on the topic and theme in question were covered by the criteria and sub-criteria of the WHO-INTEGRATE framework.

Results

We conducted nine KIIs with WHO guideline staff (n = 2), GDG chairs (n = 4), and methodologists (n = 3) with a median duration of 62 minutes (range 57-69 minutes). Two additionally intended interviews did not take place, as one participant had retired (and therefore declined) and one interview could not be scheduled despite repeated attempts. The duration of the four FGDs with between seven and seventeen participants ranged from 95-150 minutes (Table 1, Supplementary file 2). The topics were: tuberculosis guidelines and decentralized actions related to tuberculosis control (Brazil), the (health) effects of an increase of isoglucose in food and potential countermeasures (Germany), health services related to sexual reproductive health and rights of adolescents (Nepal), and the management of untreated wastewater, including sewage from septic tanks and fecal sludge from pit latrines (Uganda).

Overview of Focus Group Discussions

For conciseness, we aim to provide a synthesis rather than detailed account of all four FGDs here. Supplemental data is available by contacting the authors. The next section outlines the first phase of the discussions, as this varied, and summarizes selected main themes.

Focus Group Discussion in Brazil

The FGD in Brazil was concerned with the development of tuberculosis guidelines and manuals within the Brazilian national plan to fight tuberculosis.

During the first phase, participants discussed their experiences with the development of this plan, covering both general challenges in guideline development (eg, alignment with national and supranational strategies) and specific considerations in making recommendations; most often they referred to testing strategies as an example. Criteria of relevance brought up during this phase included affordability, availability, acceptability of healthcare access, adherence to treatments, economic and financial feasibility, cost-effectiveness, and political importance.

A central consensual theme was the need to accommodate the realities of a heterogeneous country ("multiple Brazils") comprising municipalities with high- versus low-disease burdens, and the related needs to address subsidiarity and empowerment of municipalities to develop locally adapted approaches and to target social determinants within a health in all policies approach.

A second major theme was the criterion of acceptability: participants agreed on the importance of socio-cultural acceptability, especially among those intended to implement the intervention and the intended beneficiaries, stressing that acceptability could vary greatly even on a local level (eg, between healthcare institutions) and across population groups (eg, across different ethnicities or religious groups). Participants thought it unlikely to achieve socio-cultural acceptability overall. Several stated that they were unsure about how to handle the acceptability criterion based on a lack of or false knowledge among those rejecting an intervention (eg, the belief that a vaccination was developed to kill elderly).

An important point of controversy related to the question whether a separate sub-criterion regarding the right to health needed to be added. The argument in favour of adding a sub-criterion was the framework's focus on health, the argument against doing so was that the right to health is already covered within the broader sub-criterion on human rights.

Focus Group Discussion in Germany

The FGD in Germany focused on food safety and food regulation, notably the expiration of an quota system of the European Union (EU) on the market share of isoglucose, which is expected to increase high-fructose corn syrup (HFCS) in foods and beverages. Due to concerns about adverse health effects of HFCS among the general public and parliamentarians, participants discussed whether the Bavarian Health and Food Safety Authority should issue a recommendation on countermeasures.

During a first phase, participants were presented with a rapid literature review, concluding that there is an ongoing controversy regarding adverse health effects of HFCS. The participants then engaged in lively discussions on whether countermeasures such as labelling, taxation and/or prohibition of products should be taken.

A central theme in the debate was the legal feasibility of such measures and lack of clarity regarding the responsible political level (ie, federal state, nation state, EU): participants discussed whether it would be in line with EU and national regulations and law if the Bavarian government would adopt and implement regulations.

A major point of controversy was the need to balance the expected avoidance of harm, the (lack of) certainty of evidence, and the intrusiveness of the intervention. Some participants argued that in view of inconclusive evidence of harm the government does not have a mandate to act. Others argued with the precautionary principle, which allows enacting regulations to protect public health despite

unresolved uncertainties.

A further controversy was on the role of evidence in decision-making: methodological and ethical challenges can be prohibitive in proving harm beyond a reasonable doubt. Conflicts of interest can distort the evidence, both with respect to industry-funded research and evidence generated by researchers with ideological or personal interests (white hat bias[43]).

Focus Group Discussion in Nepal

The FGD in Nepal focused on adolescent sexual and reproductive health and rights (ASRH). During the first phase, participants voiced their views on important criteria for developing and implementing guidelines focused on ASRH. These included: the capacity of healthcare providers, privacy, and user-friendliness. The discussion then shifted towards experience with adapting international guidelines (eg, from WHO) and implementing ASRH programs. While international guidelines were considered useful for procuring resources from the government, the group agreed that simply transferring global guideline recommendations to local realities can be challenging (eg, due to limited acceptability or resources).

A major topic of discussion was the issue of socio-cultural acceptability. Importantly, guideline recommendations and programs cannot achieve their intended goals if they do not meet the needs and expectations of adolescents. Participants suggested that this could be achieved by engaging adolescents in developing guidelines and thus creating ownership of the program. Acceptability also encompassed community norms and values, eg, regarding gender issues. If these are not considered, a program's effectiveness and implementation would suffer.

Another central theme was the need to take local realities into account with respect to feasibility considerations. This covers locally available infrastructure as well as financial, technical and human resources. Local coordination within and beyond the health sector was regarded as essential.

Focus Group Discussion in Uganda

The FGD in Uganda focused on the management of untreated wastewater, including sewage from septic tanks and fecal sludge from pit latrines within the larger context of the WHO guidelines on sanitation and health.

During the first phase, participants discussed the importance of managing wastewater, septage and sludge, as well as the reasons for developing (international) guidelines on the topic and for potentially opposing such guidelines. Participants brought up considerations with respect to health implications of the measure, implementation and maintenance costs, and feasibility considerations.

A central theme was the interlinkage between financial costs, resource availability, feasibility, and acceptability of the intervention. Participants agreed that these aspects needed to be reflected both from the perspective of the national as well as the local government and end users. Guideline recommendations might be rejected if they were regarded as too imposing in terms of cost and resource claims or regarded as unachievable under local circumstances.

A related central topic was the need to consider resource requirements broadly in guideline recommendations: These need to reflect the required institutional infrastructure beyond the immediate needs for implementing and maintaining the intervention, such as infrastructure for planning, budgeting, or procuring resources as well as monitoring and evaluation.

Differences and Commonalities Across the FGDs

Health implications of interventions were discussed in all four FGDs, as was the need for multisectoral collaboration, implications for the health system, and consequences beyond the health sector.

While all FGDs emphasized the importance of evidence (regarding effectiveness), the debate in Germany focused on the trustworthiness of the evidence, while the other FGDs emphasized transferability and generalizability from an international level to local realities.

Socio-cultural acceptability of the intervention was discussed across all FGDs but the focus on different stakeholder groups varied: the FGD in Nepal concentrated on intended beneficiaries and the general population, the FGDs in Brazil and Uganda discussed the acceptability among those implementing and those intended to benefit from the intervention, and the German FGD was focused on political acceptability.

Furthermore, the need to consider local perspectives was raised in all FGDs. In Brazil, discussions primarily regarded the different needs and realities of municipalities, in Uganda and Nepal this focus lay on acceptability, resources availability, and feasibility, and in Germany with a concern was legal feasibility (of passing laws on the level of federal states). In this context, participants in the FGDs in Brazil, Nepal and Uganda emphasized procedural considerations in guideline development, highlighting the need to involve affected stakeholders in the process.

The importance of international guidelines and recommendations (eg, from WHO) were discussed in the FGDs in Brazil, Nepal, and Uganda, eg, regarding their usefulness in developing local guidelines or gaining political support. These were not addressed in the German FGD. In Nepal and Uganda, there was limited controversy: FGD participants seemed to strive for consensus, eg, on regarding missing criteria. In contrast, within the FGD in Germany and, to a lesser extent, in Brazil, discussions were more controversial. To some degree this may be due to cultural norms (eg, regarding deference to authority or conflict tolerance).

General Reception of the WHO-INTEGRATE Framework

The majority of interviewees in the KIIs made positive remarks about the framework, notably its usefulness and comprehensiveness. One participant remarked that the new framework covers many important issues and expressed a clear preference for this framework compared to the one in the WHO guideline handbook.[27] Several interviewees

explicitly stated that the criteria covered in the framework are important and none could or should be dropped in order to reduce the workload. Explicit positive statements were made regarding the new or expanded criteria *Societal implications, Balance of health benefits and harms, Equity, equality and non-discrimination, Human rights and socio-cultural acceptability,* and *Feasibility and health system considerations* and their sub-criteria, for example:

> *"I think that the framework in its new form with this additional guidance is really informative, and useful, and helpful to participants in these panels and hopefully leads to good recommendations"* (KII_Methodologist).

One interviewee felt the WHO-INTEGRATE framework did not go far enough by following the same approach as the GRADE EtD framework, namely starting with a defined intervention, gathering evidence and deciding whether a recommendation should be made; accordingly, a more appropriate approach might be to focus on beneficiaries and ask what should be done to improve health and well-being. Another interviewee remarked that focusing on high quality, quantitative evidence of effectiveness for a clearly defined intervention and outcome may not be feasible for complex interventions.

It was also noted that many aspects in the framework were highly context-dependent and may therefore be less applicable in the development of global guidelines. Furthermore, two interviewees questioned the added value beyond current practice and implied that all newly added sub-criteria could also be addressed as part of the GRADE EtD framework. One interviewee, however, recognized that more explicit sub-criteria could function as "signposts" for less experienced methodologists:

> *"I like the idea of making it more explicit so that you do think of these things. But if you're quite a high-level expert, you would automatically do that […]"* (KII_Methodologist).

Participants in all four FGDs made positive remarks regarding the framework and its criteria, notably their comprehensiveness. They explicitly mentioned the importance of separating individual and population perspectives regarding health benefits and harms, the range of feasibility considerations, and the broad perspective beyond mere health implications of an intervention. No general critical remarks about the framework were made in the FGDs.

Suggestions Towards Modifying the WHO-INTEGRATE Framework

Table 2 provides an overview of suggestions for improvement derived from the KIIs or FGDs.

Wording and Definitions

Participants in the KIIs and FGDs made several specific suggestions to expand upon and offer more guidance on selected criteria and sub-criteria; notably the criterion *Societal implications* was described as "fuzzy and vague" along with the sub-criteria *Accordance with human rights, Environmental implications* and *Intrusiveness of the intervention.*

> *"These [criteria] [...] 'impact on health system,' 'social impact,' they are very vague"* (KII_WHO-Staff).

Other suggestions included rewording "impact" to "implications," distinguishing affordability more clearly from financial considerations, and clarifying the types of stakeholders that should be considered with respect to acceptability.

Missing Aspects

Several interviewees stated explicitly that no criterion seemed to be missing in the framework; others suggested that the framework might not be sufficiently conducive for reflecting on underserved populations and vulnerable groups. They further recommended that a legal expert should assess whether supportive legal environments were sufficiently covered.

FGD discussants noted several potentially missing aspects including intervention sustainability, reliability and quality of an intervention, and outcomes related to well-being, for instance:

> *"The benefits […] we define as […] professional[s] and [..] that adolescent[s] would define [..] is different: The pleasure of being together with a partner, physical contacts, enjoying beer and cigarette for them is special. I am not sure if [these] benefits [are] considered"* (FGD_Nepal).

Furthermore, participants in the FGDs discussed whether political feasibility (eg, political and administrative facilitators and barriers) was sufficiently covered, in particular in regard to political feasibility on the local administrative and political level.

Order and Grouping of Criteria and Sub-criteria

Several interviewees commented on the classification of *Patients'/beneficiaries' values in relation to health outcomes* as a sub-criterion. As this is a (main) criterion in the current EtD framework,[27] some interviewees were concerned that this aspect may not receive enough attention if only addressed as a sub-criterion.

> *"I feel like [patients'/beneficiaries' values in relation to health outcomes] is not really balance of benefits and harms. [...] So I wonder if maybe this can be part of the acceptability and values. Or something like that"* (KII_WHO-Staff).

Discussants recommended a separation of *human rights and acceptability considerations* into two distinct criteria. Also, non-discrimination could be framed as a human rights consideration, rather than an aspect under *Equity and equality*. Furthermore, they suggested combining societal impact and health impact into one broad impact-oriented criterion.

Overlap, Redundancies and Delineation of Criteria and Sub-criteria

Several interviewees commented on blurred boundaries between criteria and sub-criteria, eg, between the criterion *Health equity, equality and non-discrimination* and the sub-criterion *Social impact*, between the sub-criterion *Interaction with and impact on the health system* and the criterion *Financial and economic considerations,* as well as between acceptability considerations and the sub-criterion *Patients'/*

Table 2. Overview of Suggestions for Modifications of Framework, Criteria or Sub-criteria Based on FGDs and KIIs

Criteria and sub-criteria	Suggestions for Modifications of Framework, Criteria or Sub-criteria, Based on Statements in One or More FGDs and/or KIIs			
	Wording and Definition	Missing Aspects	Order and Grouping	Overlap, Redundancy and Delineation
Balance of health benefits and harms		FGD		
Efficacy or effectiveness on health of individuals	FGD	FGD		
Efficacy or effectiveness on health of population				
Patients'/beneficiaries' values in relation to health outcomes	KII, FGD	KII	KII	
Safety-risk-profile of intervention			FGD	
Broader positive or negative health-related impacts				
Human rights and socio-cultural acceptability		FGD	FGD	
Accordance with universal human rights standards	KII, FGD			
Socio-cultural acceptability to beneficiaries and those	KII, FGD	KII		KII, FGD
Socio-cultural acceptability of intervention to the public and other stakeholders	KII, FGD	FGD		
Impact on autonomy of concerned stakeholders				
Intrusiveness of intervention	FGD	FGD		
Equity, equality and non-discrimination	FGD	KII, FGD		FGD
Impact on health equality and/or health equity				
Distribution of benefits and harms of intervention				
Affordability of intervention	KII		FGD	KII
Accessibility of intervention				
Lack of a suitable alternative				
Societal implications	KII, FGD		FGD	
Social impact	KII, FGD			KII
Environmental impact	KII, FGD			
Financial and economic considerations				
Financial impact	KII	FGD		FGD
Impact on economy		FGD		
Ratio of costs and benefits		FGD		
Feasibility and health system considerations				
Legislation		KII		FGD
Leadership and governance	FGD	FGD		
Interaction with and impact on health system	FGD	FGD		KII
Need for, usage of and impact on health workforce and human resources		FGD		
Need for, usage of and impact on infrastructure	FGD	FGD		
Quality of evidence (meta-criterion)				
Suggestions regarding missing criteria		FGD		
Suggestions regarding the order of criteria				FGD

Abbreviations: FGD, focus group discussion; KII, key informant interview.
An expanded version of this table is provided as a supplement. Supplementary file 4 details the suggested changes to the WHO-INTEGRATE framework based on KII and FGD, and Supplementary file 5 provides exemplary quotes based on the FGDs.

beneficiaries' values in relation to health outcomes.

"I think it is just equity and non-discrimination and societal impact, there are some things that are overlapping. […] How would you really delineate?" (KII_WHO-Staff).

Participants in the FGDs whether the financial and economic as well as the resource considerations were adequately delineated in light of multiple payers on several geographical levels.

Relevance of Criteria and Sub-criteria Based on Focus Group Discussions Only

Depending on the theme and topic of the FGD, different criteria dominated the discussions; nevertheless, references to all six criteria were identified in all four FGDs (Table 3, Supplementary file 6). For example, the FGD in Germany was dominated by the challenge to balance the intrusiveness of interventions and the resultant limitations inflicted on

Table 3. Overview of Passages in the FGDs Containing a Reference to a Criterion or Sub-criteria Covered by the WHO-INTEGRATE Framework or Passages Mentioning a Criterion or Sub-criteria as Relevant for a Decision-Making Process

Criteria and Sub-criteria	Brazil	Germany	Nepal	Uganda
Balance of health benefits and harms	Yes	Yes	Yes	Yes
Efficacy or effectiveness on health of individuals	Yes			
Efficacy or effectiveness on health of population	Yes			Yes
Patients'/beneficiaries' values in relation to health outcomes		Yes	Yes	
Safety-risk-profile of intervention	Yes			
Broader positive or negative health-related impacts			Yes	Yes
Human rights and socio-cultural acceptability	Yes	Yes	Yes	Yes
Accordance with universal human rights standards	Yes			Yes
Socio-cultural acceptability to beneficiaries and those	Yes		Yes	Yes
Socio-cultural acceptability of intervention to the public and other stakeholders	Yes	Yes	Yes	
Impact on autonomy of concerned stakeholders	Yes	Yes	Yes	Yes
Intrusiveness of intervention		Yes		Yes
Equity, equality and non-discrimination	Yes	Yes	Yes	Yes
Impact on health equality and/or health equity		Yes		Yes
Distribution of benefits and harms of intervention		Yes		Yes
Affordability of intervention	Yes		Yes	Yes
Accessibility of intervention	Yes	Yes	Yes	Yes
Lack of a suitable alternative		Yes		
Societal implications	Yes	Yes	Yes	Yes
Social impact			Yes	
Environmental impact				Yes
Financial and economic considerations	Yes	Yes	Yes	Yes
Financial impact	Yes		Yes	Yes
Impact on economy		Yes		Yes
Ratio of costs and benefits	Yes			Yes
Feasibility and health system considerations	Yes	Yes	Yes	Yes
Legislation	Yes	Yes	Yes	Yes
Leadership and governance		Yes		Yes
Interaction with and impact on health system	Yes			
Need for, usage of and impact on health workforce and human resources	Yes		Yes	
Need for, usage of and impact on infrastructure	Yes		Yes	
Quality of evidence (meta-criterion)	Yes	Yes	Yes	Yes

Abbreviations: FGD, focus group discussion; WHO, World Health Organization.

individual liberties with the potential health impacts and the available evidence. However, not all sub-criteria were discussed in every FGD. For example, implications for the (natural) environment were only explicitly discussed in the FGD in Uganda (focused on wastewater management). The most discussed themes included various (health) implications of the intervention, acceptability, accessibility, and autonomy and feasibility considerations. When asked, participants from all four FGD judged the framework to cover their reasoning well.

> *"I think everything here needs to be kept. You'd rather furnish the decision-makers with more information than they need than less. And as far as I'm concerned, whatever's in here would be really relevant"* (FGD_Uganda).

Implications for Using the WHO-INTEGRATE Framework in the WHO Guideline Development Process Based on Key Informant Interviews Only

Several interviewees were concerned that the complexity and the additional workload associated with the WHO-INTEGRATE framework might be overwhelming for the guideline development process. This may lead to the process merely paying lip-service to criteria, such as *Societal implications*, and skipping over important domains. Budget constraints and limited time had to be considered when applying the framework.

> *"I think that the guideline panels will find it [the expanded criteria and sub-criteria] more burdensome because to discuss all of these things will take longer. […] I think the panels get exhausted. They get tired and then they start skipping over, and they skip quite a lot"* (KII_Staff).

In contrast, one interviewee stated that "cutting corners" to reduce the workload would diminish the value of the final product. This participant emphasized the need to raise the appropriate resources for a guideline to be "done right" and that using the framework as part of a well-coordinated process would not necessarily lead to a more expensive endeavor.

Several participants stressed the need for additional

guidance, including general guidance on how to apply the WHO-INTEGRATE framework and more specific guidance on how to select and interpret criteria and sub-criteria.

Several interviewees further remarked that identifying appropriate evidence eg, for *Health system and feasibility considerations*, *Financial and Economic considerations,* or *Societal impact* might be challenging due to limited availability, low certainty and high context-dependency of evidence.

Discussion

Discussion of Key Findings

In this qualitative study, we received feedback on the WHO-INTEGRATE framework from WHO guideline developers as well as national public health and health policy decision-makers, and identified suggestions for modifications (Table 2, Supplementary file 5). Overall, the framework, its underlying conceptualization and its comprehensive nature, as well as the detailed criteria and sub-criteria were positively received. Some key informants voiced concerns regarding the implications of applying this framework with its very comprehensive set of criteria in everyday guideline development processes. A need for practical guidance was emphasized.

All criteria and sub-criteria of the WHO-INTEGRATE framework were discussed or mentioned as relevant in at least one of the four FGDs (Table 3). Moreover, interviewees and participants in FGDs commented positively on the framework's comprehensiveness. The developers of the WHO-INTEGRATE framework considered the few criteria highlighted to be missing by KII or FGD participants to be covered, pointing to a need to revise wording and provide clarification (see below). Since the FGDs and KIIs did not identify relevant gaps, this outcome aligns with the A4R framework's[17] condition of relevance, which depends on decision-makers using the framework properly. Our findings therefore suggest that the criteria included in the WHO-INTEGRATE framework can be considered both comprehensive and relevant for real-world public health decision-making.

These findings are noteworthy in view of the heterogeneity of topics and settings in the FGDs and the diverse WHO guidelines selected for the KIIs (Table 3, Supplementary file 5). Likely, the overview of systematic reviews of real world decision criteria,[16,38] undertaken to develop the WHO-INTEGRATE EtD framework played an important role: within the included systematic reviews,[9,10,44,45] KIIs[9,46,47] and FGDs[9,47-49] were employed and the criteria used or suggested for use in these studies were similar to those discussed or mentioned in our KIIs and FGDs.[16,38]

Beyond the aspects explicitly mentioned in KIIs and FGDs, the developers of the WHO-INTEGRATE framework (JMS, ER, IBS) noted some additional domains where a modification of the framework may be warranted. For example, while "individual well-being" was reported as missing in one FGD, the developers considered this aspect covered by the broad WHO concept of health, which was a key building block towards the WHO INTEGRATE framework. Similarly, the broad conceptualization of health systems[50] (beyond healthcare systems) may not have been clear to all participants and may warrant changes in wording and definitions. Furthermore, in a future revision process the sub-criteria legislation and leadership and governance may need to be more fully described and the criteria relating to availability, accessibility or lack of a suitable alternative may need to be refined with reference to complex public health interventions (eg, labeling interventions).

A concern voiced in the KIIs was the potential additional burden that the use of the WHO-INTEGRATE framework could impose on the guideline development process. This concern was not expressed in the FGDs. These different perceptions are also reflected in the literature: On the one hand, not adequately considering relevant criteria and the views of public health and health policy decision-makers as end-users was found to be a barrier to guideline adherence and implementation.[51,52] On the other hand, the balancing act between rigorous methods and finite resources, notably limited time, repeatedly emerged as an obstacle in structured decision-making processes such as guideline development,[53,54] and the necessity of pragmatic approaches is frequently emphasized.[53-55] As highlighted in the publication of the WHO-INTEGRATE framework version 1.0,[16] one solution to resolving the conflict between the comprehensiveness and granularity of the framework (both well-received) and the implications for guideline development (viewed with some concern) is to insist on a broad approach by considering all six criteria as well as the meta-criterion *quality of evidence* while allowing for much flexibility in terms of sub-criteria to be considered and evidence to be collected towards these (or not).[56]

While EtD frameworks can support decision-makers in identifying relevant criteria for public health and health policy decisions, this does not supersede the need to address the value-laden nature of these decisions in other parts of the process.[16] The substantive criteria put forward in the WHO-INTEGRATE framework need to be integrated with procedural considerations that address issues of fairness, participation, transparency and the right to appeal,[6,14,17,22] eg, as suggested in what is referred to as evidence-informed deliberative processes.[6,14,57] Major efforts should be made to achieve a balanced composition of the committee preparing for or making decisions, and to ensure representation of affected stakeholders. A balanced committee using an appropriate framework is poised to produce fair and reasonable decisions that are perceived as acceptable and legitimate – a point emphasized by interviewees and participants in the FGDs alike.

Strengths and Limitations

This study followed a comprehensive and rigorous approach to capture the perspectives of those developing WHO guidelines and those potentially adapting and implementing WHO guidelines and/or developing national recommendations across four different countries and continents. Participants represented diverse roles and backgrounds, and two modes of obtaining insights (KIIs and FGDs) were pursued. While the development of other frameworks for structured decision-making[58-62] mostly employed confirmatory approaches (eg,

surveys to rate the importance of criteria[63]), our FGDs pursued a more open-ended approach stimulating participants to reflect deeply on issues of relevance for a given topic and to discuss the framework, its structure and criteria more freely.

While the analysis was led by a researcher of German origin, data collection and analysis were conducted with local researchers to allow for an inter-cultural and interdisciplinary perspective. A thematic area and topic of relevance were proposed by local experts as the basis for the FGD in each country. Despite these strengths, the discussion represented a theoretical decision-making scenario, and criteria other than those discussed might arise in real-world decision-making. Two FGDs (Nepal and Brazil) were translated and it cannot be ruled out that nuances of the discussions were lost in translation, although efforts have been made to preserve the meaning in the process (eg, through contextualization of statements by local researchers and involvement of the local researchers in the analysis). A further limitation is that we did not systematically assess whether participants had financial or non-financial conflicts of interest in relation to the topic in all four FGDs, and analyzed how this may have affected their positioning in the discussion.

Given the multiple dimensions of heterogeneity present in the FGDs (notably in terms of countries and topics), it is difficult to assess whether saturation was reached. Due to the differences in the nature of the topics, the discussions focused on different aspects and criteria (eg, the FGD in Nepal concerned with sexual and reproductive health did not address environmental implications, while this was an important consideration in the FGD in Uganda concerned with the management of untreated wastewater).[64] Additional FGDs (among others in the WHO Western Pacific region) on an even broader set of topics might provide further considerations for advancing the framework. While we captured the perspective of a relatively small sample of users (four FGDs) and developers (nine KIIs), as discussed above, the WHO-INTEGRATE framework itself[65] builds on a much more comprehensive sample of similar real-world events.[38]

Conclusion

Our study suggests that the WHO-INTEGRATE framework can be a valuable resource for better-informed public health and health systems decisions. Reacting to the suggestions for improvement made by potential end-users, the developers are in the process of developing practical guidance for applying the WHO-INTEGRATE framework. Moreover, the applicability and added value of the framework will need to be tested in real-world guideline development and other decision-making processes, as planned with several upcoming WHO guidelines. The WHO-INTEGRATE framework was explicitly published as a living document. The findings presented here provide a valuable starting point to advance the framework towards a version 2.0.

Acknowledgements

The authors would like to thank all focus group participants and all key informants for their time and engagement. We are very grateful to Susan L. Norris, Anayda Portela, and Rebekah Thomas Bosco as well as other WHO staff for helping with the identification of suitable guidelines and key informants. Moreover, we would like to thank John Ssempebwa for facilitating the FGD in Uganda, and local assistants Monica Lima and Pedro Bossonario involved with supporting the authors AAM and CEMR in conducting the FGDs in Brazil. We are also grateful to the student assistants Sarah Thumbeck and Katja Krug for contributing to data analysis.

Ethical issues

Ethical approval was obtained from the Research Ethics Review Committees (RERC) of the WHO and the LMU Munich, as well as from local RERCs in Brazil, Nepal, and Uganda. All participants provided written and verbal consent prior to participation.

Competing interests

JMS and IBS: Developer of the WHO-INTEGRATE EtD framework. EAR: Developer of the WHO-INTEGRATE EtD framework, member of the GRADE Working Group. JMS reports grants from WHO Department of Maternal, Newborn, Child and Adolescent Health, personal fees from Bavarian Health and Food Safety Authority, other from Norwegian Agency for Development Cooperation (NORAD), during the conduct of the study.

Authors' contributions

JMS and EAR designed and directed the project with contribution from IBS. JMS developed with interview guide for KIIs with input from EAR and IBS. JMS and EAR recruited the participants with the help from Susan L. Norris, Anayda Portela, Rebekah Thomas Bosco, and other WHO staff. JMS conducted the KIIs. KII-transcripts were analyzed by JMS with contribution from IBS, ST, and KK. DP, JO; KES, CEMdR, AAM, MF suggested topic and recruited participants for the FGDs as well as prepared and facilitated focus group discussions. DP, JO, KES, CEMR, AAM, MF revised and adapted the discussion guide drafted by JMS. JS supported this process in Uganda and ML and PB supported the facilitation of the FGDs in Brazil. JMS, DP, SES, CEMR, and AAM conducted the FGDs. DP, JO, KES, CEMR, AAM transcribed, translated and contextualized the FGD-audio records. JMS analyzed the FGD transcripts with contribution with contribution from IBS, ST, and KK, as well as input from all authors. JMS took the lead in writing the manuscript with input from all authors. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Funding

The WHO-INTEGRATE EtD Framework Version 1.0 was developed with funding provided by the WHO Department of Maternal, Newborn, Child and Adolescent Health through grants received from the United States Agency for International Development and the Norwegian Agency for Development Cooperation. We also gratefully acknowledge that JMS's position was funded by the Bavarian Health and Food Safety Authority, and that IBS's input was funded by the Norwegian Agency for Development Cooperation (NORAD). The United States Agency for International Development, the Bavarian Health and Food Safety Authority and NORAD did not have any influence on the research process or content of this manuscript.

Authors' affiliations

[1]Institute for Medical Information Processing, Biometry, and Epidemiology – IBE, LMU Munich, Munich, Germany. [2]Pettenkofer School of Public Health, Munich, Germany. [3]Save the Children, Kathmandu, Nepal. [4]The Water Institute, Department of Environmental Sciences and Engineering, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [5]National Supplementary Health Agency, Ministry of Health, Brasília, Brazil. [6]Department of Maternal-Infant Nursing and Public Health, College of Nursing, University of São Paulo, Ribeirão Preto, Brazil. [7]Department of Disease Control and Environmental Health, Makerere University School of Public Health, Kampala, Uganda. [8]Department of Global Health, Norwegian Institute of Public Health, Oslo, Norway. [9]Bavarian Health and Food Safety Authority, Munich, Germany.

Supplementary files

Supplementary file 1. Interview Guide for KIIs and FGDs.
Supplementary file 2. Details on KIIs and FGDs.
Supplementary file 3. Coding Frames Used to Code Transcripts of KIIs and FGDs.

Supplementary file 4. Suggested Changes to the WHO-INTEGRATE Framework Based on KIIs and FGDs.
Supplementary file 5. Suggested Changes to the WHO-INTEGRATE Framework Based on FGDs With Exemplary Quotes.
Supplementary file 6. Coverage of Decision Aspects Stated in the FGDs by Criteria and Sub-criteria of the WHO-INTEGRATE Framework.

References

1. Sanderson I. Intelligent policy making for a complex world: pragmatism, evidence and learning. *Polit Stud*. 2009;57(4):699-719. doi:10.1111/j.1467-9248.2009.00791.x
2. Petticrew M, Knai C, Thomas J, et al. Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Glob Health*. 2019;4(Suppl 1):e000899. doi:10.1136/bmjgh-2018-000899
3. Rutter H, Savona N, Glonti K, et al. The need for a complex systems model of evidence for public health. *Lancet*. 2017;390(10112):2602-2604. doi:10.1016/s0140-6736(17)31267-9
4. Lewin S, Hendry M, Chandler J, et al. Assessing the complexity of interventions within systematic reviews: development, content and use of a new tool (iCAT_SR). *BMC Med Res Methodol*. 2017;17(1):76. doi:10.1186/s12874-017-0349-x
5. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol*. 2009;43(3-4):267-276. doi:10.1007/s10464-009-9229-9
6. Baltussen R, Jansen MP, Mikkelsen E, et al. Priority setting for universal health coverage: we need evidence-informed deliberative processes, not just more evidence on cost-effectiveness. *Int J Health Policy Manag*. 2016;5(11):615-618. doi:10.15171/ijhpm.2016.83
7. Baltussen R, Mikkelsen E, Tromp N, et al. Balancing efficiency, equity and feasibility of HIV treatment in South Africa - development of programmatic guidance. *Cost Eff Resour Alloc*. 2013;11(1):26. doi:10.1186/1478-7547-11-26
8. Kapiriri L, Lee NM, Wallace LJ, Kwesiga B. Beyond cost-effectiveness, morbidity and mortality: a comprehensive evaluation of priority setting for HIV programming in Uganda. *BMC Public Health*. 2019;19(1):359. doi:10.1186/s12889-019-6690-8
9. Guindo LA, Wagner M, Baltussen R, et al. From efficacy to equity: literature review of decision criteria for resource allocation and healthcare decisionmaking. *Cost Eff Resour Alloc*. 2012;10(1):9. doi:10.1186/1478-7547-10-9
10. Youngkong S, Kapiriri L, Baltussen R. Setting priorities for health interventions in developing countries: a review of empirical studies. *Trop Med Int Health*. 2009;14(8):930-939. doi:10.1111/j.1365-3156.2009.02311.x
11. United Nations. *Transforming our World: The 2030 Agenda for Sustainable Development*. New York: United Nations; 2015.
12. Holm S. The second phase of priority setting. Goodbye to the simple solutions: the second phase of priority setting in health care. *BMJ*. 1998;317(7164):1000-1002.
13. Kapiriri L, Martin DK. A strategy to improve priority setting in developing countries. *Health Care Anal*. 2007;15(3):159-167. doi:10.1007/s10728-006-0037-1
14. Baltussen R, Jansen MPM, Bijlmakers L, et al. Value assessment frameworks for HTA agencies: the organization of evidence-informed deliberative processes. *Value Health*. 2017;20(2):256-260. doi:10.1016/j.jval.2016.11.019
15. Norris SL, Rehfuess EA, Smith H, et al. Complex health interventions in complex systems: improving the process and methods for evidence-informed health decisions. *BMJ Glob Health*. 2019;4(Suppl 1):e000963. doi:10.1136/bmjgh-2018-000963
16. Rehfuess EA, Stratil JM, Scheel IB, Portela A, Norris SL, Baltussen R. The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. *BMJ Glob Health*. 2019;4(Suppl 1):e000844. doi:10.1136/bmjgh-2018-000844
17. Daniels N, Sabin J. The ethics of accountability in managed care reform. *Health Aff (Millwood)*. 1998;17(5):50-64. doi:10.1377/hlthaff.17.5.50
18. Daniels N. Accountability for reasonableness. *BMJ*. 2000;321:1300-1301. doi:10.1136/bmj.321.7272.1300
19. Daniels N, Sabin JE. Accountability for reasonableness: an update *BMJ*. 2008;337:a1850. doi:10.1136/bmj.a1850
20. Daniels N, Sabin J. *Just Health: Meeting Health Needs Fairly*. 1st ed. Cambridge: Cambridge University Press; 2008.
21. Daniels N, Sabin J. The ethics of accountability in managed care reform. *Health Aff (Millwood)*. 1998;17(5):50-64. doi:10.1377/hlthaff.17.5.50
22. Friedman A. Beyond accountability for reasonableness. *Bioethics*. 2008;22(2):101-112. doi:10.1111/j.1467-8519.2007.00605.x
23. Singer PA, Martin DK, Giacomini M, Purdy L. Priority setting for new technologies in medicine: qualitative case study. *BMJ*. 2000;321(7272):1316-1318. doi:10.1136/bmj.321.7272.1316
24. Tyler TR. Psychological perspectives on legitimacy and legitimation. *Annu Rev Psychol*. 2006;57:375-400. doi:10.1146/annurev.psych.57.102904.190038
25. Greenberg J, Folger R. Procedural justice, participation, and the fair process effect in groups and organizations. In: Paulus PB, eds. *Basic Group Processes*. New York, NY: Springer; 1983:235-256. doi:10.1007/978-1-4612-5578-9_10
26. Saltelli A, Giampietro M. What is wrong with evidence based policy, and how can it be improved? *Futures*. 2017;91:62-71. doi:10.1016/j.futures.2016.11.012
27. World Health Organization (WHO). *WHO Handbook for Guideline Development*. 2nd ed. Geneva, Switzerland: WHO; 2014.
28. Ham C. Priority setting in health care: learning from international experience. *Health Policy*. 1997;42(1):49-66. doi:10.1016/s0168-8510(97)00054-7
29. Niessen LW, Bridges J, Lau BD, et al. *Assessing the Impact of Economic Evidence on Policymakers in Health Care-A Systematic Review*. Rockville, MD: Agency for Healthcare Research and Quality (US); 2012.
30. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*. 2016;353:i2016. doi:10.1136/bmj.i2016
31. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc*. 2006;4:14. doi:10.1186/1478-7547-4-14
32. Nutt PC. Surprising but true: half the decisions in organizations fail. *Acad Manag Perspect*. 1999;13(4):75-90. doi:10.5465/ame.1999.2570556
33. Robinson R. Limits to rationality: economics, economists and priority setting. Health Policy. 1999;49(1-2):13-26. doi:10.1016/s0168-8510(99)00040-8
34. Gigerenzer G. Decision making: nonrational theories. In: Smelser NJ, Baltes PB, eds. *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Elsevier Science; 2001:3304-3309.
35. Moberg J, Oxman AD, Rosenbaum S, et al. The GRADE Evidence to Decision (EtD) framework for health system and public health decisions. *Health Res Policy Syst*. 2018;16(1):45. doi:10.1186/s12961-018-0320-2
36. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. 2017;390(10092):415-423. doi:10.1016/s0140-6736(16)31592-6
37. Gopinathan U, Hoffman SJ. Institutionalising an evidence-informed approach to guideline development: progress and challenges at the World Health Organization. *BMJ Glob Health*. 2018;3(5):e000716. doi:10.1136/bmjgh-2018-000716
38. Stratil JM, Baltussen R, Scheel I, Nacken A, Rehfuess EA. Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making. *Cost Eff Resour Alloc*. 2020;18:8. doi:10.1186/s12962-020-0203-6
39. World Health Organization (WHO). *Consolidated Guideline on Sexual and Reproductive Health and Rights of Women Living with HIV*. Geneva, Switzerland: WHO; 2017.
40. World Health Organization (WHO). *Communicating Risk in Public Health Emergencies: A WHO Guideline for Emergency Risk Communication (ERC) Policy and Practice*. Geneva, Switzerland: WHO; 2018.
41. World Health Organization (WHO). *WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience*. Geneva, Switzerland: WHO; 2014.

42. Mayring P. *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. Klagenfurt, Austria: Social Science Open Access Repository; 2014.
43. Cope MB, Allison DB. White hat bias: examples of its presence in obesity research and a call for renewed commitment to faithfulness in research reporting. *Int J Obes (Lond)*. 2010;34(1):84-88. doi:10.1038/ijo.2009.239
44. Cromwell I, Peacock SJ, Mitton C. 'Real-world' health care priority setting using explicit decision criteria: a systematic review of the literature. *BMC Health Serv Res*. 2015;15:164. doi:10.1186/s12913-015-0814-3
45. Noorani HZ, Husereau DR, Boudreau R, Skidmore B. Priority setting for health technology assessments: a systematic review of current practical approaches. *Int J Technol Assess Health Care*. 2007; 23(3):310-315. doi:10.1017/s026646230707050x
46. Asante AD, Zwi AB. Factors influencing resource allocation decisions and equity in the health system of Ghana. *Public Health*. 2009; 123(5):371-377. doi:10.1016/j.puhe.2009.02.006
47. Lasry A, Carter MW, Zaric GS. Allocating funds for HIV/AIDS: a descriptive study of KwaDukuza, South Africa. *Health Policy Plan*. 2011;26(1):33-42. doi:10.1093/heapol/czq022
48. Baltussen R, Stolk E, Chisholm D, Aikins M. Towards a multi-criteria approach for priority setting: an application to Ghana. *Health Econ*. 2006;15(7):689-696. doi:10.1002/hec.1092
49. Baltussen R, ten Asbroek AH, Koolman X, Shrestha N, Bhattarai P, Niessen LW. Priority setting using multiple criteria: should a lung health programme be implemented in Nepal? *Health Policy Plan*. 2007;22(3):178-185. doi:10.1093/heapol/czm010
50. World Health Organization (WHO). *Everybody's Business -- Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action*. Geneva: WHO; 2007.
51. Ryan MA. Adherence to clinical practice guidelines. *Otolaryngol Head Neck Surg*. 2017;157(4):548-550. doi:10.1177/0194599817718822
52. Fischer F, Lange K, Klose K, Greiner W, Kraemer A. Barriers and strategies in guideline implementation-a scoping review. *Healthcare (Basel)*. 2016;4(3). doi:10.3390/healthcare4030036
53. McDonald S, Elliott JH, Green S, Turner T. Towards a new model for producing evidence-based guidelines: a qualitative study of current approaches and opportunities for innovation among Australian guideline developers. *F1000Res*. 2019;8:956. doi:10.12688/f1000research.19661.1
54. Turner TJ. Developing evidence-based clinical practice guidelines in hospitals in Australia, Indonesia, Malaysia, the Philippines and Thailand: values, requirements and barriers. *BMC Health Serv Res*. 2009;9:235. doi:10.1186/1472-6963-9-235
55. Florez ID, Morgan RL, Falavigna M, et al. Development of rapid guidelines: 2. A qualitative study with WHO guideline developers. *Health Res Policy Syst*. 2018;16(1):62. doi:10.1186/s12961-018-0329-6
56. Setty K, Jiménez A, Willetts J, Leifels M, Bartram J. Global water, sanitation and hygiene research priorities and learning challenges under Sustainable Development Goal 6. *Dev Policy Rev*. 2020; 38(1):64-84. doi:10.1111/dpr.12475
57. Tromp N, Prawiranegara R, Siregar A, et al. Translating international HIV treatment guidelines into local priorities in Indonesia. *Trop Med Int Health*. 2018;23(3):279-294. doi:10.1111/tmi.13031
58. Heintz E, Lintamo L, Hultcrantz M, et al. Framework for systematic identification of ethical aspects of healthcare technologies: the SBU approach. *Int J Technol Assess Health Care*. 2015;31(3):124-130. doi:10.1017/s0266462315000264
59. Norheim OF, Baltussen R, Johri M, et al. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Eff Resour Alloc*. 2014;12:18. doi:10.1186/1478-7547-12-18
60. Johnson AP, Sikich NJ, Evans G, et al. Health technology assessment: a comprehensive framework for evidence-based recommendations in Ontario. *Int J Technol Assess Health Care*. 2009;25(2):141-150. doi:10.1017/s0266462309090199
61. Treweek S, Oxman AD, Alderson P, et al. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. *Implement Sci*. 2013;8:6. doi:10.1186/1748-5908-8-6
62. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*. 2016;353:i2016. doi:10.1136/bmj.i2016
63. Vogel JP, Oxman AD, Glenton C, et al. Policymakers' and other stakeholders' perceptions of key considerations for health system decisions and the presentation of evidence to inform those considerations: an international survey. *Health Res Policy Syst*. 2013; 11:19. doi:10.1186/1478-4505-11-19
64. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res*. 2017;27(4):591-608. doi:10.1177/1049732316665344
65. Stratil JM, Baltussen R, Scheel I, Nacken A, Rehfuess EA. Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making. *Cost Eff Resour Alloc*. 2020;18:8. doi:10.1186/s12962-020-0203-6

7. References

1. Rehfuess E, Stratil J, Scheel I, et al. The WHO-INTEGRATE evidence to decision framework version 1.0: Integrating WHO norms and values and a complexity perspective *BMJ Global Health* 2019(4):e000844. doi: 10.1136/bmjgh-2018-000844
2. Stratil JM, Paudel D, Setty KE, et al. Advancing the WHO-INTEGRATE Framework as a Tool for Evidence-Informed, Deliberative Decision-Making Processes: Exploring the Views of Developers and Users of WHO Guidelines. *J International Journal of Health Policy and Management* 2020:-. doi: 10.34172/ijhpm.2020.193
3. Stratil J, Baltussen R, Scheel I, et al. Development of the WHO-INTEGRATE evidence-to-decision framework: An overview of systematic reviews of decision criteria for health decision-making. *Cost Effectiveness and Resource Allocation* 2020;8(18)
4. Baltussen R, Jansen MP, Mikkelsen E, et al. Priority Setting for Universal Health Coverage: We Need Evidence-Informed Deliberative Processes, Not Just More Evidence on Cost-Effectiveness. *Int J Health Policy Manag* 2016;5(11):615-18. doi: 10.15171/ijhpm.2016.83
5. Ciliska D, Thomas H, Buffett C. An Introduction to Evidence-Informed Public Health and A Compendium of Critical Appraisal Tools for Public Health Practice. Hamilton: National Collaborating Centre for Methods and Tools (NCCMT) 2008.
6. Kohatsu ND, Robinson JG, Torner JC. Evidence-based public health: an evolving concept. *American journal of preventive medicine* 2004;27(5):417-21. doi: 10.1016/j.amepre.2004.07.019 [published Online First: 2004/11/24]
7. Yost J, Dobbins M, Traynor R, et al. Tools to support evidence-informed public health decision making. *BMC Public Health* 2014;14(1):728. doi: 10.1186/1471-2458-14-728
8. Norris SL, Rehfuess EA, Smith H, et al. Complex health interventions in complex systems: improving the process and methods for evidence-informed health decisions. *BMJ Global Health* 2019;4(Suppl 1):e000963. doi: 10.1136/bmjgh-2018-000963
9. World Health Organisation W. WHO Handbook for Guideline Development 2nd edition ed. Geneva, Switzerland: World Health Organization 2014.
10. Baltussen R, Jansen MPM, Bijlmakers L, et al. Value Assessment Frameworks for HTA Agencies: The Organization of Evidence-Informed Deliberative Processes. *Value in Health* 2017;20(2):256-60. doi: https://doi.org/10.1016/j.jval.2016.11.019
11. Goetghebeur MM, Wagner M, Khoury H, et al. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. *Medical decision making : an international journal of the Society for Medical Decision Making* 2012;32(2):376-88. doi: 10.1177/0272989x11416870 [published Online First: 2011/10/12]
12. Mathes T, Willms G, Polus S, et al. Health technology assessment of public health interventions: an analysis of characteristics and comparison of methods-study protocol. *Systematic reviews* 2018;7(1):79. doi: 10.1186/s13643-018-0743-4 [published Online First: 2018/05/25]
13. Refolo P, Sacchini D, Brereton L, et al. Why is it so difficult to integrate ethics in Health Technology Assessment (HTA)? The epistemological viewpoint. *European review for medical and pharmacological sciences* 2016;20(20):4202-08. [published Online First: 2016/11/11]
14. Husereau D, Boucher M, Noorani H. Priority setting for health technology assessment at CADTH. *International journal of technology assessment in health care* 2010;26(3):341-7.
15. Noorani HZ, Husereau DR, Boudreau R, et al. Priority setting for health technology assessments: A systematic review of current practical approaches. *International Journal of Technology Assessment in Health Care* 2007;23(3):310-15.
16. Rajan D, Koch K, Rohrer K, et al. Governance of the Covid-19 response: a call for more inclusive and transparent decision-making. *BMJ Global Health* 2020;5(5):e002655. doi: 10.1136/bmjgh-2020-002655
17. Pearce W. Trouble in the trough: how uncertainties were downplayed in the UK's science advice on Covid-19. *Humanities and Social Sciences Communications* 2020;7(1):122. doi: 10.1057/s41599-020-00612-w
18. Sanderson I. Intelligent Policy Making for a Complex World: Pragmatism, Evidence and Learning. *Political Studies* 2009;57(4):699-719. doi: 10.1111/j.1467-9248.2009.00791.x
19. Petticrew M, Knai C, Thomas J, et al. Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Glob Health* 2019;4(Suppl 1):e000899. doi: 10.1136/bmjgh-2018-000899 [published Online First: 2019/02/19]
20. Rutter H, Savona N, Glonti K, et al. The need for a complex systems model of evidence for public health. *The Lancet* 2017;390(10112):2602-04. doi: 10.1016/s0140-6736(17)31267-9
21. Lancaster K, Rhodes T, Rosengarten M. Making evidence and policy in public health emergencies: lessons from COVID-19 for adaptive evidence-making and intervention. *Evidence & Policy: A Journal of Research, Debate and Practice* 2020 doi: 10.1332/174426420X15913559981103

22. Stratil JM, Voss M, Arnold L. WICID framework version 1.0: criteria and considerations to guide evidence-informed decision-making on non-pharmacological interventions targeting COVID-19. 2020;5(11):e003699. doi: 10.1136/bmjgh-2020-003699 %J BMJ Global Health
23. Ioannidis JP. A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data. *STAT* 2020
24. Ioannidis JPA. Coronavirus disease 2019: The harms of exaggerated information and non-evidence-based measures. *European Journal of Clinical Investigation* 2020;50(4):e13222. doi: 10.1111/eci.13222
25. Lewin S, Hendry M, Chandler J, et al. Assessing the complexity of interventions within systematic reviews: development, content and use of a new tool (iCAT_SR). *BMC medical research methodology* 2017;17(1):76. doi: 10.1186/s12874-017-0349-x [published Online First: 2017/04/28]
26. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *American journal of community psychology* 2009;43(3-4):267-76. doi: 10.1007/s10464-009-9229-9 [published Online First: 2009/04/25]
27. Rhodes T, Lancaster K. Evidence-making interventions in health: A conceptual framing. *Social Science & Medicine* 2019;238:112488. doi: https://doi.org/10.1016/j.socscimed.2019.112488
28. Baltussen R, Mikkelsen E, Tromp N, et al. Balancing efficiency, equity and feasibility of HIV treatment in South Africa – development of programmatic guidance. *Cost Effectiveness and Resource Allocation* 2013;11(1):26. doi: 10.1186/1478-7547-11-26
29. Kapiriri L, Lee N-M, Wallace LJ, et al. Beyond cost-effectiveness, morbidity and mortality: a comprehensive evaluation of priority setting for HIV programming in Uganda. *BMC Public Health* 2019;19(1):359. doi: 10.1186/s12889-019-6690-8
30. Youngkong S, Kapiriri L, Baltussen R. Setting priorities for health interventions in developing countries: A review of empirical studies. *Tropical Medicine and International Health* 2009;14:930-39.
31. United Nations (UN). Transforming our world: the 2030 agenda for sustainable development. A/RES/70/1. New York: UN 2015.
32. Guindo LA, Wagner M, Baltussen R, et al. From efficacy to equity: Literature review of decision criteria for resource allocation and healthcare decisionmaking. *Cost Effectiveness and Resource Allocation* 2012;10
33. Abbasi M, Majdzadeh R, Zali A, et al. The evolution of public health ethics frameworks: systematic review of moral values and norms in public health policy. *Medicine, health care, and philosophy* 2018;21(3):387-402. doi: 10.1007/s11019-017-9813-y [published Online First: 2017/11/11]
34. Lee LM. Public health ethics theory: review and path to convergence. *The Journal of law, medicine & ethics : a journal of the American Society of Law, Medicine & Ethics* 2012;40(1):85-98. doi: 10.1111/j.1748-720X.2012.00648.x [published Online First: 2012/03/31]
35. Keeling M, Bellefleur O. 'Principlism' and Frameworks in Public Health Ethics. Montréal: National Collaborating Centre for Healthy Public Policy 2016.
36. Krebs K, Urwin J. Public health: Ethical issues. London: Nuffield Council on Bioethics 2009.
37. Petrini C. Theoretical models and operational frameworks in public health ethics. *International journal of environmental research and public health* 2010;7(1):189-202. doi: 10.3390/ijerph7010189 [published Online First: 2010/03/03]
38. Marckmann G, Schmidt H, Sofaer N, et al. Putting public health ethics into practice: a systematic framework. *Frontiers in public health* 2015;3:23. doi: 10.3389/fpubh.2015.00023 [published Online First: 2015/02/24]
39. Childress JF, Faden RR, Gaare RD, et al. Public health ethics: mapping the terrain. *The Journal of law, medicine & ethics : a journal of the American Society of Law, Medicine & Ethics* 2002;30(2):170-8. [published Online First: 2002/06/18]
40. Burchett H, Mounier-Jack S, Griffiths U, et al. National decision-making on adopting new vaccines: A systematic review. *Health Policy and Planning* 2012;27(SUPPL.2):ii62-ii76. doi: 10.1093/heapol/czr049
41. Cromwell I, Peacock SJ, Mitton C. Real-world health care priority setting using explicit decision criteria: A systematic review of the literature. *BMC Health Services Research* 2015;15(1)
42. MacLeod TE, Harris AH, Mahal A. Stated and Revealed Preferences for Funding New High-Cost Cancer Drugs: A Critical Review of the Evidence from Patients, the Public and Payers. *Patient* 2016;9:201-22.
43. Majone G. Evidence, Argument, and Persuasion in the Policy Process. New Haven: Yale University Press 1989.
44. Pfadenhauer LM, Gerhardus A, Mozygemba K, et al. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. *Implement Sci* 2017;12(1):21. doi: 10.1186/s13012-017-0552-5 [published Online First: 2017/02/17]
45. Daniels N, Sabin J. The Ethics Of Accountability In Managed Care Reform. *Health Affairs* 1998;17(5):50-64. doi: 10.1377/hlthaff.17.5.50
46. Daniels N, Sabin J. Just Health: Meeting Health Needs Fairly. 1st ed. Cambridge: Cambridge University Press. 2008.
47. Friedman A. Beyond accountability for reasonableness. *Bioethics* 2008;22(2):101-12. doi: 10.1111/j.1467-8519.2007.00605.x [published Online First: 2008/02/07]

48. Daniels N. Accountability for reasonableness. *BMJ* 2000;7272(321):1300-01.
49. Bærøe K, Baltussen R. Legitimate Healthcare Limit Setting in a Real-World Setting: Integrating Accountability for Reasonableness and Multi-Criteria Decision Analysis. *Public Health Ethics* 2014;7(2):98-111. doi: 10.1093/phe/phu006
50. Thompson A, Faith K, Gibson J, et al. Pandemic influenza preparedness: an ethical framework to guide decision-making. *BMC Medical Ethics* 2006;7(12)
51. Tannahill A. Beyond evidence - To ethics: A decision-making framework for health promotion, public health and health improvement. *Health promotion international* 2008;23(4):380-90. doi: 10.1093/heapro/dan032
52. Daniels N, Sabin J. The ethics of accountability in managed care reform. *Health affairs (Project Hope)* 1998;17(5):50-64. [published Online First: 1998/10/14]
53. Tyler TR. Psychological Perspectives on Legitimacy and Legitimation. *Annual Review of Psychology* 2005;57(1):375-400. doi: 10.1146/annurev.psych.57.102904.190038
54. Greenberg J, Folger R. Procedural Justice, Participation, and the Fair Process Effect in Groups and Organizations. In: Paulus P, ed. In Basic Group Processes. New York: Springer 1983:235-56.
55. Saltelli A, Giampietro M. What is wrong with evidence based policy, and how can it be improved? *Futures* 2017;91:62-71. doi: https://doi.org/10.1016/j.futures.2016.11.012
56. Edelenbos J, Klijn E-H. Managing Stakeholder Involvement in Decision Making: A Comparative Analysis of Six Interactive Processes in the Netherlands. *Journal of Public Administration Research and Theory* 2006;16(3):417-46. doi: 10.1093/jopart/mui049
57. Ham C. Priority setting in health care: learning from international experience. *Health policy (Amsterdam, Netherlands)* 1997;42(1):49-66. [published Online First: 1997/09/04]
58. Niessen LW, Bridges J, Lau BD, et al. ssessing the Impact of Economic Evidence on Policymakers in Health Care—A Systematic Review. Methods Research Report (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract No. 290-2007-10061-I). AHRQ Publication No. 12(13)-EHC133-EF. Rockville: Agency for Healthcare Research and Quality 2012.
59. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016;353:i2016.
60. Booth A, Moore G, Flemming K, et al. Taking account of context in systematic reviews and guidelines considering a complexity perspective. *BMJ Global Health* 2019;4(Suppl 1):e000840. doi: 10.1136/bmjgh-2018-000840
61. Flemming K, Booth A, Garside R, et al. Qualitative evidence synthesis for complex interventions and guideline development: clarification of the purpose, designs and relevant methods. *BMJ Glob Health* 2019;4(Suppl 1):e000882. doi: 10.1136/bmjgh-2018-000882 [published Online First: 2019/02/19]
62. Noyes J, Booth A, Moore G, et al. Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Global Health* 2019;4(Suppl 1):e000893. doi: 10.1136/bmjgh-2018-000893
63. Higgins JPT, López-López JA, Becker BJ, et al. Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health* 2019;4(Suppl 1):e000858. doi: 10.1136/bmjgh-2018-000858
64. World Health Organization (WHO). Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948. 1946
65. Kass NE. An ethics framework for public health. *American journal of public health* 2001;91(11):1776-82. [published Online First: 2001/10/31]
66. Ten Have M, De Beaufort ID, Mackenbach JP, et al. An overview of ethical frameworks in public health: can they be supportive in the evaluation of programs to prevent overweight? *BMC Public Health* 2010;10:638. doi: 10.1186/1471-2458-10-638 [published Online First: 2010/10/26]
67. Upshur R. Principles for the justification of public health intervention. *Canadian Journal of Public Health* 2002;93(2):101-03.
68. Baum N, Gollust S, Goold S, et al. Looking Ahead: Addressing Ethical Challenges in Public Health Practice. *Journal of Law, Medicine & Ethics* 2007;35(4):657-67.
69. Swain G, Burns K, Etkind P. Preparedness: Medical Ethics Versus Public Health Ethics. *Journal of Public Health Management Practice* 2008;14(4):354-57. .
70. World Health Organization (WHO). WHO Handbook for Guideline Development 2nd ed. Geneva: World Health Organization 2014.
71. Committee on Economic Social and Cultural Rights (CESCR). CESCR General Comment No. 14: The Right to the Highest Attainable Standard of Health (Art. 12) Adopted at the Twenty-second Session of the Committee on Economic, Social and Cultural Rights, on 11 August 2000 (Contained in Document E/C.12/2000/4) Office of the High Comissioner for Human Rights, 2000.

72. World Health Organization (WHO). Rio Political Declaration on Social Determinants of Health - Rio de Janeiro, Brazil, 21 October 2011 Rio de Janeiro: World Health Organization, 2011.
73. Bustreo F, Hunt P. Women's and Children's Health: Evidence of Impact of Human Rights. Geneva: World Health Organization 2013.
74. Comission on Social Determinants of Health (CSDH). Closing the gap in a generation: Health equity through action on the social determinants of health. Final Report of the Comission on Social Determinants of Health. Geneva, Switzerland: WHO 2008.
75. World Health Organization (WHO). The world health report 2000 - Health systems: improving performance. Geneva: World Health Organization 2000.
76. World Health Organization (WHO). The world health report 2008 - Primary health care: now more than ever. Geneva: World Health Organization 2008.
77. World Health Organization (WHO). Everybody's Business: Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action. Geneva: World Health Organization 2007.
78. United Nations (UN). United Nations General Assembly A/69/L.6. Draft outcome document of the United Nations summit for the adoption of the post-2015 development agenda. 2015
79. World Health Organization (WHO). A69/15. Health in the 2030 Agenda for Sustainable Development. 08. April. 2016. Geneva: World Health Organization, 2016.
80. World Health Organization (WHO). A69/11. Health in the 2030 Agenda for Sustainable Development. 28 May. 2016. Geneva: WHO 2016.
81. World Health Organization (WHO). EB138.R5. Strengthening essential public health functions in support of the achievement of universal health coverage. 30 January 2016. Geneva: World Health Organization, 2016.
82. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. 2017;358:j4008. doi: 10.1136/bmj.j4008 %J BMJ
83. Scott A, Hofmann B, Gutierrez-Ibarluzea I, et al. Q-SEA - a tool for quality assessment of ethics analyses conducted as part of health technology assessments. *GMS Health Technology Assessment* 2017;13:Doc02. doi: 10.3205/hta000128
84. World Health Organization (WHO). Communicating Risk in Public Health Emergencies. A WHO Guideline for Emergency Risk Communication (ERC) policy and practice. Geneva, Switzerland: World Health Organization 2018.
85. World Health Organization (WHO). Consolidated guideline on sexual and reproductive health and rights of women living with HIV. Geneva, Switzerland: World Health Organization 2017.
86. World Health Organization (WHO). WHO recommendations on antenatal care for a positive pregnancy experience. Geneva, Switzerland: World Health Organization 2014.
87. Mayring P. Qualitative content analysis: theoretical foundation, basic procedures and software solution. . Klagenfurt (Austria): Social Science Open Access Repository 2014.
88. Movsisyan A, Rehfuess E, Norris SL. When complexity matters: a step-by-step guide to incorporating a complexity perspective in guideline development for public health and health system interventions. *BMC medical research methodology* 2020;20(1):245. doi: 10.1186/s12874-020-01132-6
89. Maaløe N, Ørtved AMR, Sørensen JB, et al. The injustice of unfit clinical practice guidelines in low-resource realities. *The Lancet Global Health* doi: 10.1016/S2214-109X(21)00059-0
90. World Health Organization (WHO). Guidelines on Sanitation and Health. Geneva: WHO 2018.
91. World Health Organization (WHO). Evidence to recommendations: COVID-19 mitigation in the aviation sector, interim guidance 27 November 2020. Geneva: World Health Organization 2020.
92. World Health Organisation W. WHO Guideline on parenting to prevent child maltreatment and promote positive development in children aged 0-17 years Geneva: World Health Organization; 2020 [updated 01.03.2020; cited 2021 27.04.2021]. Available from: https://www.who.int/news/item/01-03-2020-guideline-on-parent-training-to-prevent-child-maltreatment-and-promote-the-positive-development-of-children-aged-0-17-years accessed 27.04.2021 2021.
93. World Health Organisation W. Chapter 18. Incorporating a complexity perspective into WHO guidelines. In: WHO, ed. WHO Handbook for Guideline Development 2nd edition. Geneva: WHO 2020.
94. Deutsche Gesellschaft für Epidemiologie D, Deutsche Gesellschaft für Public Health D, Deutsche Gesellschaft für Kinder- und Jugendmedizin D, et al. S3-Leitlinie - Maßnahmen zur Prävention und Kontrolle der SARS-CoV-2-Übertragung in Schulen | Lebende Leitlinie. AWMF-Registernummer 027-076. Version 1, Februar 2021 AWMF 2021.
95. Schweitzer J. Interview mit Prof. Dr. Rehfuess: "Ansteckungen unter Schülern kann die Leitlinie nicht ganz verhindern". 08. Februar 2021. *Zeit Online* 2021 08.02.2021.
96. Hilton Boon M, Thomson H, Shaw B, et al. Challenges in applying the GRADE approach in public health guidelines and systematic reviews: a concept article from the GRADE Public Health Group. *Journal of Clinical Epidemiology* 2021;135:42-53. doi: 10.1016/j.jclinepi.2021.01.001

8. Appenix A: Paper III: Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making

Stratil, J.M., Baltussen, R., Scheel, I. et al. (2020) Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making. Cost Eff Resour Alloc 18, 8. doi: 10.1186/s12962-020-0203-6 50

Stratil *et al. Cost Eff Resour Alloc* (2020) 18:8
https://doi.org/10.1186/s12962-020-0203-6

Cost Effectiveness and Resource Allocation

REVIEW Open Access

Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making

J. M. Stratil[1*], R. Baltussen[2], I. Scheel[3], A. Nacken[1] and E. A. Rehfuess[1]

Abstract

Background: Decision-making in public health and health policy is complex and requires careful deliberation of many and sometimes conflicting normative and technical criteria. Several approaches and tools, such as multi-criteria decision analysis, health technology assessments and evidence-to-decision (EtD) frameworks, have been proposed to guide decision-makers in selecting the criteria most relevant and appropriate for a transparent decision-making process. This study forms part of the development of the WHO-INTEGRATE EtD framework, a framework rooted in global health norms and values as reflected in key documents of the World Health Organization and the United Nations system. The objective of this study was to provide a comprehensive overview of criteria used in or proposed for real-world decision-making processes, including guideline development, health technology assessment, resource allocation and others.

Methods: We conducted an overview of systematic reviews through a combination of systematic literature searches and extensive reference searches. Systematic reviews reporting criteria used for real-world health decision-making by governmental or non-governmental organization on a supranational, national, or programme level were included and their quality assessed through a bespoke critical appraisal tool. The criteria reported in the reviews were extracted, de-duplicated and sorted into first-level (i.e. *criteria*), second-level (i.e. *sub-criteria*) and third-level (i.e. *decision aspects*) categories. First-level categories were developed a priori using a normative approach; second- and third-level categories were developed inductively.

Results: We included 36 systematic reviews providing criteria, of which one met all and another eleven met at least five of the items of our critical appraisal tool. The criteria were subsumed into 8 *criteria*, 45 *sub-criteria* and 200 *decision aspects*. The first-level of the category system comprised the following seven substantive criteria: "Health-related balance of benefits and harms"; "Human and individual rights"; "Acceptability considerations"; "Societal considerations"; "Considerations of equity, equality and fairness"; "Cost and financial considerations"; and "Feasibility and health system considerations". In addition, we identified an eight criterion "Evidence".

Conclusion: This overview of systematic reviews provides a comprehensive overview of criteria used or suggested for real-world health decision-making. It also discusses key challenges in the selection of the most appropriate criteria and in seeking to implement a fair decision-making process.

*Correspondence: stratil@ibe.med.uni-muenchen.de
[1] Institute for Medical Information Processing, Biometry and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany
Full list of author information is available at the end of the article

Keywords: Decision-making, Decisionmaking, Resource allocation, Priority-setting, HTA, Health technology assessment, Criteria, WHO, WHO-INTEGRATE

Background

Decision-making in public health and health policy is complex [1–3]. Processes that consider evidence and other considerations in a structured manner require a careful deliberation of many and often conflicting normative and technical criteria [4–10]. The choice of which of these criteria should be employed in the form of criteria has a profound impact on the outcome of the decision-making process. In many decision-making processes directly addressing health, criteria have mostly been concerned with effectiveness and cost [10–13]. This is at odds with the complexity of real-world decision making, were normative and feasibility considerations may act as key drivers of decisions (e.g. infringement of population health interventions on individual rights or interactions of interventions with other components of a health system) [14, 15]. The values and perceptions of different stakeholders with respect to normative and technical considerations often vary greatly both within and across societies. As there are various reasonable and defendable perceptions of which values and principles should guide the decision-making process and as there is no consensus on the *right* or *best* criteria, reasonable disagreement about the *right* decision or action is likely in pluralist societies [16].

Of course, many health-relevant decisions in public health policy and practice are made without adhering to structured decision-making processes based on predefined sets of criteria, populating those with evidence and weighting the results. But in various areas of public health policy and practice, such structured processes are relied on in the evaluation or comparison of alternative interventions or modes of actions. This inter alia includes the allocation of resources [9], the setting of research priorities [8, 17], decision-making about public health interventions [18], the assessment of health technologies for funding or reimbursement [19–21], or investment or disinvestment considerations [22]. Selecting the most appropriate and relevant criteria is a challenging but critical task in all of these structured decision-making processes.

The criteria used across different types of decisions have been addressed in multiple reviews. Several reviews have explored the criteria used when applying multi-criteria decision analysis (MCDA) [13, 23–28], an"umbrella term to describe a collection of formal approaches which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter" [29]. Other reviews have explored the criteria employed in the context of health technology assessments (HTA), which intend to examine social, economic, organizational and ethical considerations in relation to health technologies in a comprehensive manner [30]; these covered both the criteria to inform decisions about health technologies by national or sub-national HTA institutions [19, 31–33], and the criteria used for selecting the technologies or interventions a HTA is to be conducted on [8, 34]. In general, reviews have addressed criteria used for making decisions on funding or implementing health interventions or technologies [9, 10, 22, 35–38], prioritizing research topics [39–41] or coverage decisions [7, 12, 42, 43]. Reviews include criteria used on various levels of decision making (national, regional, or local), in different contexts (e.g. high- vs. low-income countries), and proposed by various stakeholder groups (e.g. decision makers, beneficiaries/patients).

Against this background, we conducted this study as part of a larger research project to develop a new evidence-to-decision (EtD) framework. The WHO-INTEGRATE EtD framework was developed to be firmly rooted in WHO norms and values and reflective of the changing global health landscape, and to encompass a comprehensive set of criteria suitable for decision making on clinical practice, public health, and health system interventions [15]. Within the development process of the framework, we conducted this overview of systematic reviews de-novo with the objective to provide a comprehensive overview of criteria used or intended to be used in real-world health decisions. More details on the role of this review in the development process of the WHO-INTEGRATE EtD framework is provided in our publication Rehfuess/Stratil et al. [15].

Methods

Our search strategy combined the terms "decision-making", "decision maker*", "decision analysis", "multi-criteria decision analysis", "priority setting", "resource allocation", "policy-making" and "policy-maker*" and their synonyms with the terms "criterion" and "criteria" as well as the terms "review*", "literature search", "mapping", "meta analysis" and their synonyms. Searches were conducted in PubMed and focused on the occurrence of these search terms in title and abstract. As the term "criteria" is used in many adjacent fields (e.g. referring to treatment or diagnostic decisions), we complemented these

systematic searches with hand searches of the references of all included studies.

Title and abstract screening was conducted using the software Rayyan [44]. Title and abstract screening as well as full-text screening was independently undertaken by two authors (JMS and AN), based on the inclusion and exclusion criteria shown in Table 1. We included studies which had conducted systematic searches of the literature and had comprehensively reported real-world criteria used in health decision-making. Studies focused on clinical decision-making (i.e. concerned with the decisions of individual patients) as well as studies focussing on the concepts or measurements of individual criteria (e.g. cost-effectiveness) were excluded. Where discrepancies could not be resolved by the two screening authors, a third author (ER) was consulted. Screening of the records identified through the updated literature searches was conducted by the author (JMS) and a research assistant (ST). The original literature searches were undertaken in September 2016 and updated in July 2018.

Information extracted from the included reviews were (i) study objective, (ii) type of health decision, (iii) the types of studies included, (iv) the strategy used to identify primary studies or documents, (v) information on how the criteria were compiled in the primary studies, (vi) the topic of the health decision in the primary studies (e.g. public health interventions, pharmaceuticals), (vii) the regional context of the primary studies (e.g. high- or low-income countries), (viii) the decision-making level (e.g. national, regional, local) and, importantly, (ix) the criteria themselves.

We critically appraised included studies. As no adequate, validated critical appraisal tool was available, we adapted items of the CASP systematic review checklist [45] and AMSTAR 2 [46] to our research question (Additional file 1). Our critical appraisal tool focuses on (i) the formulation of a clear research question regarding the decision-making process to be explored, (ii) a comprehensive search strategy, (iii) the adequate selection of eligible studies, (iv) the comprehensive extraction of criteria, (v) the critical appraisal of primary studies, (vi) the adequate description of the identified criteria (vi) the consideration of potential conflicts of interest, and (vii) the use of pre-established methods. The critical appraisal was conducted by one author (JS) and cross-checked by a research assistant (ST).

Given the intended primary use of the WHO-INTEGRATE framework in the development of WHO guidelines, the analysis focuses on substantive criteria (i.e. "What are the considerations or criteria a decision should be based on?"; e.g. cost, health benefit, available resources) rather than procedural criteria (i.e. "How should the process through which a decision is made be organized?"; e.g. transparency, participation of key stakeholders, opportunity for revising decisions).

This information was extracted onto an Excel spreadsheet by one author (JMS) and spot checked by a research assistant (ST). Wherever possible, criteria were extracted as stated in the primary studies. Where the reviews only reported synthesised criteria without a direct link to the primary studies, we extracted these synthesised criteria (e.g. "disease burden" and "burden of illness" as reported in primary studies summarized in a single "burden of disease" criterion in the included review). Categories, referring to the classification system developed or used in the reviews, were not extracted (e.g. "cost" and "cost-effectiveness" as reported in primary studies summarized under the criterion "financial considerations" in the included review). The criteria were then re-organised by one author (JMS) by combining (i) identical criteria (e.g. "burden of disease" and "burden of disease") and (ii) criteria described through similar terms with the same meaning (e.g. "burden of disease", "burden of illness" or "disease burden").

The criteria were then synthesised in a mixed inductive and deductive approach:

Table 1 Inclusion and exclusion criteria for the overview of systematic reviews

Inclusion criteria	Exclusion criteria
The publication is a review based on systematic searches of the literature	The publication is not a literature review
The publication is concerned with the criteria considered as part of a health decision between two or more options or the weighting of multiple options, made by a governmental or non-governmental organization on a supranational, national, or programme level	The publication did not utilize a systematic search strategy
	The publication is focused on selected criteria (e.g. cost-effectiveness) rather than sets of criteria guiding a decision
The health decision is made from a population perspective regarding the general population, patients, healthcare personnel, health decision-makers or other similar stakeholders in public health and healthcare	The publication relates to a health decision made outside of an organizational context (e.g. General Practitioner's office) and/or from an individual perspective (e.g. treatment choices for an individual patient in clinical practice)
The publication reports on a comprehensive set of criteria identified through the literature searches (at least 3 criteria)	The publication primarily addresses issues of how to measure, weight or calculate a criterion (e.g. cost-effectiveness, quality of life)
The publication is written in English, German, Spanish or Italian	

For the deductive approach, we used an intermediate step in the development of the WHO-INTEGRATE framework [15], the seven so-called preliminary criteria "Health-related balance of benefits and harms", "Human and individual rights", "Acceptability considerations", "Societal considerations", "Considerations of equity, equality and fairness", "Cost and financial considerations", "Feasibility and health system considerations", as well as "Evidence considerations" (Note that these categories were revised in the final WHO-INTEGRATE Framework [15]). "Evidence considerations" was singled out to align with the role of evidence as a meta-criterion in the WHO-INTEGRATE Framework: rather than taking evidence as one of several substantive decision-making criteria into account, the framework argues for reflecting on the quality of evidence of each criterion and considering these aspects alongside. We used these—what we refer to as—*criteria* as level one of the category system. During the synthesis, we remained open-minded about revisions of the category system to be able to capture new considerations relevant for decision making in an appropriate manner. For the inductive approach, we started from the criteria as reported in primary studies and reviews and grouped similar criteria into groups of—what we refer to as—*sub-criteria* (level two of the category system) and *decision aspects* (level three of the category system). Criteria relating to decision-making principles, procedural criteria and research priority setting were extracted and categorized separately.

In cases where the exact meaning of a criterion was unclear, the primary publication was consulted wherever possible. Were uncertainty remained, these cases were discussed with a research assistant (ST) or with other members of the research team (ER; RB). After an initial sorting of criteria identified through the included reviews into the three-level category system, this was discussed and refined through discussions between JMS, ST, RB and ER; one author (JMS) subsequently conducted a second round of sorting of the extracted criteria to ensure that all *criteria, sub-criteria* and *decision aspects* would be placed correctly within the category system.

Results

The literature search yielded 4448 unique records, of which 106 were assessed for eligibility based on their full text. A further 88 records were identified through hand searching (see Additional file 2 for PRISMA diagram).

We included 36 reviews in this overview of systematic reviews [4–8, 10–13, 19–23, 27, 28, 31–37, 39, 41–43, 47–54]. All of these were published after 2006, with 15 reviews published in 2018 or 2017 and only 5 reviews published before 2010. 16 reviews provided criteria used for or intended to guide various priority setting exercises [5, 6, 8, 11, 12, 31, 32, 34–36, 38, 41, 42, 49, 51, 55], with one review focused on research priority setting (in the field of child health and nutrition) [39]. Six reviews were framed in the context of multi-criteria decision analysis [6, 13, 23, 27, 28, 48]. Three reviews explored criteria used to guide investment or disinvestment decisions [22, 27, 37]. Two reviews assessed criteria to guide the selection of topics for HTA [8, 34], and eight reviews captured criteria used in HTA [8, 11, 19, 20, 31, 32, 48]. Four reviews focused specifically on the evaluation of or decisions on vaccines [4, 21, 47, 55]. 19 reviews exclusively included studies or documents from high-income countries while five had an explicit focus on decision-making processes in low- and middle-income countries [10, 35, 52, 55]. The number of criteria extracted from each publication ranged from 31 [4] to 360 [6].

The *criteria*, *sub-criteria* and *decision aspects* based on the -criteria extracted from the reviews are provided in Table 2. An additional category containing synthesized criteria extracted from the included reviews is provided in Additional file 3. The first level of the category system encompasses seven substantive *criteria*, i.e. "Health-related balance of benefits and harms", "Human and individual rights", "Acceptability considerations", "Societal considerations", "Considerations of equity, equality and fairness", "Cost and financial considerations", and "Feasibility and health system considerations". In addition to these substantive criteria, we also identified an eight criterion "evidence" (Table 3).

As criteria may be used in different decision-making processes and different decision-making contexts, not all *criteria* may apply. One important distinction, for example, is whether the problem to be addressed (e.g. a specific disease) has already been decided on or not. If so, the decision is about selecting one out of several options to address the problem, and considerations regarding the priority of the problem itself (e.g. burden or severity of disease or disability) are no longer relevant.

As noted in the methods section, we sorted criteria into a category system based on content. This way of organizing the criteria could be modified by adding additional dimensions. For example, one could also adopt a temporal perspective where criteria may relate to the point in time before an intervention is decided on or implemented the process of implementing the intervention or the short-term or longer-term outcome of the intervention. As an illustration, equity considerations can be framed as relating to the starting point (e.g. priority of a given health issue due to high health inequity), as an criterion of relevance to the implementation process (e.g. distribution of adverse events across all those affected by the intervention) or as an outcome (e.g. reduced health inequity several years

Table 2 Overview of substantive *criteria, sub-criteria,* and *decision aspects*

Criteria	Sub-criteria	Decision aspects
Health related balance of benefits and harms	General considerations surrounding benefit/effect	Benefits/effect/efficacy/effectiveness/impact [1, 2, 4–10, 12–15, 17–27, 29, 33–*36*]
		Health related Benefits/effect/efficacy/effectiveness/impact [1, 2, 4, 8, 11–15, 21, 23, 26, 31, 34, 35]
		Uptake of intervention [15, 20, 34, 35]
		Magnitude of benefit/effect/impact [2, 4, 11, 14, 18, 31, 33, *36*]
		Additional or indirect effects [2, 6, 33, 34]
	Type and composition of effect/benefit/impact [2]	Impact on mortality, survival, longevity and life expectancy [1, 2, 4, 11, 16, 19, 21, 24–26, 28, 34–*36*]
		Last chance therapies [23, 24]
		Impact on morbidity and disability [1, 2, 16, 35]
		Potential changes in health consequences [24, 25]
		Impact on (health-related) quality of life [2, 8, 11, 12, 14, 19, 20, 22, 25, 26, 28, 29, 31, 33, 35, *36*]
		Impact on patient-reported outcomes [2, 12, 16, 21, 26]
		Valuation of health outcomes by patients and desirability of the effects [2]
		Preventive benefits/effects or preventive approaches [1, 2, 4, 5, 16, 21, 25, 26, 31]
	Character of benefit or effect	Onset of effect and time until benefit [2, 11, 13]
		Duration, sustainability and lasting effect [2, 11, 13, 15, 27, 31]
	Individual and population level of benefit	Clinical benefits/effectiveness/impact [2, 3, 5, 11, 14, 17–22, 24, 25, 28, 31–33, *36*]
		Individual level benefit, effectiveness or impact [1, 2, 4, 18, 21, 22, 25, 26, *36*]
		Marginal benefits (for every patient) [20, 26]
		Population level benefit, effectiveness or impact [1–5, 9, 12, 16, 18, 19, 24–27, 31, 34, 35]
		Threshold effectiveness on populations (herd immunity) [15, 34]
	Balance of benefits and harms	Balance of (health) benefits and harms [5, 14, 18, 19, 23, 24, 34]
	General considerations surrounding harm/risk	General safety, risk and tolerability of intervention [2, 4, 5, 7–9, 11–17, 19–26, 29, 31, 33–*36*]
		Magnitude and likelihood of adverse events [26, 33–35]
		Valuation of health outcomes by patients and desirability of the effects regarding harms [34]
		Short and long term risk and safety profile [26]
		Over diagnosis and over treatment [16, 26]
		Stigmatization [2, 26]
		Risk of failure of intervention [15, 34, 35]
		Burden of treatment [2, 26, 33]
		Risk of inappropriate use [8, 16, 25, 26]
		Impact on disease patterns and reduced long-term effectiveness [15, 34]
		Other or additional adverse events [2]
	Health-related need and priority	Health-related needs: in general [1, 2, 11, 14, 22, 23, 24]
		Burden and impact of disease: in general [1, 2, 4, 7–10, 13–16, 21, 23–26, 33, 34, *36*]
		Magnitude of the problem [4, 10, 25, 34]
		Burden of disease measured through epidemiological indicators [1, 2, 4, 5, 8, 11, 13–15, 21, 23–26, 28, 33–*36*]

Table 2 (continued)

Criteria	Sub-criteria	Decision aspects
		Size of affected population and number of potential beneficiaries [1, 2, 4, 5, 8, 10, 11, 13–17, 20–26, 28, 31–34, *36*]
		Maximum potential for disease burden reduction [10]
		Severity of disease/condition: in general [1, 2, 4, 5, 9–12, 14, 16–19, 21–26, 28, 29, 31, 33–*36*]
		Severity of disease/condition: long term outcomes [34, 35]
		Severity of disease/condition: life threatening disease/condition and prognosis without treatment [1, 5, 11, 16, 21, 28, 29, 34]
		Severity of disease/condition: late stage or end-of life status of disease/condition [5, 21, 24, 25, 28, 29, *36*]
		Outbreaks and epidemic potential [15, 34]
		Urgency and emergencies [1, 2, 13, 25]
Human and individual rights		Human rights considerations [2, 11, 12, 20, *36*]
		Autonomy and informed consent [2, 11, 35]
		Privacy and confidentiality [26, 35]
		Intrusiveness of intervention [22]
Acceptability considerations	Perceived priority of the problem	Public perception of disease burden, disease risk or severity [15, 34, 35]
	Acceptability	Acceptability in general [2, 4, 6, 14, 15, 24, 26, 27, 34–*36*]
		Acceptability of cost and financial outcomes [2, 14, 25, 26, 34]
	Acceptability by beneficiaries	Acceptability by beneficiaries: in general [2, 6, 8, 11, 15, 16, 26, 35]
		Comfort, convenience and user experience [2, 11, 12, 14, 15, 20, 24–26, 31, 33, 35]
	Acceptability by those providing intervention	Acceptability by those providing intervention [15, 16, 34, 35]
	Social and cultural acceptability	Social and cultural acceptability [2, 4, 8, 9, 11, 15, 18, 21, 22, 25, 26, 31–34, *36*]
		Ethical/moral acceptability [2, 6, 11, 12, 15, 18, 22, 25, 26, 32–34, *36*]
	Stakeholder demand, interests and pressures	Advocacy and stakeholder (in general) interests and pressures [1, 2, 6, 14, 15, 21, 25, 33, *36*]
		Demands, interest and pressures of the public [1, 2, 8, 14, 24–26, *36*]
		Demands, interest and pressures by industry [2]
		Pressures, demand and interest of beneficiaries and patient representatives [2, 17, 20, 22, 25, 32, 33]
		Pressures, demand and interest of those providing intervention [2, 17, 20, 25, 26, 32, 33]
		Media attention and coverage [1, 32]
Societal considerations	Societal needs and priority	Social burden of disease (individual/population) [8, 18, 21, 29, 34, 35]
		Social needs [24]
		Economic burden of disease on society (in general) [32]
	Social and societal impact	Social impact or benefits [2, 3, 5, 6, 14, 21, 23–26, 31, 33, 35]
		Impact on non-health outcomes (in general) [15, 18, 26, 34]
		Impact on poverty [1, 2, 4, 14, 25, 26, *36*]
		Relevance to social development of the country [10]
		Value of hope [11]
		Raise profile of condition [14]
	Impact on economy	Impact on economy [2, 25, 26]
		Impact on productivity and population in productive age [2, 12, 16, 25]
		Relevance to economic development of the country
		Innovativeness (potential to encourage innovation) [2, 9, 10, 14, 21, 23–25, 31]
	Environmental impact	Environmental and/or ecological impact of intervention (in general) [14, 16, 21, 29–31, 33]
	Impact on future generations	Impact on future generations [19]

Table 2 (continued)

Criteria	Sub-criteria	Decision aspects
Considerations of equity, equality and fairness	Equity and equality	Equity/equality considerations: in general [1–4, 9–11, 14, 15, 18, 19, 22–24, 26, 27, 29, 33–*36*]
		Fairness considerations: in general [2, 3, 6, 11, 16, 26, *36*]
		Impact on (health) (in-)equity/(in-)equality [2, 6, 10, 14, 26, 31, 34, 35]
		Distribution of benefits and harms [10, 26]
	Accessibility	Accessibility in general [2–4, 6, 9, 14–17, 19, 21, 24, 29, 31, 33–35]
		Equity in accessibility [1, 2, 4, 6, 7, 11, 25, 26, 35]
		Physical and spacial accessibility [2, 4, 6, 16, 26]
		Timeliness of access (time spent waiting for treatment) [2, 3, 22, *36*]
		Informational accessibility [11]
		Financial accessibility of intervention/Affordability [4, 9, 10, 12, 14, 19, 25, 26, 34–*36*]
		Affordability: risk of catastrophic health expenditure [11, 16, 25]
		Affordability: cost and Financial impact on beneficiaries [11, 13, 14, 16, 22, 25, 26, 31, 34, 35]
	Availability	Availability/lack of suitable alternatives [1, 2, 5, 11, 12, 14, 17, 19–25, 28, 32, 34–*36*]
		Limitations of alternative interventions [2, 11, 14, 15, 25, 26, 34, 35]
		Rare diseases/orphan disease [12, 14, 16, 21, 24, 28]
		Unmet needs [11, 12, 16, 21, 24, 29]
	Responsibility	Ability to reduce own health risk and conditions arising from patient behavior [4, 5, 11, 16, 24, 25]
	Non-discrimination	Non-discrimination [11, 23, 24]
	Consideration regarding specific populations	Consideration of high-risk populations [2, 11, 14, 26]
		Consideration of vulnerable populations [1, 2, 4, 11, 21, 25, 26]
		Consideration of Socio-economic status [4, 16, 25, 26, *36*]
		Consideration of sex, gender, and/or sexual orientation [2, 6, 16, 25, 26, 34]
		Consideration of race/ethnicity [25, *36*]
		Consideration of care giver responsibilities [29]
		Consideration of age-groups [1, 2, 4, 11, 14, 16, 26, 28]
		Consideration of place of living [11, 16, 25]
		Consideration of identity and ideology [25]
		Other group related considerations [11]
Cost and financial considerations	Financial burden of disease	Financial burden of disease or current intervention on health system [2, 15, 16, 25, 32, 34, 35]
	Cost and budget impact of intervention	Cost/budget impact: in general [1, 2, 4–6, 8, 9, 11–19, 21–23, 25–31, 33–*36*]
		Cost per unit/usage [1, 2, 5, 11, 15, 17, 20, 23–26, 33, 34, *36*]
		Cost over time [20, 21, 25, 32, 34]
		Long term cost/budget impact [25]
		Overall cost/budget impact [1, 2, 11, 22, 26, *36*]
		Direct cost [11, 25, 32, 34, 35]
		Indirect/additional/hidden cost [11, 34–*36*]
		Marginal cost [11, 25]
		Opportunity cost [2, 11, 21, 31]

Table 2 (continued)

Criteria	Sub-criteria	Decision aspects
		Impact on other spending/investments [2, 6, 14, 25]
		Investment/start-up cost [2, 22, 30]
		Operating cost [2, 30]
		Lifecycle cost of intervention/technology [30]
		Economic/financial benefits and cost-minimization potential [1, 3, 8, 11, 13, 22, 25, 27, 28, 34–*36*]
		Cost to/budget impact on government or society [1, 2, 5, 18, 25, 26, 29, 33, *36*]
	Relation of cost and benefits	Relation of cost and benefits [1–7, 9–16, 18–26, 28, 31–*36*]
	Financial context	Appropriateness [2, 14]
	Financial feasibility	Affordability to health system Availability or lack of funds/funding [1–4, 6, 10–12, 14, 15, 21–25, 31, 34, *36*]
	Financial sustainability	Financial sustainability of intervention and consistency of funding [4, 11, 15, 31, 34]
Feasibility and health system considerations	Health-system related needs and priority	Burden of disease on health system [2, 25, 26, 31, 33, 34]
		Needs of the health-system
	Feasibility and capacity to implement	General feasibility considerations [2–4, 7, 9, 10, 14–16, 22–27, 31, 33, 34, *36*]
		Technical feasibility considerations [6, 26, 33]
		Practical feasibility considerations [4, 6, 26]
		Capacity to implement [1, 2, 6, 14, 21, 25, 29]
	Considerations of management and organization of health system	Availability of, capacity of and need for management and organizational structure [1, 2, 16, 22, 26, 34]
		Impact on management and organizational structure [2, 8, 11, 25]
		Logistical considerations [2, 15, 34]
		Availability of, capacity of and need for monitoring, surveillance and information system [34, 35]
	Resource considerations	Availability of, capacity of and need for resources (in general) [1, 2, 10, 14, 30, *36*]
		Impact on resources (in general) [2, 22–24]
		Efficiency of resource use [18, 24, 31]
	Considerations of human resources and their skills	Availability of, capacity of and need for human resources [2, 10, 13, 14, 18, 22, 26, 30]
		Availability of, capacity of and need for skill levels/knowledge of human resources [2, 3, 14, 16, 22, 25, 30, 35]
		Impact on human resources and skill levels [2, 14, 18, 25, 31]
	Considerations of Non-financial physical resources (equipment, infrastructure)	Capacity of, availability of need for of physical resources and infrastructure [16, 18, 22, 25, 26, 34–*36*]
		Impact on non-financial physical resources and infrastructure [18]
	Interaction with and impact on health system	Impact on performance of health system and impact on other services [1, 2, 4, 6, 8, 10, 11, 14, 16, 18, 20, 22, 26, 31, 33–35]
		Interaction and compatibility with other health system components [1–3, 6, 14, 22, 25, 33–35]
		Ease of use, application and burden of intervention [13, 15, 22, 26, 35]
	Appropriateness within health system	Appropriateness (in general) [2, 3, 11, 21, 25, 26, 31, *36*]
		Appropriateness of intervention for specific context [2, 16, 30]
	Legislative and regulatory considerations	Adherence to legal requirements, constrains and implications [2, 8, 10, 15, 16, 18, 25, 26, 31–35]
		Adherence to other directives, standards and requirements [2, 6, 14, 22, 30, 34]
	Political considerations	Political acceptability, interests and pressures [1–4, 6, 10, 11, 14, 18, 21, 25, 26, 33–*36*]
		Donor and global interests and pressures [2, 3]
		Political impact [2, 25, 33, 35]
		Alignment with priorities [1–4, 6, 13–15, 21, 23, 24, 26, 29, 30, 34, 35]
		Mission, mandate and goals of health system [2, 13, 14, 21, 25]

Table 2 (continued)

Criteria	Sub-criteria	Decision aspects
	Strategic considerations	Strategic planning and considerations [1, 2, 13, 19, 25, 31]
		Existing cooperation [2, 10, 14, 22]
		Decisions and practice of other institutions or stakeholders [2, 11, 17, 24, 29, 34]
		Alignment with recommendations, guidelines and standards [1–3, 7, 11, 14, 16, 21, 24, 26, 29, 34]
		Historical context and past decisions [1–3, 11, 13, 14, 17, 19–22, 25]
		Availability of incentives [1, 2]
		Impact on future decisions [6, 14]
		Keeping promises and commitments [2, 13]
	Characteristic of intervention	(technical) Complexity of intervention [2, 4, 6, 16, 21]
		Scalability of intervention [9, 27]
		Ability to evaluate intervention [2, 33–35]
		Reversibility of intervention [2]
		Flexibility of implementation [2]
		Uniqueness [21]
		Frequency of use and expected level of usage/activity [1, 2, 6, 8, 14, 25, 26, 32]
		Dependence on maintenance [26, 30]
		Sustainability of intervention utilization (e.g. Supply of parts, vaccines) [2, 14, 15, 27, 29, 34]
		Position of intervention in care pathway [5, 12, 20, 24]
		Additional uses of intervention [5, 21, 23–25]

1 Youngkong et al. 2009 [10], 2 Guindo et al. 2012 [6]; 3 Waithaka et al. 2018 [36]; 4 Wiseman et al. 2016 [35], 5 Fischer 2012 [43], 6 Wahlster et al. 2015 [13], 7 Ricciardi et al. 2015 [47], 8 Specchia et al. 2015 [34], 9 Hayati 2018 [42], 10 McGregor et al. 2014 [41], 11 MacLeod et al. 2016 [7], 12 Angelis et al. 2018 [48], 13 Barasa et al. 2015 [57], 14 Cromwell et al. 2015 [5], 15 Burchett et al. 2012 [4], 16 Varela-Lema et al. 2016 [38], 17 Vuorenkoski et al. 2008 [12], 18 Cowles et al. 2017 [50], 19 Golan et al. 2011 [33], 20 Erntoft et al. 2011 [51], 21 Friedmann et al. 2018 [28], 22 Ølholm et al. 2015 [37], 23 Stafinski et al. 2011a [19], 24 Stafinski et al. 2011b [20], 25 Mobinizadeh et al. 2016 [32], 26 Marsh et al. 2014 [27], 27 Rudan et al. 2017 [39], 28 Ghijben et al. 2018 [31], 29 Drake et al. 2017 [23], 30 Diaconu et al. 2017 [52], 31 Polisena et al. 2013 [22], 32 Noorani et al. 2007 [8], 33 Johnson et al. 2009 [53], 34 González-Lorenzo et al. 2015 [54], 35 Piso et al. 2009 [21], *36 Niessen et al. 2012* [11]

The publication by Niessen et al. (Ref. 36) is highlighted in italic, as it is the only study meeting all criteria of our critical appraisal tool

after introducing the intervention). Further additional dimensions could be a focus on individuals, populations or systems (e.g. clinical health benefits for the individual, reduction of the disease burden of a population, or impact on the performance of a health system following an intervention). In the organization of the criteria, we kept such additional organizational dimensions in mind.

The most frequently reported criteria were health-related impact of interventions, cost, cost-effectiveness and political interests or priorities; these were covered in all of the included reviews. Rarely used criteria were concerned with the environmental or societal impacts of interventions, and (non-financial) resource availability/needs. The granularity (level of detail with respect to *sub-criteria/aspects*) varied widely depending on the criterion: the criteria related to cost or financial considerations included general (e.g. "resource use" or "cost") as well as very specific usages (e.g. distinct ways to quantify cost-effectiveness). In contrast, criteria related to the societal or environmental impacts of interventions, as well as considerations regarding equity or equality were usually reported in very generic terms.

Evidence in general or evidence regarding specific criteria was mentioned in most included reviews, most often using generic terms, such as "evidence". In some cases, the criterion evidence referred to specific measures, primarily "evidence of effectiveness/efficacy" and sometimes "evidence on cost". In other reviews, this included criteria regarding the relevance of the available evidence for a given context (e.g. "relevance of evidence" or "generalizability of evidence") and criteria regarding the quality of evidence (e.g. "certainty of evidence", "credibility of evidence" or "validity of evidence") (see Table 3 and Additional file 4).

We also identified several considerations of specific relevance to research priority setting, covering considerations regarding the answerability of the research question, research ethics or avoidance of duplication of research. As those were not the primary focus of this

Table 3 Overview of evidence considerations

Criteria	Decision aspects
Evidence Considerations	evidence in general [1, 2, 11–13, 15, 17, 21, 23, 31, 33, *36*]
	strength of evidence in general [2, 11, 12, 21, 25, *36*]
	Certainty of evidence (in general) [2, 5, 6, 11, 16, 21–24, 26, 28, 33, *36*]
	Quality of evidence [2, 4–6, 11, 14–22, 25, 26, 28, 29, 33, 34, *36*]
	Completeness of evidence [2, 6, 12, 14, 25, 26]
	Validity of evidence [2, 6, 25, 26, 29, *36*]
	Credibility of evidence [29, 34, *36*]
	directness of evidence [2, 11, 28]
	Consistency of evidence [2, 6, 14, 25, 26, 29, 33]
	Precision of evidence effect [2]
	Relevance of evidence [2, 6, 11, 25–29]
	Applicability and generalizability of evidence [2, 11, 22]
	Type and quality of evidence sources [2, 11, 17, 21, 34, 35]
	Experience based evidence [26, 34, 35]
	Evidence requirements [2, 5, 25]

publication, they are not further discussed here but listed in Additional file 4.

Furthermore, the included publications reported several decision-making *principles* (i.e. guiding concepts from which different criteria derive). The distinction between decision-making principles and substantive criteria is not always clear cut. For example, *human rights* can be regarded as an underlying principle from which other criteria derive (as used in the human rights-based framework by Bustreo et al. [56]), as well as a specific criterion (assessing whether the intervention is in accordance with human rights). From the publications included in our overview of systematic reviews, the following criteria were extracted: Beneficence, non-maleficence, fairness, diversity, fair innings, proportional shortfall, concern for the worse off, justice, formal justice, social justice, distributive justice, principles of human rights, principle of human dignity; marginal utility principle, principle of need and solidarity, collectivism, cohesion, mutuality, rule of rescue and Rawls' difference principle (see Additional file 4).

Table 2 provides an overview of the *criteria, sub-criteria* and *decision aspects* for the seven substantive criteria. Criteria relating to evidence are reported in Table 3.

The results of the critical appraisal are provided in Additional file 5. Only one publication, Niessen et al. [11], met all eight items; 11 out of 36 publications met five or more items. Most publications did not conduct a critical appraisal of included studies, did not report independent extraction of criteria by two reviewers, and did not state explicitly, that the review had been undertaken based on a protocol or otherwise pre-established methods (Additional file 5).

Discussion

Summary of findings

Drawing on 36 included reviews, we identified a set of 200 unique *decision aspects*. These were sorted into 7 substantive *criteria* und 45 *sub-criteria* as well as a separate criterion on evidence. The substantive criteria cover health-related balance of benefits and harms; human and individual rights; acceptability considerations; societal considerations; considerations of equity, equality and fairness; cost and financial considerations; and feasibility and health system considerations. We found that some *criteria, sub-criteria* and decision *aspects* are well developed in the literature, such as those referring to the health implications of an interventions or to the costs of an intervention. In contrast, several others lacked a clear conceptualisation, notably those relating to societal implications or equity and equality considerations.

The wide range of *decision aspects* were used to refine the *criteria* and *sub-criteria* in the WHO-INTEGRATE EtD framework, as well as to inform the development of definitions and guiding questions provided as part of the framework.

In addition to their use in the WHO-INTEGRATE framework, we postulate that the list of *criteria, sub-criteria* and *decision aspects* can be helpful to decision-makers in their own right: To the best of our knowledge, this is the most comprehensive and up-to-date list of real-world criteria available for health decision-making. It could therefore provide a valuable tool for informing decision-makers wishing to select those criteria relevant for a given type of decision and decision-making context. This comprehensive list is likely to be most relevant to decisions in public health or healthcare. Due to the focus of the present study, the applicability for research priority setting or the evaluation of diagnostic or testing devices is likely to be more limited, as we may not have covered all relevant publications.

Most of the reviews included in our overview of systematic reviews did not meet all or even a majority of the items of our critical appraisal tool. This finding does, however, need to be interpreted in view of the following considerations. First, a validated critical appraisal tool appropriate for the topic does not exist—neither at the level of systematic reviews nor at the level of primary studies. Only three of the included reviews undertook some form of critical appraisal: Whaitaka et al. [36] and Burchett et al. [4] used an adapted CASP Qualitative Checklist and Niessen et al. [11] used custom quality-of-research assessment scales. Second, the low score of

some of the included reviews is likely due to poor reporting rather than poor conduct (e.g. regarding pre-established methods, or data extraction in duplicate).Third, the value and relevance of criteria for a given decision-making process does not necessarily depend on the quality of the review they were derived from. For example, even if the criterion "environmental impact" was merely mentioned in a single systematic review of low quality, this would not invalidate its relevance for a decision-making process focused on interventions with pronounced environmental (adverse) effect (e.g. large-scale usage of DDT in malaria prevention).

Contextualization of findings

With our overview of systematic reviews, we build on several previously published reviews, notably, the review by Guindo and colleagues, which represented the most extensive general overview of criteria until now [6]. Rather than focusing on specific decision-making processes (e.g. priority setting in low- and middle-income countries), we sought to cover the full range and heterogeneity of criteria and their use across various health fields.

We followed an approach focusing on descriptive ("what criteria are used?"), rather than prescriptive ("what criteria should be used?") approach. Several overviews of more prescriptive frameworks have been published in the field of public-health ethics in recent years [58–60]. A similar undertaking—providing decision-makers with a basis to select appropriate criteria—was conducted by Vermeulen and Krabbe, who provided an overview of the most widely recognized arguments and principles used in decision-making [18]. Their more prescriptive publication, which explores decision arguments and principles, and our more descriptive publication complement each other.

In contrast to some of the other reviews of criteria for decision making [6, 10], we abstained from quantifying how often criteria were cited for several reasons: First, the focus of this publication was to provide an overview of criteria that can be used for decision-making, rather than to provide an overview of which criteria are (widely) used in different decision making settings, as was the purpose in other publications [5, 6, 35]. Second, the quantification of how often or rarely a criterion is used does not necessarily imply its relevance for a given decision-making process: we believe that relevance should be informed by normative considerations. Third, there is a pronounced heterogeneity in the included studies: this begs the question, whether a criterion used in decision-making in a local hospital should count as much as the criteria used in the health technology assessment process of a national or supra-national organization. Finally, the quantification of the use of criteria is complicated: not only were many studies cited in several included reviews [61], but some of the reviews referred to other reviews as their data sources [6, 38].

Strengths and limitations

Our focus of the literature search on a single data base (PubMed) and the reliance on a selection of terms such as "criteria/criterion" might have missed relevant studies conducted on this issue. These decisions were made due to significant time and resource constraints relating to the development of the WHO-INTEGRATE framework over a relatively short period of time. We countered this potential limitation by thoroughly searching the references of all included studies, which yielded some additional publications. Furthermore, during the extraction of the criteria from included reviews we noted that we seemed to have reached saturation, as from the mid-way point, additional extracted studies yielded no or minimal additional criteria. Expanding the search to additional databases, especially those in the fields of political sciences and health economics, with a more inclusive search strategy may yield valuable additional insights from a broader range of disciplines.

A significant strength of our publication is that—to the best of our knowledge—it is the most extensive overview of criteria used in or proposed for health decision making. We included studies from several different health fields, conducted on various levels of decision-making and topics and in heterogeneous contexts around the world. We classified this comprehensive and diverse set of criteria according to a theory-based categorization system comprising three levels, i.e. *criteria, sub-criteria* and *decision aspects*. In doing so, as a team we critically reflected on extracted criteria and their underlying rationale, seeking to be as consistent as possible in how we sorted criteria reported in included reviews into higher-order categories.

Implications for policy and practice

The very large number of criteria and sub-criteria identified in this publication highlights the complexity of health decision-making It can serve as a resource when considering which criteria to include in sound multi-criteria approaches (i.e. adhering to principles of completeness, lack of redundancy, mutual independence, operationalizability and clustering) and how to use these.

The challenge of selecting the right criteria

At the centre of any decision-making process will be the challenge of who selects which criteria and how they should be weighted or ranked against each other. As various stakeholders with diverging but reasonable motives

are likely to disagree on which criteria are the *right* ones, the focus often shifts from selecting the *right* criteria to making decisions using a *good* or *fair* process [16, 62]. Numerous procedural conditions which characterize such a fair process have been proposed, including in the Accountability for Reasonableness framework [16], among others [63–66]. A fair and transparent process and especially an adequate representation and participation of all relevant stakeholder groups is essential for achieving legitimacy [62, 63].

One approach to overcoming reasonable disagreement about criteria for decision making is to reflect on the underlying normative principles and to make them explicit, e.g. by exploring the roots of a conflict which may lie in (potentially) conflicting normative arguments, e.g. if improving the life of a large number of people has to be weighed against the interest of those suffering from rare diseases with no alternative treatment. While we extracted such principles in our overview of reviews, others have focused explicitly on these [18, 58, 60] and several frameworks to guide the discussions and selection process have been proposed in the public health ethics literature [59, 60, 64, 65, 67–74].

Furthermore, underlying motives and drivers of stakeholders should be taken into account when reflecting on proposed criteria, as these can manifest themselves as trojan horses cloaked in ethical rhetoric [75]. Some calls for strengthening the consideration of criteria beyond evidence of effectiveness or incremental cost-effectiveness ratios are motivated by vested interest in a specific outcome [27, 76]. Such conflicts of interest should not necessarily lead to dismiss the arguments made, but it should lead to a critical reflection regarding the relevance and appropriateness of the proposed criteria for a given decision-making process and the power relations in the discourse [75, 77].

The challenge of resolving conflicts within and between criteria

The criteria by themselves are often highly interconnected and at times conflicting. An example is the criterion "age", which can serve as a "surrogate" criterion for other normative and (harder to measure) considerations. For example, a focus on interventions targeting younger people may be motivated by their potential for achieving a longer life span (greater health impact) or to reduce productivity losses (positive impact on the economy). At the same time, "age" can be considered with respect to non-discrimination or equity: explicitly reflecting on age in order not to prioritize one group of people based on age as a characteristic (ageism). An explicit reflection on and discussion of such conflicts within criteria is important.

Furthermore, the criteria identified in the included reviews are partly overlapping (e.g. cost, effectiveness, and cost-effectiveness). Depending on the decision-making process and the tools used (e.g. MCDA), accounting for overlaps and redundancies may be of relevance. This can, for example, be achieved through selecting non-overlapping criteria or through increasing the granularity of the criteria. In particular the MCDA-literature has developed methods and guidance on how to identify and handle overlapping criteria [78].

The example of "age" as a criterion that can have conflicting interpretations highlights the need to set up a mechanism for handling conflicts within criteria and balancing interests in place. The same holds true for conflicts between criteria (e.g. positive impact on population health, negative impact on the natural environment), which occur on a regular basis in decision-making processes.

The challenge of using criteria

Populating criteria with evidence presents a third important challenge [15]. Evidence collection and synthesis approaches are well developed for some criteria (e.g. health impacts) although some challenges remain. For a few criteria, approaches are virtually non-existent in the literature on health decision-making (e.g. environmental implications) while for others there is a lack of clarity regarding the best methods to be employed (e.g. societal or environmental impact assessments) [15]. It is highly likely that suitable methods exist outside of the health decision-making or broader healthcare and public health literature and learning from other disciplines may offer solutions to this challenge.

Guideline development, HTA and other decision-making processes aiming to integrate evidence and criteria for decision making in a structured manner usually operate under significant time and resource constraints. To avoid treating criteria beyond effectiveness and cost-effectiveness as an "after thought", evidence will need to be collected or analysis on these other criteria. This will require the development of rapid and pragmatic approaches to keep such decision-making processes feasible.

Conclusion

The comprehensive list of criteria from and for real-world health decision-making presented here was an essential building block in the development of the WHO-INTEGRATE framework. We postulate that it can also be a useful stand-alone tool to inform health decision-making processes not employing an EtD framework. To make the best possible use of this list, solutions to the challenges of

selecting criteria, of resolving conflicts between criteria or their interpretation, and of identifying and appraising evidence towards these criteria will need to be found. The WHO-INTEGRATE framework seeks to address some of these challenges, by providing a set of criteria selected based on a strong normative basis and by offering a methodological toolbox, which suggests both comprehensive and pragmatic approaches to populating criteria with evidence [15].

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12962-020-0203-6.

Additional file 1. Critical appraisal tool.

Additional file 2. PRISMA flow diagram.

Additional file 3. Overview of criteria on abstraction levels 1–3 and levels 1–4.

Additional file 4. Overview of evidence considerations and principles.

Additional file 5. Results of the critical appraisal of included studies.

Abbreviations

MCDA: Multi criteria decision analysis; EtD: Evidence to decision; HTA: Health technology assessment.

Acknowledgements

We would like to acknowledge the contribution of the research assistants Anja Nacken (AN) and Sarah Thumbeck (ST) who supported us in the data analysis. We furthermore want to acknowledge of all those involved in the development of the WHO-INTEGRATE EtD Framework V1.0, in particular Anayda Portela and Susan L Norris, from the World Health Organization, who supported us on the conceptualization of our approach.

Authors' contributions

EAR conceived the overall research project on the WHO-INTEGRATE EtD Framework with substantial input from JMS. RB and IBS provided important contributions to the overall direction of the project and to the draft manuscript. JMS with support from AN conducted the literature searches as well as screening of the identified publications. JMS conducted the data extraction and analysis with support from the research assistant ST. EAR, RB and IBS supported the development of the category system and organization of the criteria. EAR, RB and IBS discussed different versions of the table of criteria and the content. JMS wrote the manuscript with input from EAR all authors critically reviewed different versions of the manuscript, suggested revisions and approved the version to be published. All authors read and approved the final manuscript.

Funding

Funding was provided by the World Health Organization's Department of Maternal, Newborn, Child and Adolescent Health through a grant received from the United States Agency for International Development (USAID). We also gratefully acknowledge that JMS's position was funded by the Bavarian Health and Food Safety Authority and that IBS's input was funded by the Norwegian Agency for Development Cooperation (Norad).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its Additional files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

This research was conducted within a research project initiated by the World Health Organization in which the WHO-INTEGRATE EtD Framework V1.0 was first developed and published by the authors. The authors declare that they have no competing interests. The funding organizations, in particular the USAID, the Bavarian Health and Food Safety Authority and Norad had no influence on the research process or content of this manuscript.

Author details

[1] Institute for Medical Information Processing, Biometry and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany. [2] Department for Health Evidence, Radboud University Medical Center, P.O.Box 9101, 6500 HB Nijmegen, The Netherlands. [3] Department of Global Health, Norwegian Institute of Public Health, PO Box 4404, Nydalen, 0403 Oslo, Norway.

Received: 9 January 2019 Accepted: 24 January 2020
Published online: 11 February 2020

References

1. Sanderson I. Intelligent policy making for a complex world: pragmatism, evidence and learning. Polit Stud. 2009;57(4):699–719.
2. Rajan D, Adam T, El Husseiny D, Porignon D, Ghaffar A, Schmets G. Briefing Note—Policy dialogue: what it is and how it can contribute to evidence-informed decision-making. Geneva: World Health Organization; 2015.
3. Holm S. The second phase of priority setting. Goodbye to the simple solutions: the second phase of priority setting in health care. BMJ. 1998;317:1000–2.
4. Burchett H, Mounier-Jack S, Griffiths U, Mills A. National decision-making on adopting new vaccines: a systematic review. Health Policy Plan. 2012;27(SUPPL.2):ii62–76.
5. Cromwell I, Peacock SJ, Mitton C. 'Real-world' health care priority setting using explicit decision criteria: a systematic review of the literature. BMC Health Serv Res. 2015;15:164. https://doi.org/10.1186/s12913-015-0814-3.
6. Guindo LA, Wagner M, Baltussen R, Rindress D, van Til J, Kind P, Goetghebeur MM. From efficacy to equity: literature review of decision criteria for resource allocation and healthcare decisionmaking. Cost Effect Resour Alloc. 2012;10:9. https://doi.org/10.1186/1478-7547-10-9.
7. MacLeod TE, Harris AH, Mahal A. Stated and revealed preferences for funding new high-cost cancer drugs: a critical review of the evidence from patients, the public and payers. Patient. 2016;9:201–22.
8. Noorani HZ, Husereau DR, Boudreau R, Skidmore B. Priority setting for health technology assessments: a systematic review of current practical approaches. Int J Technol Assess Health Care. 2007;23(3):310–5.
9. Stafinski T, Menon D, Marshall D, Caulfield T. Societal values in the allocation of healthcare resources: is it all about the health gain? Patient. 2011;4:207–25.
10. Youngkong S, Kapiriri L, Baltussen R. Setting priorities for health interventions in developing countries: a review of empirical studies. Tropical Med Int Health. 2009;14:930–9.
11. Niessen LW, Bridges J, Lau BD, Wilson RF, Sharma R, Walker DG, Frick KD, Bass EB. Assessing the impact of economic evidence on policymakers in health care—a systematic review. Methods research report (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract No. 290-2007-10061-I). AHRQ Publication No. 12(13)-EHC133-EF. Rockville: Agency for Healthcare Research and Quality; 2012.
12. Vuorenkoski L, Toiviainen H, Hemminki E. Decision-making in priority setting for medicines-A review of empirical studies. Health Policy. 2008;86(1):1–9.
13. Wahlster P, Goetghebeur M, Kriza C, Niederländer C, Kolominsky-Rabas P, on behalf of the National Leading-Edge Cluster Medical Technologies 'Medical Valley EMN'. Balancing costs and benefits at different stages of medical innovation: a systematic review of multi-criteria decision analysis (MCDA). BMC Health Serv Res. 2015;15:262. https://doi.org/10.1186/s12913-015-0930-0.

14. Dukhanin V, Searle A, Zwerling A, Dowdy DW, Taylor HA, Merritt MW. Integrating social justice concerns into economic evaluation for healthcare and public health: a systematic review. Soc Sci Med. 1982;2018(198):27–35.
15. Rehfuess EA, Stratil JM, Scheel IB, Portela A, Norris SL, Baltussen R. The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. BMJ Glob Health. 2019;4:e000844. https://doi.org/10.1136/bmjgh-2018-000844.
16. Daniels N. Accountability for reasonableness. BMJ. 2000;7272(321):1300–1.
17. Husereau D, Boucher M, Noorani H. Priority setting for health technology assessment at CADTH. Int J Technol Assess Health Care. 2010;26(3):341–7.
18. Vermeulen KM, Krabbe PFM. Value judgment of health interventions from different perspectives: arguments and criteria. Cost Effect Resour Alloc. 2018;16:16.
19. Stafinski T, Menon D, Philippon DJ, McCabe C. Health technology funding decision-making processes around the world: the same, yet different. Pharmacoeconomics. 2011;29(6):475–95.
20. Stafinski T, Menon D, Davis C, McCabe C. Role of centralized review processes for making reimbursement decisions on new health technologies in Europe. Clinicoecon Outcomes Res. 2011;3:117–86.
21. Piso B, Wild C. Decision support in vaccination policies. Vaccine. 2009;27(43):5923–8.
22. Polisena J, Clifford T, Mitton C, Elshaug AG, Russell E, Skidmore B. Case studies that illustrate disinvestment and resource allocation decision-making processes in health care: a systematic review. Int J Technol Assess Health Care. 2013;29(2):174–84.
23. Drake JI, de Hart JCT, Monleón C, Toro W, Valentim J. Utilization of multiple-criteria decision analysis (MCDA) to support healthcare decision-making FIFARMA, 2016. J Mark Access Health Policy. 2017;5(1):1360545.
24. Marsh K, Caro JJ, Hamed A, Zaiser E. Amplifying each patient's voice: a systematic review of multi-criteria decision analyses involving patients. Appl Health Econ Health Policy. 2017;15(2):155–62.
25. Mühlbacher AC, Kaczynski A. Making good decisions in healthcare with multi-criteria decision analysis: the use, current research and future development of MCDA. Appl Health Econ Health Policy. 2016;14(1):29–40.
26. Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo ZN, Lönngren T, Mussen F, Peacock S, Watkins J, et al. Multiple criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA emerging good practices task force. Value Health. 2016;19(1):1–13.
27. Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. PharmacoEconomics. 2014;32:345–65.
28. Friedmann C, Levy P, Hensel P, Hiligsmann M. Using multi-criteria decision analysis to appraise orphan drugs: a systematic review. Expert Rev Pharmacoecon Outcomes Res. 2018;18:135–46.
29. Benton V, Stewart T. Multiple criteria decision analysis—an integrated approach. New York: Springer; 2002.
30. del Llano-Señarís J, Campillo-Artero C. Health technology assessment and health policy today: a multifaceted view of their unstable crossroads. Berlin: Springer; 2015.
31. Ghijben P, Gu Y, Lancsar E, Zavarsek S. Revealed and stated preferences of decision makers for priority setting in health technology assessment: a systematic review. Pharmacoeconomics. 2018;36:323–40.
32. Mobinizadeh M, Raeissi P, Nasiripour AA, Olyaeemanesh A, Tabibi SJ. The health systems' priority setting criteria for selecting health technologies: a systematic review of the current evidence. Med J Islamic Rep Iran. 2016;30:329.
33. Golan O, Hansen P, Kaplan G, Tal O. Health technology prioritization: which criteria for prioritizing new technologies and what are their relative weights? Health Policy. 2011;102:126–35.
34. Specchia ML, Favale M, Di Nardo F, Rotundo G, Favaretti C, Ricciardi W, de Waure C. How to choose health technologies to be assessed by HTA? A review of criteria for priority setting. Epidemiol Prev. 2015;39(4 Suppl 1):39–44.
35. Wiseman V, Mitton C, Doyle-Waters MM, Drake T, Conteh L, Newall AT, Onwujekwe O, Jan S. Using economic evidence to set healthcare priorities in low-income and lower-middle-income countries: a systematic review of methodological frameworks. Health Econ. 2016;25:140–61.
36. Waithaka D, Tsofa B, Barasa E. Evaluating healthcare priority setting at the meso level: a thematic review of empirical literature. Wellcome Open Res. 2018;3:2.
37. Ølholm AM, Kidholm K, Birk-Olsen M, Christensen JB. Hospital managers' need for information on health technology investments. Int J Technol Assess Health Care. 2015;31(6):414–25.
38. Varela-Lema L, Atienza-Merino G, López-García M. Priority setting of health interventions. Review of criteria, approaches and role of assessment agencies. Gac Sanit. 2017;31(4):349–57. https://doi.org/10.1016/j.gaceta.2016.09.015.
39. Rudan I, Yoshida S, Chan KY, Sridhar D, Wazny K, Nair H, Sheikh A, Tomlinson M, Lawn JE, Bhutta ZA, et al. Setting health research priorities using the CHNRI method: VII. A review of the first 50 applications of the CHNRI method. J Glob Health. 2017;7(1):011004.
40. Tomlinson M, Rudan I, Chopra M, Hoosain N. A review of selected research priority setting processes at national level in low and middle income countries: towards fair and legitimate priority setting. Health Res Policy Syst. 2011;9:19.
41. McGregor S, Kaldor JM, Henderson KJ. How are health research priorities set in low and middle income countries? A systematic review of published reports. PLoS ONE. 2014;9(9):e108787.
42. Hayati R, Bastani P, Kabir MJ, Kavosi Z, Sobhani G. Scoping literature review on the basic health benefit package and its determinant criteria. Global Health. 2018;14:26. https://doi.org/10.1186/s12992-018-0345-x.
43. Fischer KE. A systematic review of coverage decision-making on health technologies-evidence from the real world. Health Policy (Amsterdam, Netherlands). 2012;107(2):218–30.
44. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5:210.
45. Critical Appraisal Skills Programme. CASP Systematic Review Checklist [online]. 2019. https://casp-uk.net/wp-content/uploads/2018/01/CASP-Systematic-Review-Checklist_2018.pdf. Accessed 10 Dec 2019.
46. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.
47. Ricciardi GW, Toumi M, Weil-Olivier C, Ruitenberg EJ, Dankó D, Duru G, Picazo J, Zöllner Y, Poland G, Drummond M. Comparison of NITAG policies and working processes in selected developed countries. Vaccine. 2015;33(1):3–11.
48. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. Eur J Health Econ. 2018;19(1):123–52.
49. Barasa EW, Cleary S, Molyneux S, English M. Setting healthcare priorities: a description and evaluation of the budgeting and planning process in county hospitals in Kenya. Health Policy Plan. 2017;32(3):329–37.
50. Cowles E, Marsden G, Cole A, Devlin N. A review of NICE methods and processes across health technology assessment programmes: why the differences and what is the impact? Appl Health Econ Health Policy. 2017;15(4):469–77.
51. Erntoft S. Pharmaceutical priority setting and the use of health economic evaluations: a systematic literature review. Value Health. 2011;14(4):587–99.
52. Diaconu K, Chen Y-F, Cummins C, Jimenez Moyao G, Manaseki-Holland S, Lilford R. Methods for medical device and equipment procurement and prioritization within low- and middle-income countries: findings of a systematic literature review. Global Health. 2017;13(1):59.
53. Johnson AP, Sikich NJ, Evans G, Evans W, Giacomini M, Glendining M, Krahn M, Levin L, Oh P, Perera C. Health technology assessment: a comprehensive framework for evidence-based recommendations in Ontario. Int J Technol Assess Health Care. 2009;25(2):141–50.
54. González-Lorenzo M, Piatti A, Coppola L, Gramegna M, Demicheli V, Melegaro A, Tirani M, Parmelli E, Auxilia F, Moja L, et al. Conceptual frameworks and key dimensions to support coverage decisions for vaccines. Vaccine. 2015;33:1206–17.
55. Boujaoude M-A, Mirelman AJ, Dalziel K, Carvalho N. Accounting for equity considerations in cost-effectiveness analysis: a systematic review of rotavirus vaccine in low- and middle-income countries. Cost Effect Resour Alloc. 2018;16(1):18.

56. World Health Organization. Women's and children's health: evidence of impact of human rights. Geneva: WHO; 2013.
57. Barasa EW, Molyneux S, English M, Cleary S. Setting healthcare priorities in hospitals: a review of empirical studies. Health Policy Plann. 2015;30:386–96.
58. Abbasi M, Majdzadeh R, Zali A, Karimi A, Akrami F. The evolution of public health ethics frameworks: systematic review of moral values and norms in public health policy. Med Health Care Philos. 2018;21(3):387–402.
59. Lee LM. Public health ethics theory: review and path to convergence. J Law Med Ethics. 2012;40(1):85–98.
60. ten Have M, de Beaufort ID, Mackenbach JP, van der Heide A. An overview of ethical frameworks in public health: can they be supportive in the evaluation of programs to prevent overweight? BMC Public Health. 2010;10:638.
61. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Effect Resour Alloc. 2006;4(1):14.
62. Martin D, Singer P. A strategy to improve priority setting in health care institutions. Health Care Anal. 2003;11(1):59–68.
63. Friedman A. Beyond accountability for reasonableness. Bioethics. 2008;22(2):101–12.
64. Childress JF, Faden RR, Gaare RD, Gostin LO, Kahn J, Bonnie RJ, Kass NE, Mastroianni AC, Moreno JD, Nieburg P. Public health ethics: mapping the terrain. J Law Med Ethics. 2002;30(2):170–8.
65. Tannahill A. Beyond evidence–to ethics: a decision-making framework for health promotion, public health and health improvement. Health Promot Int. 2008;23(4):380–90.
66. Sen A. The idea of justice. Cambridge: Massachusetts Belknap Press of Harvard University Press; 2009.
67. Kass NE. An ethics framework for public health. Am J Public Health. 2001;91(11):1776–82.
68. Krebs K, Urwin J. Public health: ethical issues. London: Nuffield Council on Bioethics; 2009.
69. Marckmann G, Schmidt H, Sofaer N, Strech D. Putting Public health ethics into practice: a systematic framework. Front Public Health. 2015;3:23.
70. Petrini C. Theoretical models and operational frameworks in public health ethics. Int J Environ Res Public Health. 2010;7(1):189–202.
71. Thompson A, Faith K, Gibson J, Upshur R. Pandemic influenza preparedness: an ethical framework to guide decision-making. BMC Med Ethics. 2006;7:12.
72. Upshur R. Principles for the justification of public health intervention. Can J Public Health. 2002;93(2):101–3.
73. Baum N, Gollust S, Goold S, Jacobson P. Looking ahead: addressing ethical challenges in public health practice. J Law Med Ethics. 2007;35(4):657–67.
74. Swain G, Burns K, Etkind P. Preparedness: medical ethics versus public health ethics. J Public Health Manag Pract. 2008;14(4):354–7.
75. Chanska W, Grunt-Mejer K. The unethical use of ethical rhetoric: the case of flibanserin and pharmacologisation of female sexual desire. J Med Ethics. 2016;42:701–4.
76. Marsh K, IJzerman M, Thokala P, Baltussen R, Boysen M, Kalo ZN, Lönngren T, Mussen F, Peacock S, Watkins J, et al. Multiple criteria decision analysis for health care decision making–emerging good practices: report 2 of the ISPOR MCDA emerging good practices task force. Value Health. 2016;19(2):125–37.
77. Meixel A, Yanchar E, Fugh-Berman A. Hypoactive sexual desire disorder: inventing a disease to sell low libido. J Med Ethics. 2015;41(10):859–62.
78. Marsh KD, Sculpher M, Caro JJ, Tervonen T. The use of MCDA in HTA: great potential, but more effort needed. Value Health. 2018;21(4):394–7.

Publisher's Note

9. Acknowledgements

First and foremost, I would like to offer my sincerest gratitude to my supervisors. I would like to express my deepest thanks to Eva Rehfuess for all the support in and beyond this PhD project; including all the insights into to world of scientific researched I had the privilege to gain along the way. Many thanks to Manfred Wildner, not only for his input to the project, but also to remind me of the real-world impact public health research in the end aims to achieve. Lastly, I would express my thanks to Rob Baltussen for all the support in the project, in particular during the meeting in Nijmegen and Amsterdam early in the process.

I would like to acknowledge all individuals who contributed to the project and without which the WHO-INTEGRATE EtD framework could not have been developed. I would like to express my gratitude towards Inger Scheel for her contributions to the project, but as well for her insights on life beyond work. I would like to thank Anayda Portela and Susan L Norris, who's contribution in the development of the WHO-INTEGRATE framework were crucial. I would like to acknowledge and thank the many individuals who gave freely of their time to review the WHO-INTEGRATE framework or components of it in its final or in preliminary versions and provided constructive feedback: Kristin Bakke-Lysdahl, Saskia den Boon, Randy Elder, Nathan Ford, Elizabeth Fox, Davina Ghersi, Claire Glenton, David Gough, Sean Grant, Björn Hofmann, Thomas Harder, Mike Kelly, Etienne Langlois, Stephanie Levy, Fergus Macbeth, Paul Montgomery, Graham Moore, Ani Movsisyan, Jane Noyes, Pierre Ongolo-Zogo, Wija Oortwijn, Deepak Paudel, Jennifer Petcovic, Mark Petticrew, Kevin Pottie, Pietro Refolo, Dario Sacchini, Gerard Schmets, Elham Shakibazadeh, James Thomas, Rebekah Thomas-Bosco, Andrea Tricco, Vivian Welch, Verina Wild and Teun Zuiderent-Jerak.

I would like to express my gratitude to my co-authors Deepak Paudel, Karen E. Setty, Carlos E. Menezes de Rezende, Aline A. Monroe, and Jimmy Osuret, without which the focus group discussions would not have been possible, as well as to Anja Nacken, Sarah Thumbeck, and Katja Krug, who provided valuable contribution in the data analysis. I would like to thank all focus group participants and all participants in the key informant interviews for their time and engagement. As well as Susan L. Norris, Anayda Portela, and Rebekah Thomas Bosco as well as other WHO staff for helping with the identification of suitable guidelines and key informants.

I would like to acknowledge all institutions that were involved in this thesis and in particular the IBE in Munich, the LGL in Munich, and the WHO Headquarters in Geneva. In particular, I would like to thank Prof. Dr. Ulrich Mansmann, Monika Darchinger, Dr. Magda Beer-Radermacher, and Dr. Annette Hartmann for their support.

I am deeply grateful for having had the opportunity to work in a great team of inspiring individuals that helped across different stages of this PhD project and beyond. In particular: Jacob Burns, Karin Geffert, Ani Movsisyan, Lisa Pfadenhauer, Stephanie Polus, Mike Rüb, Peter von Philipsborn, and Stefan Voss.

I would like to thank my parents, my brother, my grandparents, and my partner. Without them, all of this would not have been possible.