What Evidence to Trust?

NIPH
Norwegian Institute of Public Health

Social skills: Using GRADE and GRADE-CERQual in systematic reviews of social interventions

Heather Munthe-Kaas
Centre for Epidemic Interventions Research, Norwegian Institute of Public Health
SIA Conference, Stockholm, 24 May 2023

Declaration of interests

This presentation was created in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. I am co-founder of GRADE-CERQual and TRANSFER.

Funding for the development of GRADE-CERQual came from the following organizations:

Cochrane Methods

FONDATION BROCHER

Alliance for Health Policy and Systems Research

World Health Organization

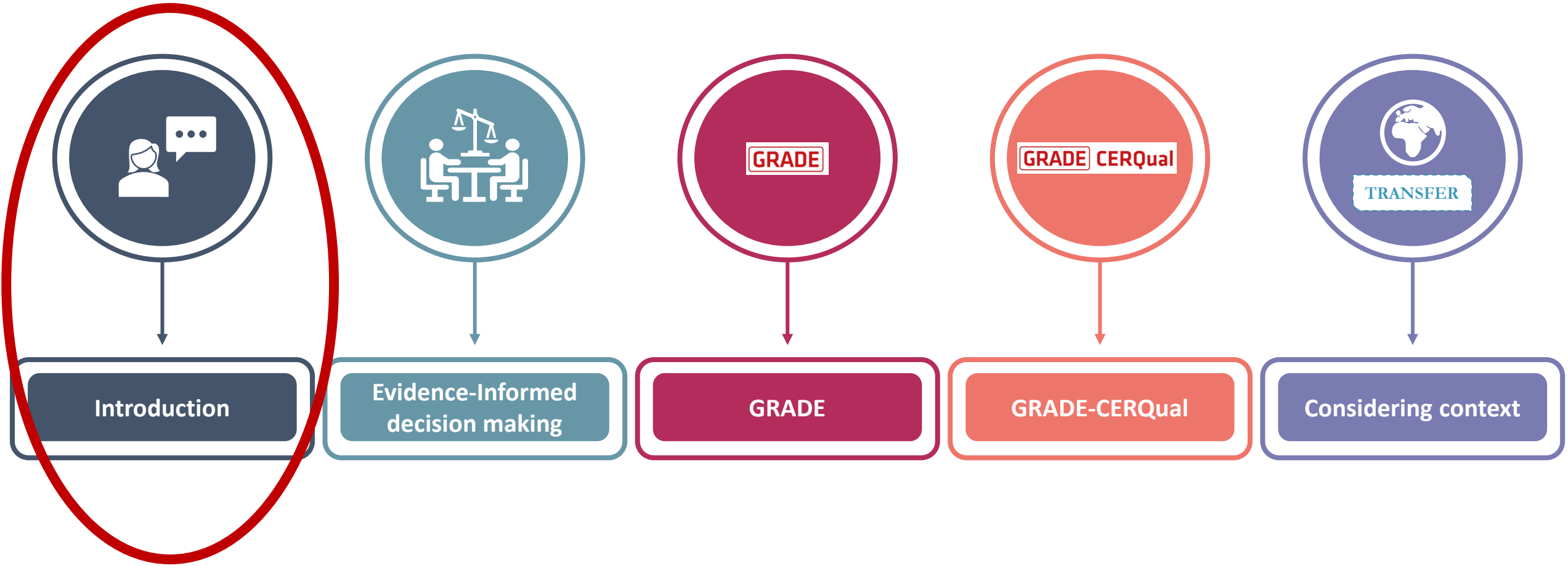
FHI Folkehelseinstituttet

Funding for the development of TRANSFER came from the following organizations:

FHI Folkehelseinstituttet

Campbell Collaboration

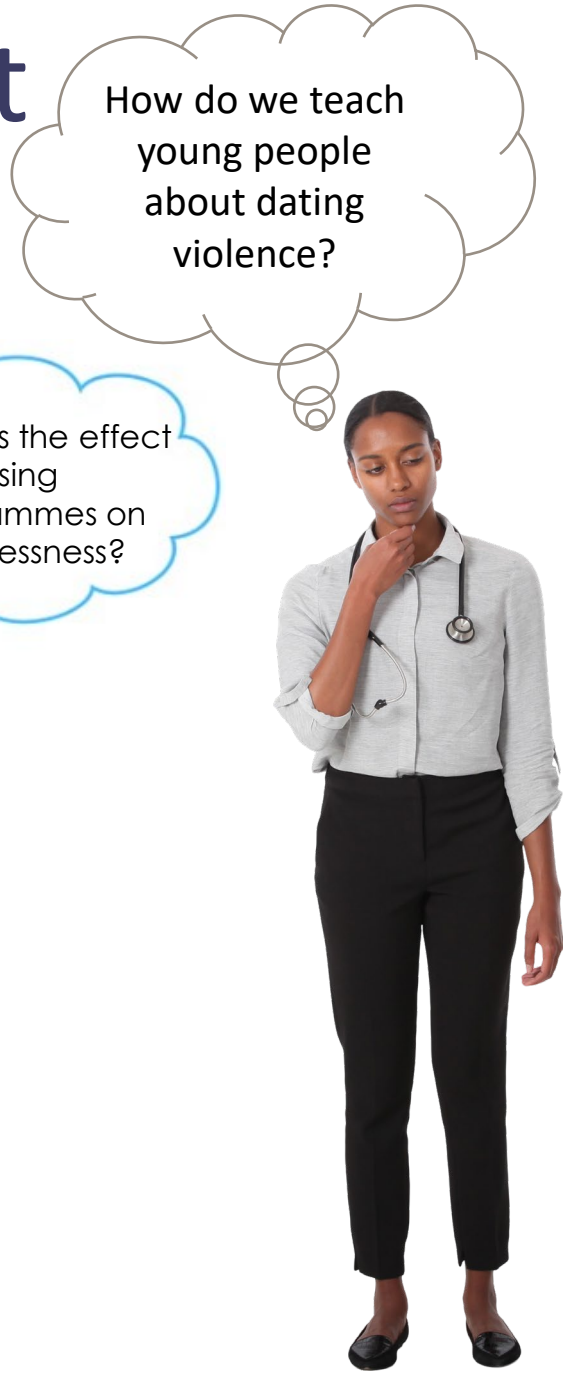
Overview


GRADE
GRADE CERQual
TRANSFER
Introduction
Evidence-Informed decision making
GRADE
GRADE-CERQual
Considering context

Systematic reviews of social

Social interventions have as much potential for good or harm as health interventions – or more.

Decision makers need evidence about interventions…



Systematic reviews help us to answer important questions about interventions

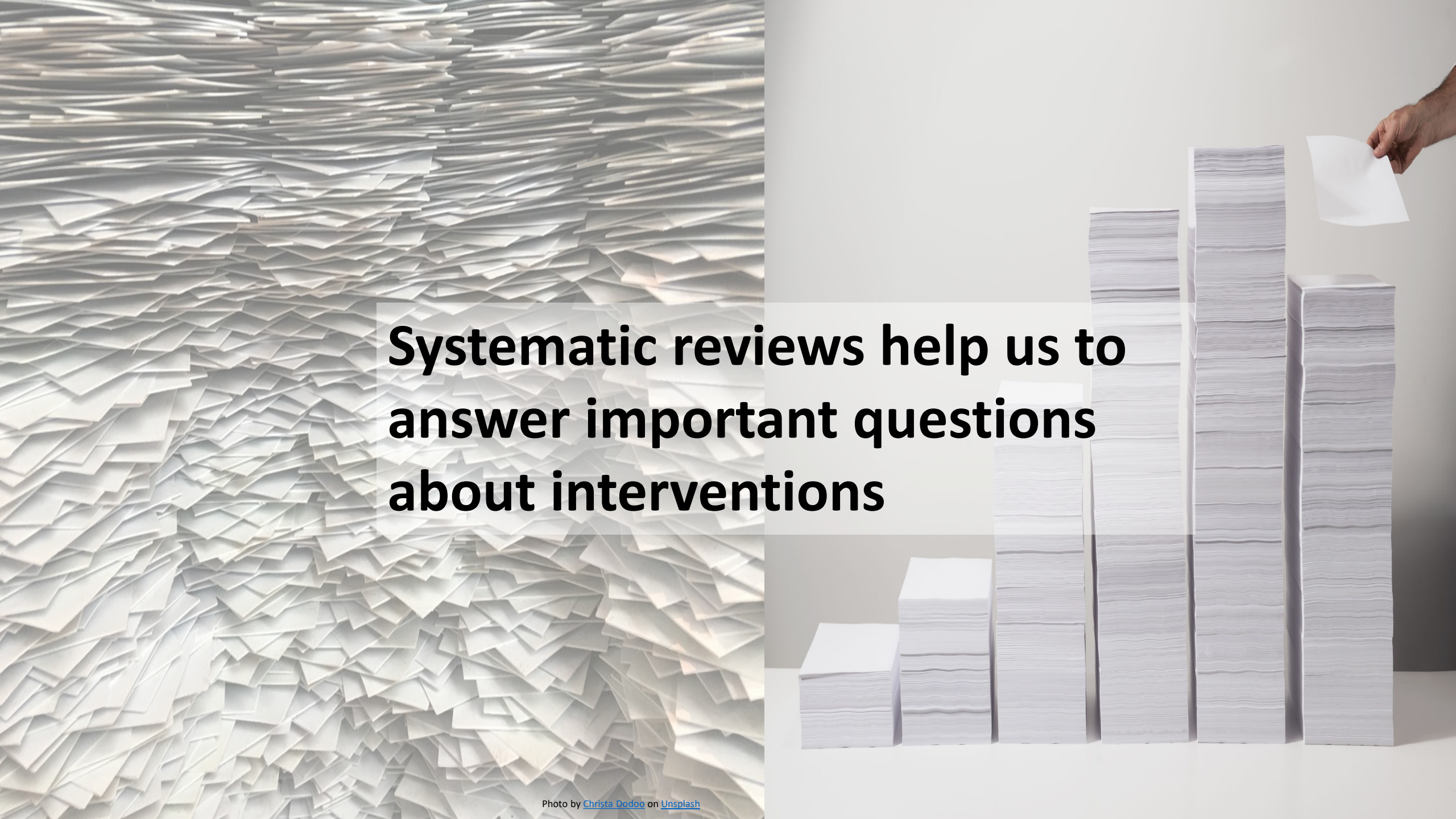
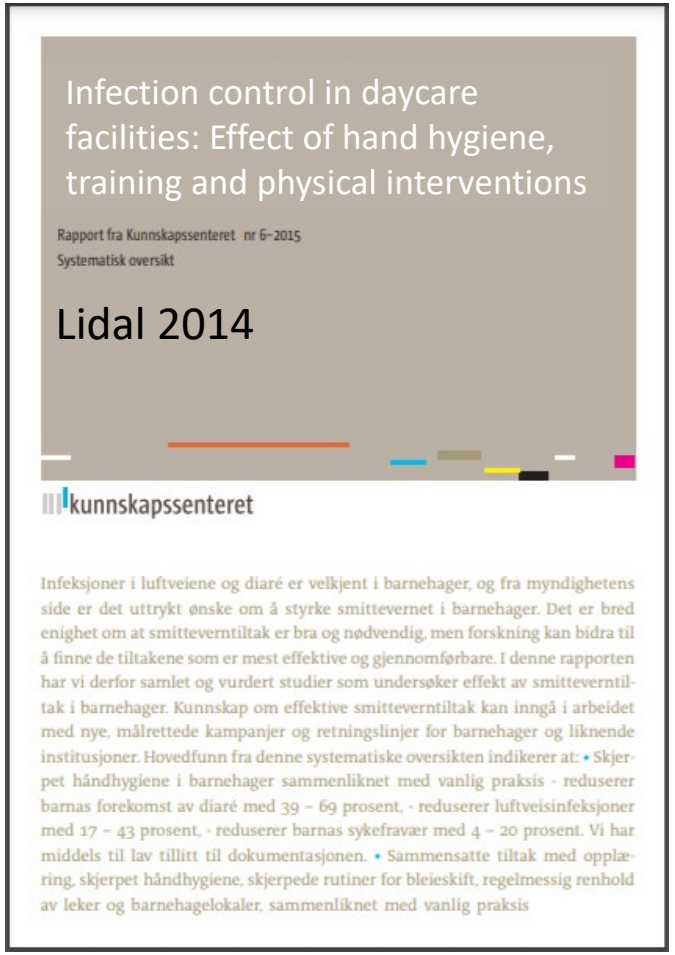


Photo by Christa Dodoo on Unsplash

Findings from systematic reviews can be used to…

…provide evidence regarding the effect of interventions

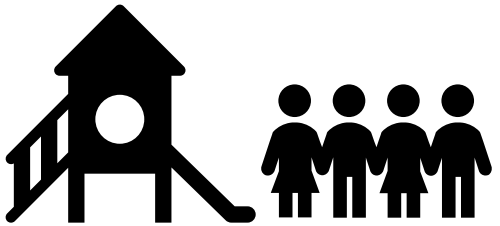

Infection control in daycare facilities: Effect of hand hygiene, training and physical interventions

Rapport fra Kunnskapssenteret nr 6–2015
Systematisk oversikt

Lidal 2014

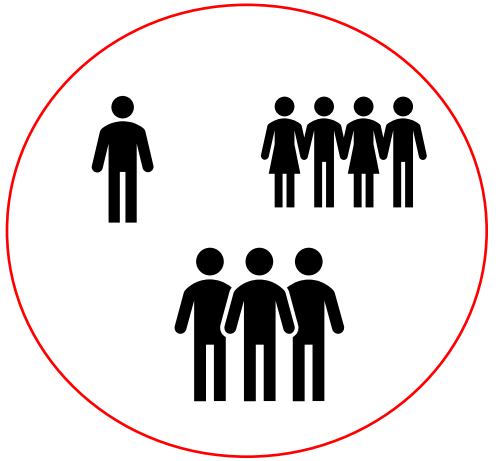
kunnskapssenteret

Infeksjoner i luftveiene og diaré er velkjent i barnehager, og fra myndighetens side er det uttrykt ønske om å styrke smittevernet i barnehager. Det er bred enighet om at smitteverntiltak er bra og nødvendig, men forskning kan bidra til å finne de tiltakene som er mest effektive og gjennomførbare. I denne rapporten har vi derfor samlet og vurdert studier som undersøker effekt av smitteverntiltak i barnehager. Kunnskap om effektive smitteverntiltak kan inngå i arbeidet med nye, målrettede kampanjer og retningslinjer for barnehager og liknende institusjoner. Hovedfunn fra denne systematiske oversikten indikerer at: • Skjerpet håndhygiene i barnehager sammenliknet med vanlig praksis - reduserer barnas forekomst av diaré med 39 – 69 prosent, - reduserer luftveisinfeksjoner med 17 – 43 prosent, - reduserer barnas sykefravær med 4 – 20 prosent. Vi har middels til lav tillitt til dokumentasjonen. • Sammensatte tiltak med opplæring, skjerpet håndhygiene, skjerpede rutiner for bleieskift, regelmessig renhold av leker og barnehagelokaler, sammenliknet med vanlig praksis

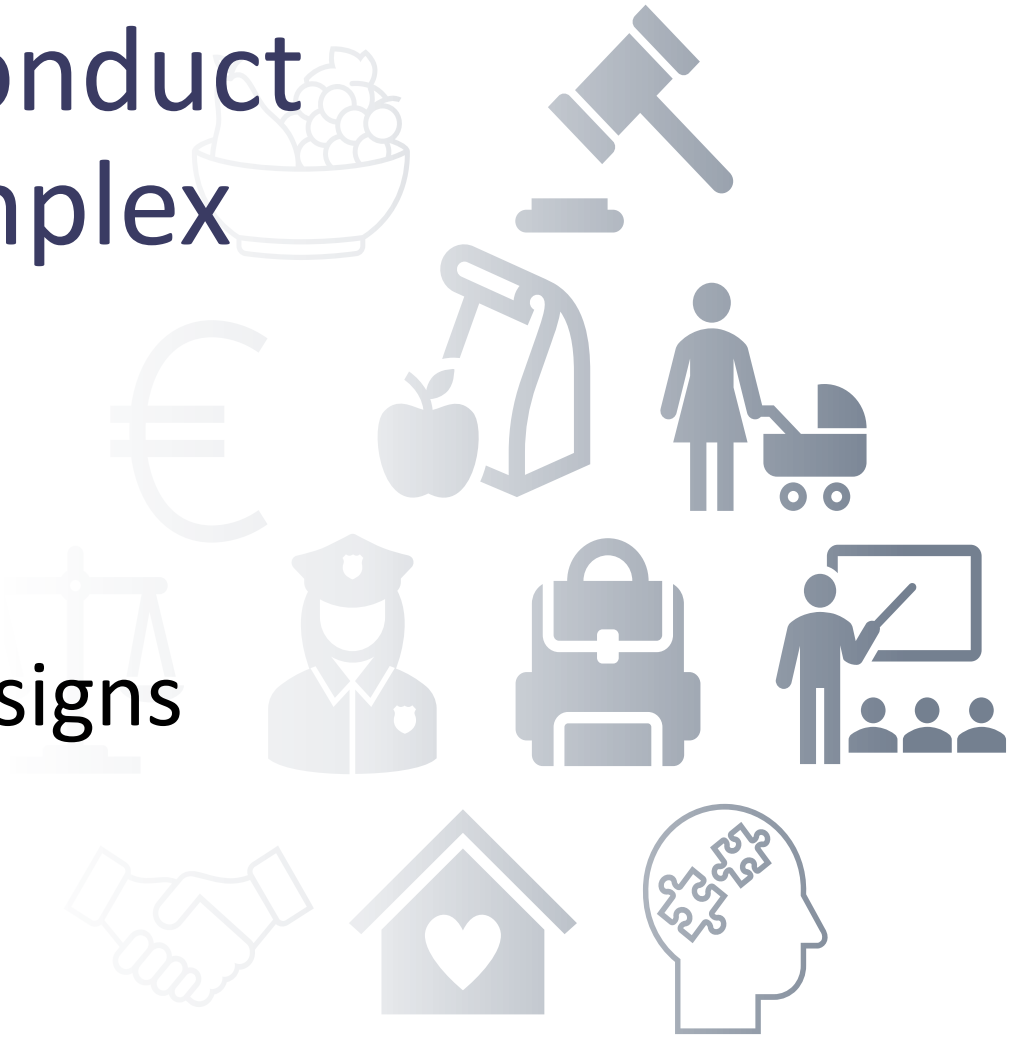


Findings from systematic reviews can be used to…

…provide evidence regarding peoples' experiences or perceptions of interventions



It can be challenging to conduct systematic reviews of complex interventions



- Complex
- Evidence from diverse study designs
- Culture

Key papers

Boon, M. H., Thomson, H., Shaw, B., Akl, E. A., Lhachimi, S. K., López-Alcalde, J., ... & GRADE Working Group. (2021). 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*, *135*, 42-53.

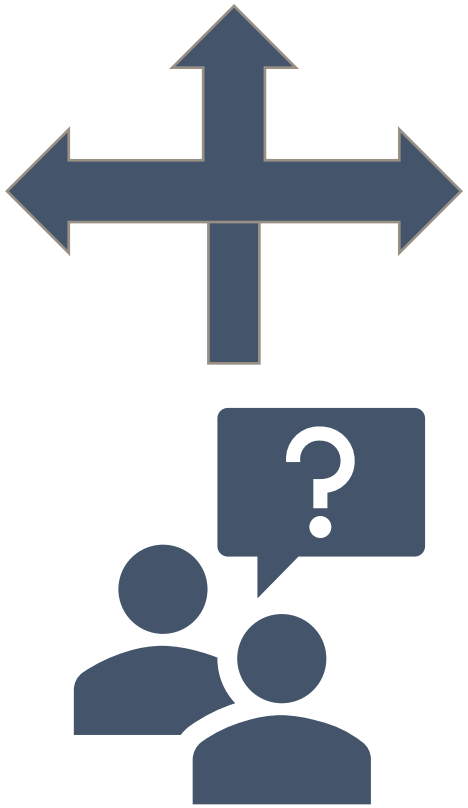
Lorenc, T., Tyner, E. F., Petticrew, M., Duffy, S., Martineau, F. P., Phillips, G., & Lock, K. (2014). Cultures of evidence across policy sectors: systematic review of qualitative evidence. *The European Journal of Public Health*, *24*(6), 1041-1047.

Mezey, G., Robinson, F., Campbell, R., Gillard, S., Macdonald, G., Meyer, D., ... & White, S. (2015). Challenges to undertaking randomised trials with looked after children in social care settings. *Trials*, *16*, 1-15.

Egan, M., Bambra, C., Petticrew, M., & Whitehead, M. (2009). Reviewing evidence on complex social interventions: appraising implementation in systematic reviews of the health effects of organisational-level workplace interventions. *Journal of Epidemiology & Community Health*, *63*(1), 4-11.

Thomson, H., Hoskins, R., Petticrew, M., Craig, N., Quinn, T., Lindsay, G., & Ogilvie, D. (2004). Evaluating the health effects of social interventions. *BMJ*, *328*(7434), 282-285.

But we need them anyway…



‘We need to be able to rely on social science and social scientists to tell us what works and why and what types of policy initiative are likely to be most effective’

(David Blunkett (then Secretary of State for Education) quoted in Boaz et al. (2002))

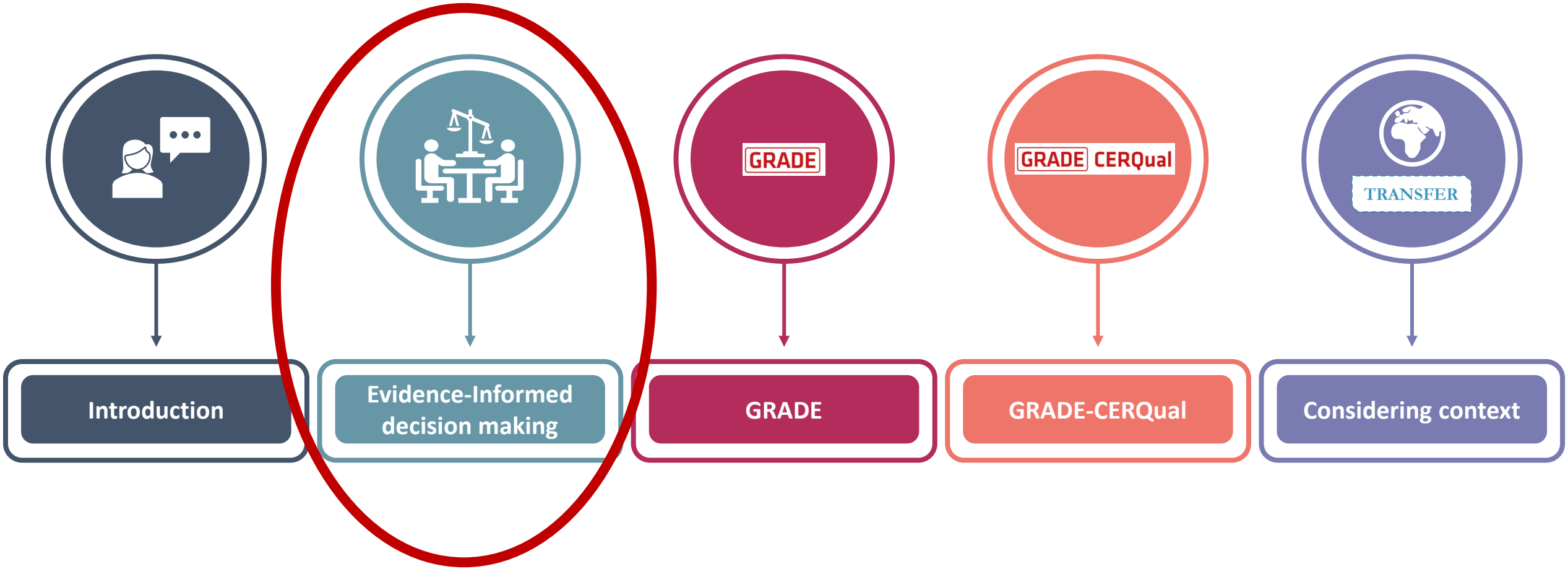
How do we know indicate how much certainty or confidence we have in the findings from systematic reviews?

GRADE

GRADE CERQual

Image from: https://blog.irisconnect.com/uk/community/blog/5-ways-to-boost-your-confidence-as-a-teacher-1/

Overview


GRADE
GRADE CERQual
TRANSFER
Introduction
Evidence-Informed decision making
GRADE
GRADE-CERQual
Considering context

NIPH
Norwegian Institute of Public Health

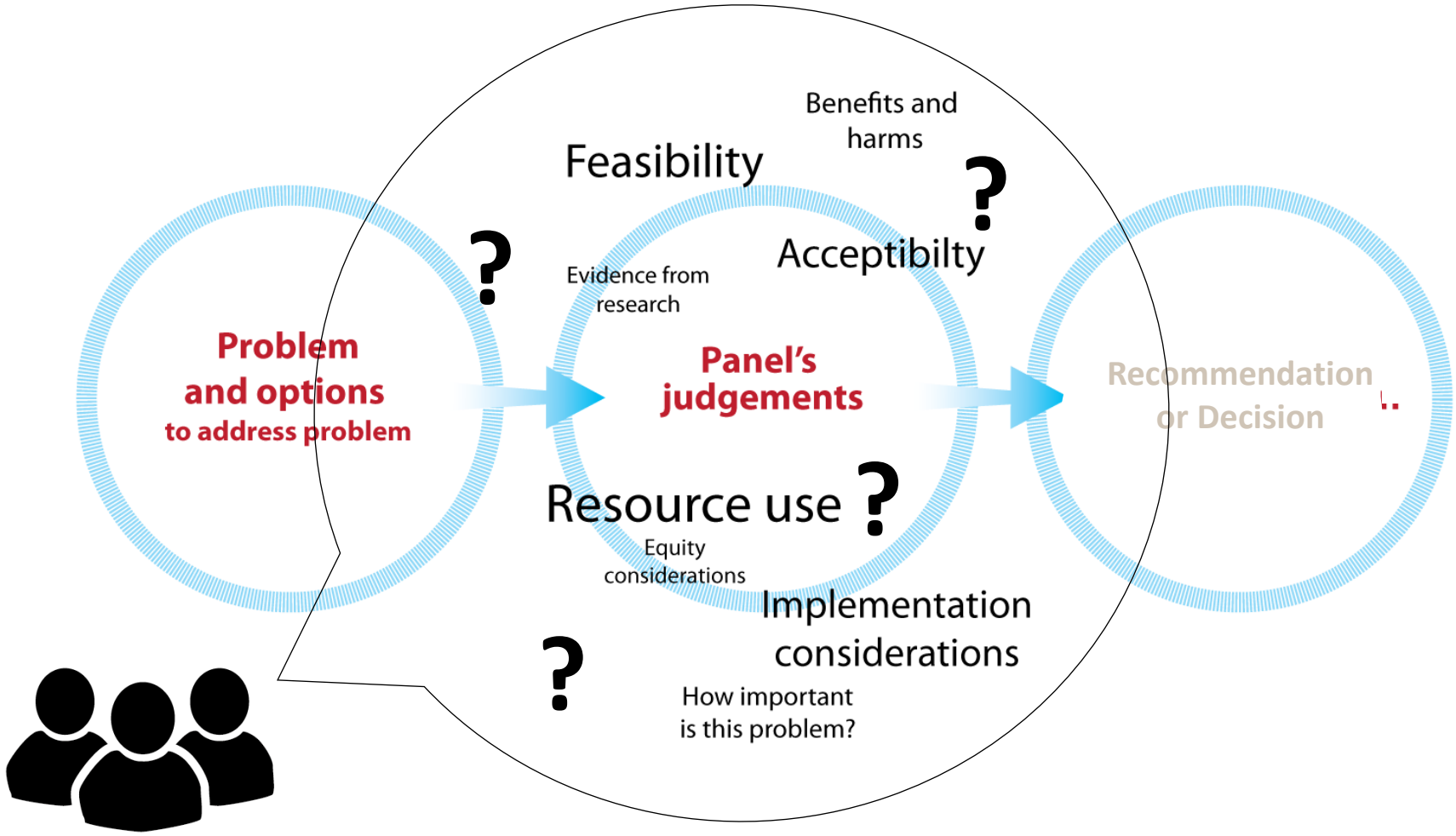
Evidence-informed decision making

Evidence-informed decision making

"Of course we'll make a decision ...
once we have considered the 5243 factors."

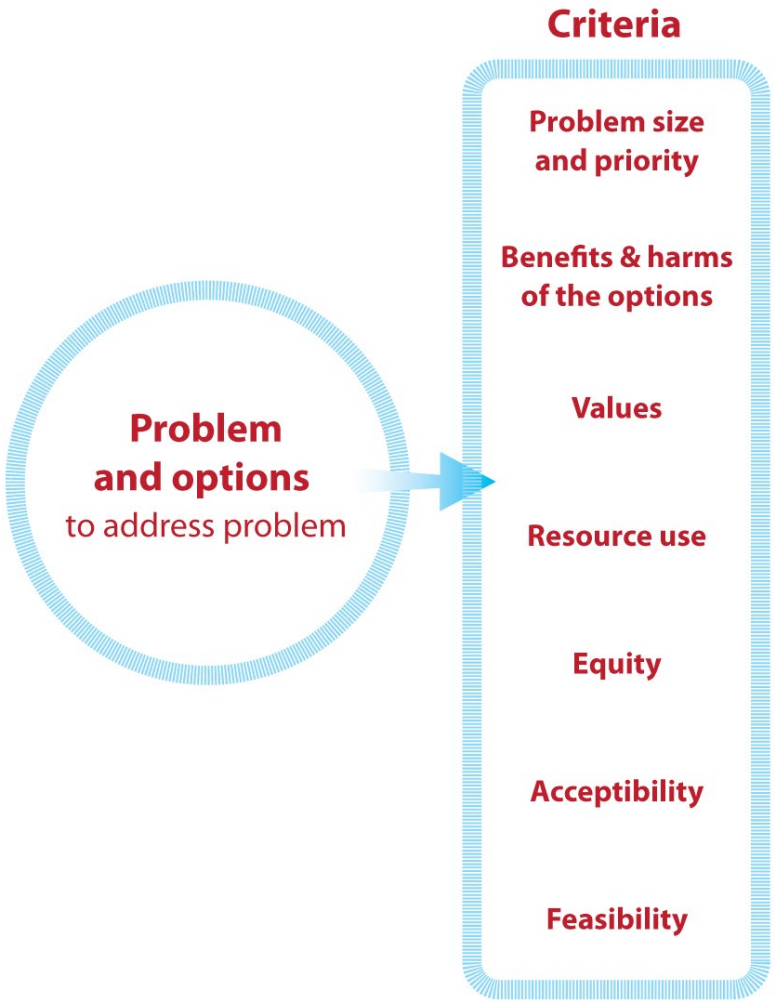
Photo by Nathalia Segato on Unsplash

Evidence-to-decision frameworks can help decision makers consider all important factors in a systematic and balanced way…

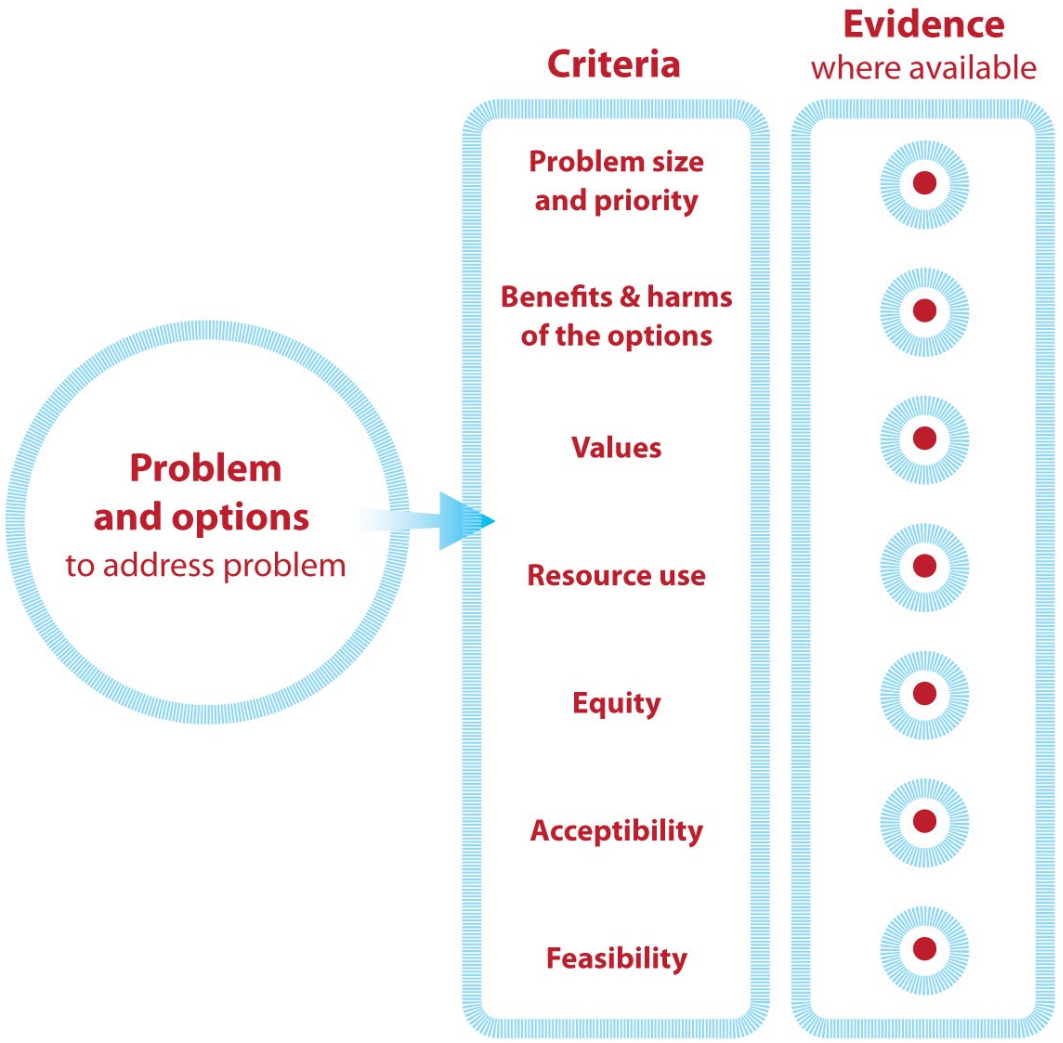


These diagrams on the GRADE Evidence-to-Decision frameworks were developed by the DECIDE Project and the GRADE Working Group

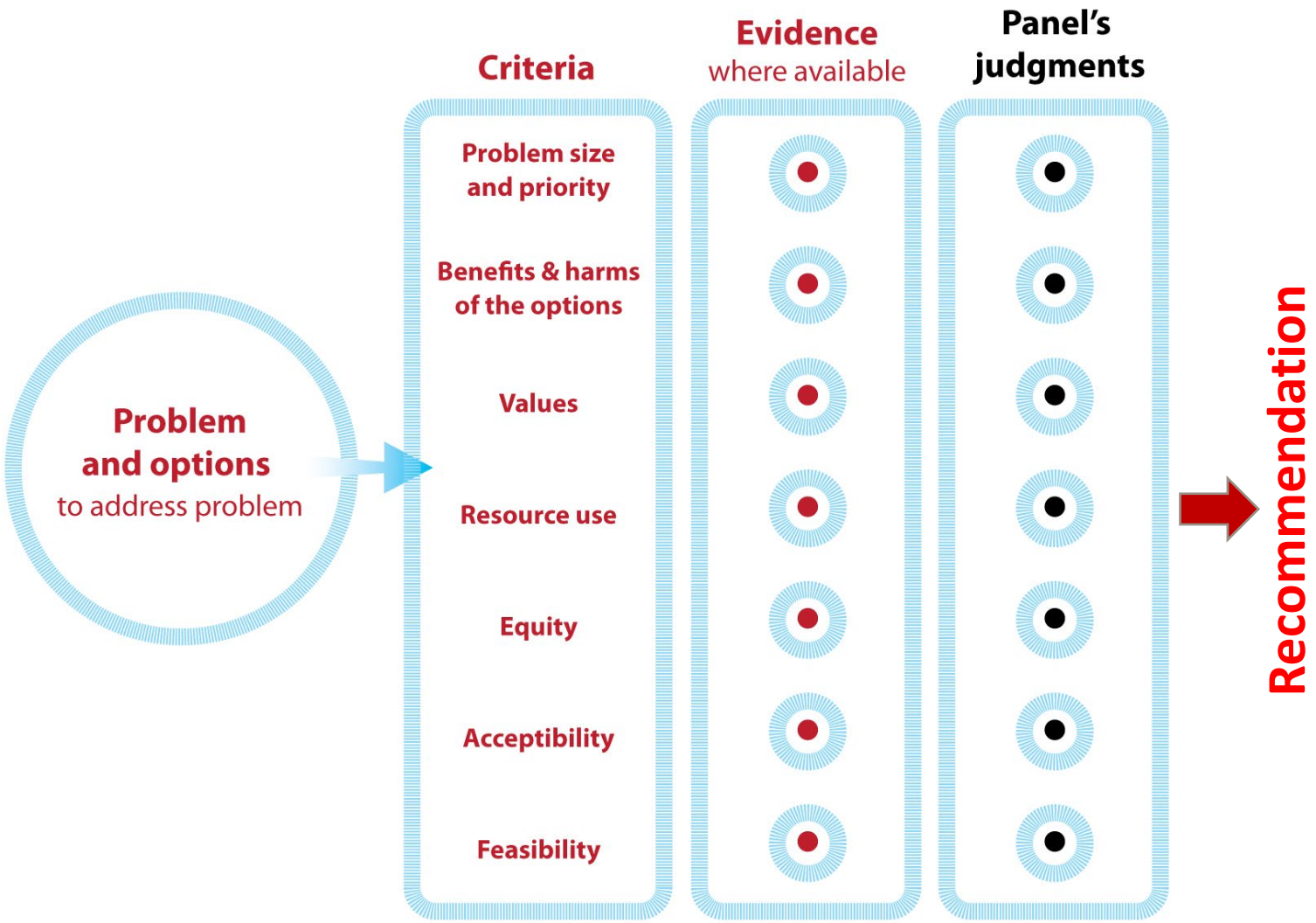
….evidence-to-decision frameworks guide decision makers through different pre-specified criteria…..



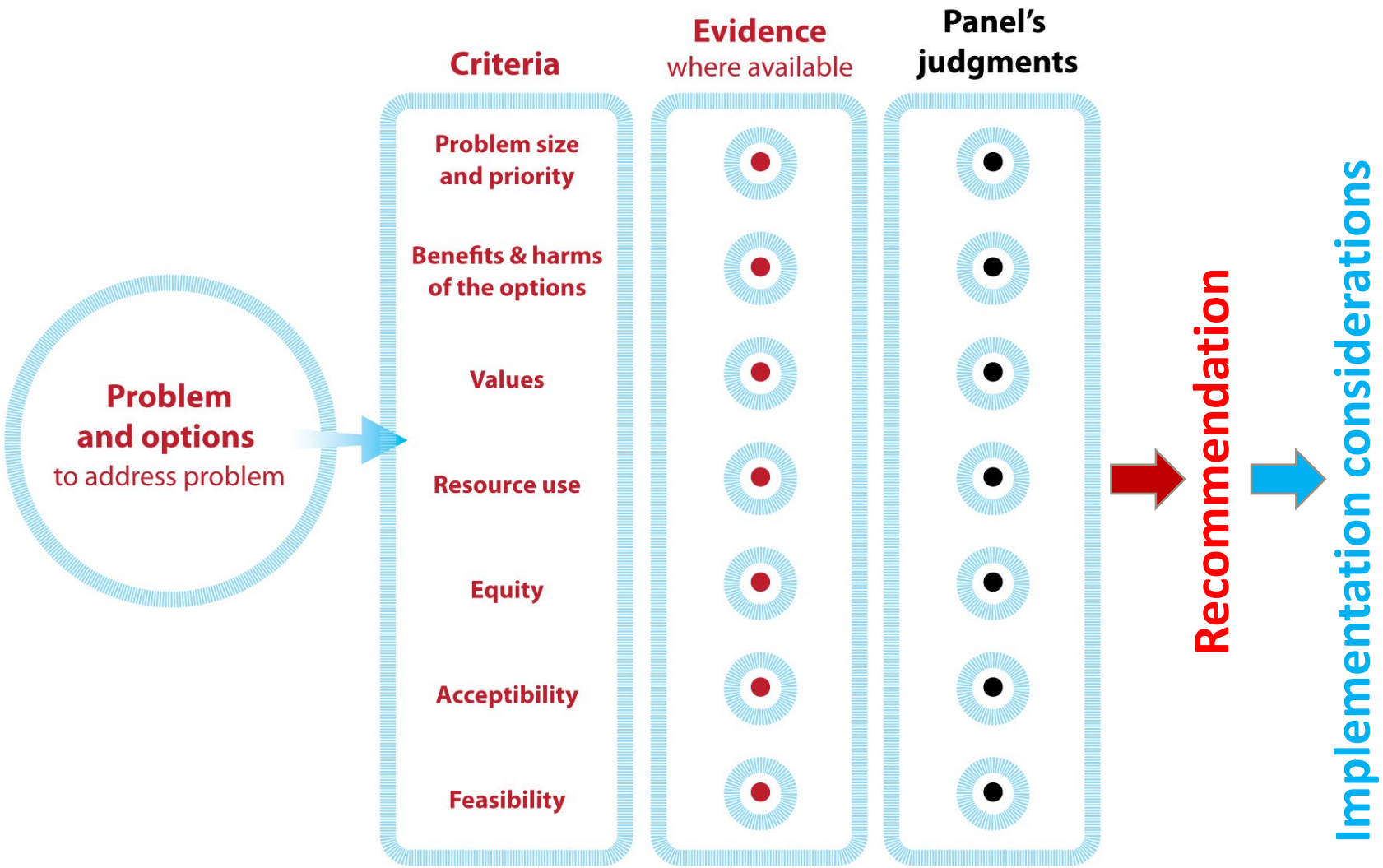
….using the most appropriate evidence for each criterion…



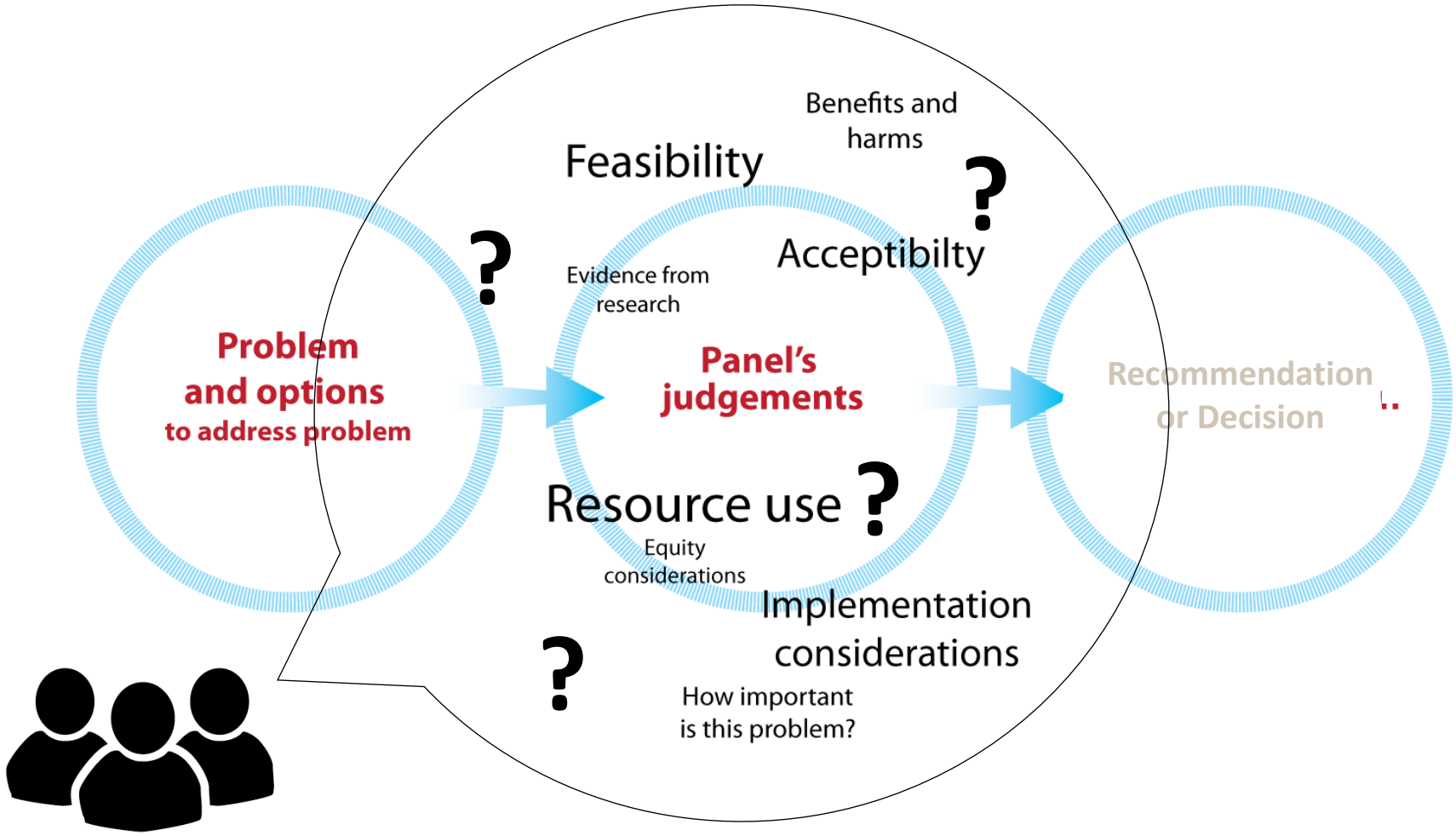
...before making a final judgments and a recommendation or decision…..



...and suggesting implementation considerations

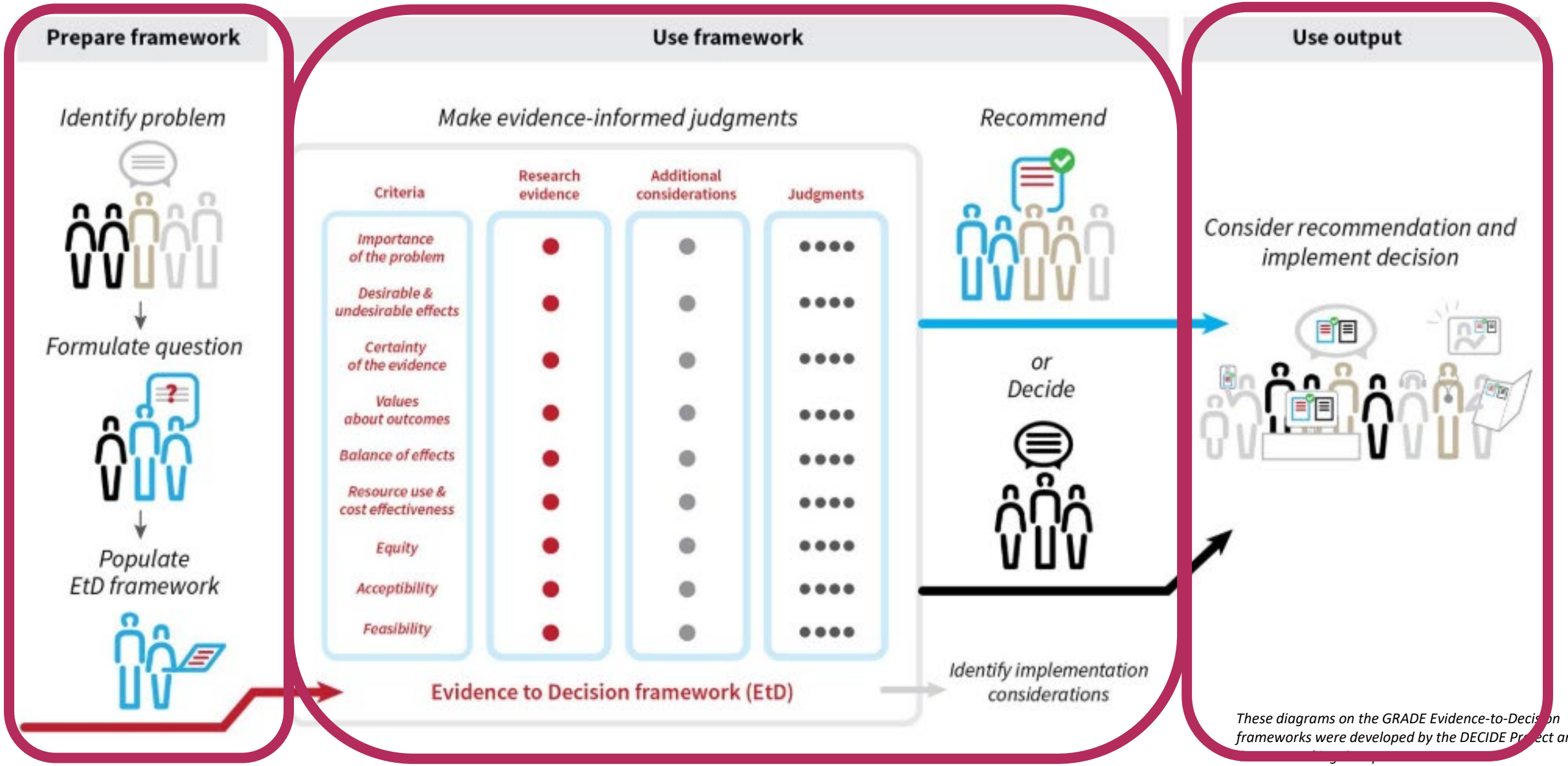


So we can og from this…



These diagrams on the GRADE Evidence-to-Decision frameworks were developed by the DECIDE Project and the GRADE Working Group

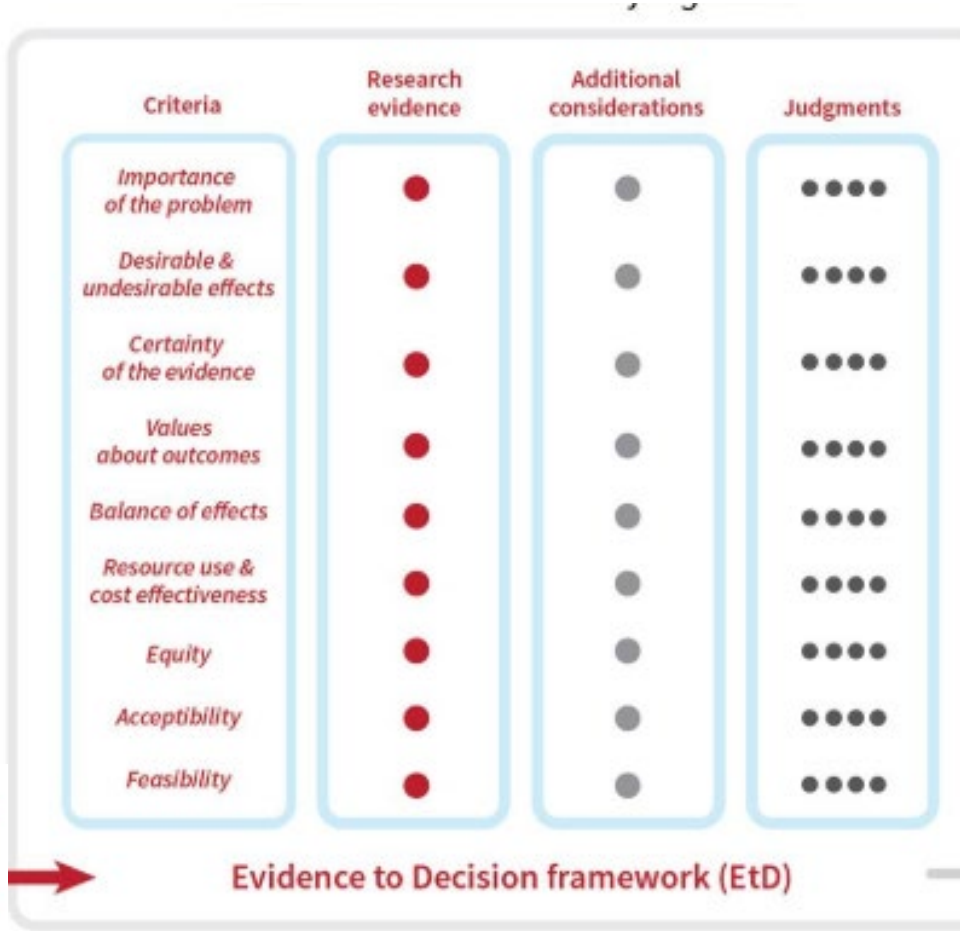
...to a more systematic and transparent assessment of relevant criteria



These diagrams on the GRADE Evidence-to-Decision frameworks were developed by the DECIDE Project and the

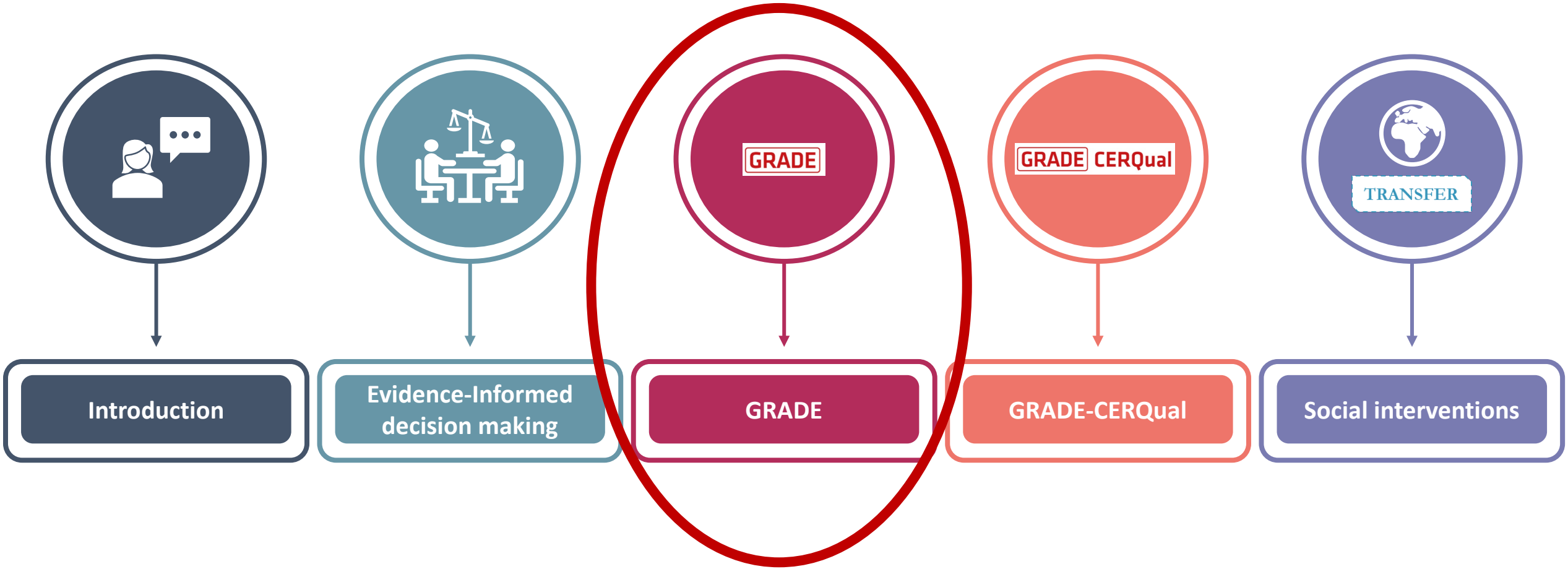
Confidence!!

Certainty!!



These diagrams on the GRADE Evidence-to-Decision frameworks were developed by the DECIDE Project and the GRADE Working Group

Overview


GRADE
GRADE CERQual
TRANSFER
Introduction
Evidence-Informed decision making
GRADE
GRADE-CERQual
Social interventions

NIPH
Norwegian Institute of Public Health

GRADE

Assessing certainty of the evidence

The GRADE approach

What? Why? How?

An approach used to rate certainty of evidence about the effect of an intervention

Considers:

- Quality of evidence
- Individual review outcomes
- Magnitude of effect

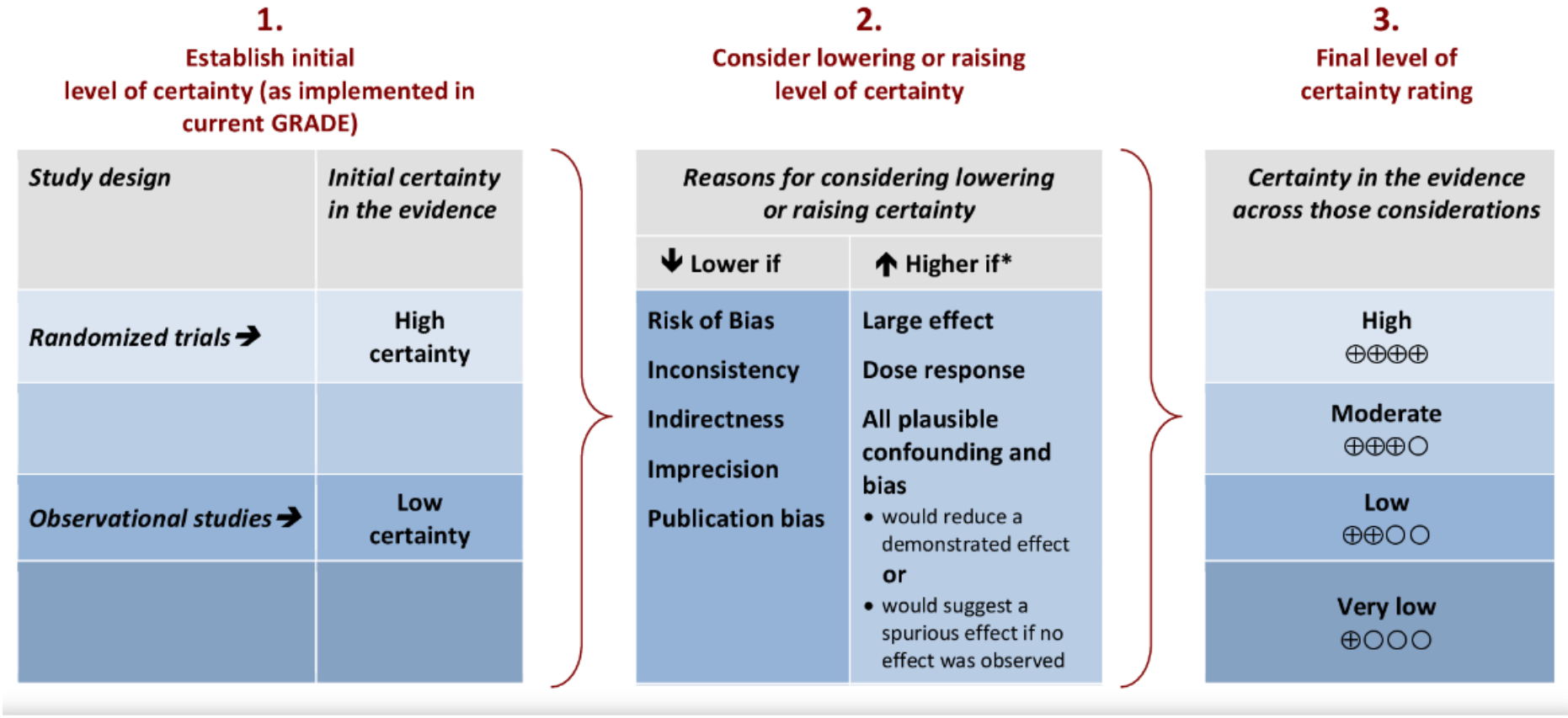
Wall Street journal

I figure there's a 40% chance of showers and a 10% chance we know what we're talking about"

GRADE

- Risk of bias
- Inconsistency
- Indirectness
- Imprecision
- Publication bias

GRADE



1. Establish initial level of certainty (as implemented in current GRADE)

Study design	*Initial certainty in the evidence*
Randomized trials ➔	High certainty
Observational studies ➔	Low certainty

2. Consider lowering or raising level of certainty

Reasons for considering lowering or raising certainty	
↓ Lower if	↑ Higher if*
Risk of Bias	Large effect
Inconsistency	Dose response
Indirectness	All plausible confounding and bias • would reduce a demonstrated effect or • would suggest a spurious effect if no effect was observed
Imprecision	
Publication bias	

3. Final level of certainty rating

Certainty in the evidence across those considerations
High ⊕⊕⊕⊕
Moderate ⊕⊕⊕◯
Low ⊕⊕◯◯
Very low ⊕◯◯◯

Table 1. Use of GRADE not considering ROBINS-I and similar tools: According to GRADE, certainty, quality, strength of the evidence or the confidence in the estimate of effect, is determined for each outcome based on a… Expand

Published in 2018

GRADE Working Group (2019). GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. Clinical Epidemiology 105-114.

H. Schünemann, E. Akl, R. Mustafa, G. Gartlehner, R. Kunz, G. Guyatt

GRADE Evidence Profile

Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Profile for high intensity case management compared to low intensity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving housing stability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Munthe-Kaas, H. M., Berg, R. C., & Blaasvær, N. (2018). Effectiveness of interventions to reduce homelessness: a systematic review and meta-analysis. *Campbell Systematic Reviews*, *14*(1), 1-281.

Study design

Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Profile for high intensity case management compared to low intensity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving housing stability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Risk of Bias

Risk of bias - randomized trials

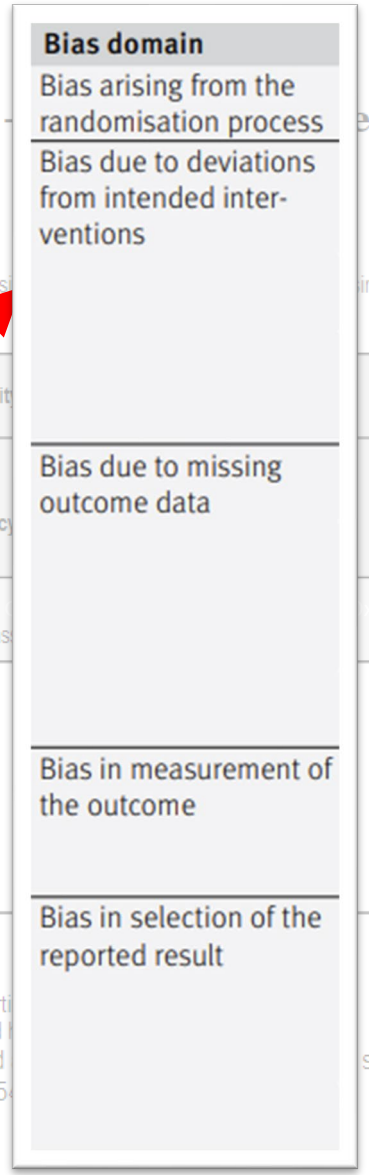


Table 8.1.2. Comparison 1.A.2 - … Profile for high intensity case management compared to low intensity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intens… …ing stability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality						№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency		Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; as…											
2	randomised trials	serious [1]	Serious[3]		none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporti…
2. Wide confidence intervals which include benefits and h…
3. Inconsistency between results from the pooled … study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.5…

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *bmj*, *366*.

ROBINS-I

Risk of bias - Non-randomized studies

Table 8.1.2. Comparison 1.A.2 – rofile for high intensity case management compared to low intensity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low inte ability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality a						№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency		Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; asse											
2	randomised trials	serious [1]	Serious[3]		none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting
2. Wide confidence intervals which include benefits and ha
3. Inconsistency between results from the pooled a y that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54,

Domain
Pre-intervention
Bias due to confounding
Bias in selection of participants into the study
At intervention
Bias in classification of interventions
Post-intervention
Bias due to deviations from intended interventions
Bias due to missing data
Bias in measurement of outcomes
Bias in selection of the reported result

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *bmj, 355*.

Inconsistency

Consider in a meta-analysis
- Variation in effect size
- Confidence intervals
- Statistical test for heterogeneity is $p<0.05$
- I^2 is large

Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Profile for high int… management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving ho…g stability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Inconsistency

Study or Subgroup	HICM Events	HICM Total	Usual services Events	Usual services Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Bond 1990	30	42	17	40	24.6%	1.68 [1.12, 2.53]
Garety 2006	59	71	54	73	75.4%	1.12 [0.95, 1.33]
Total (95% CI)		**113**		**113**	**100.0%**	**1.26 [1.07, 1.49]**
Total events	89		71			
Heterogeneity: Chi² = 3.64, df = 1 (P = 0.06); I² = 73%						
Test for overall effect: Z = 2.69 (P = 0.007)						

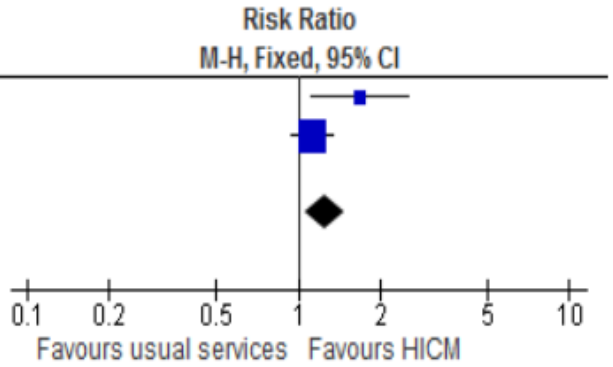


Table 8.1.2. Comparison 1.A.2 – GRADE I management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management f
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⊕⊕◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Inconsistency

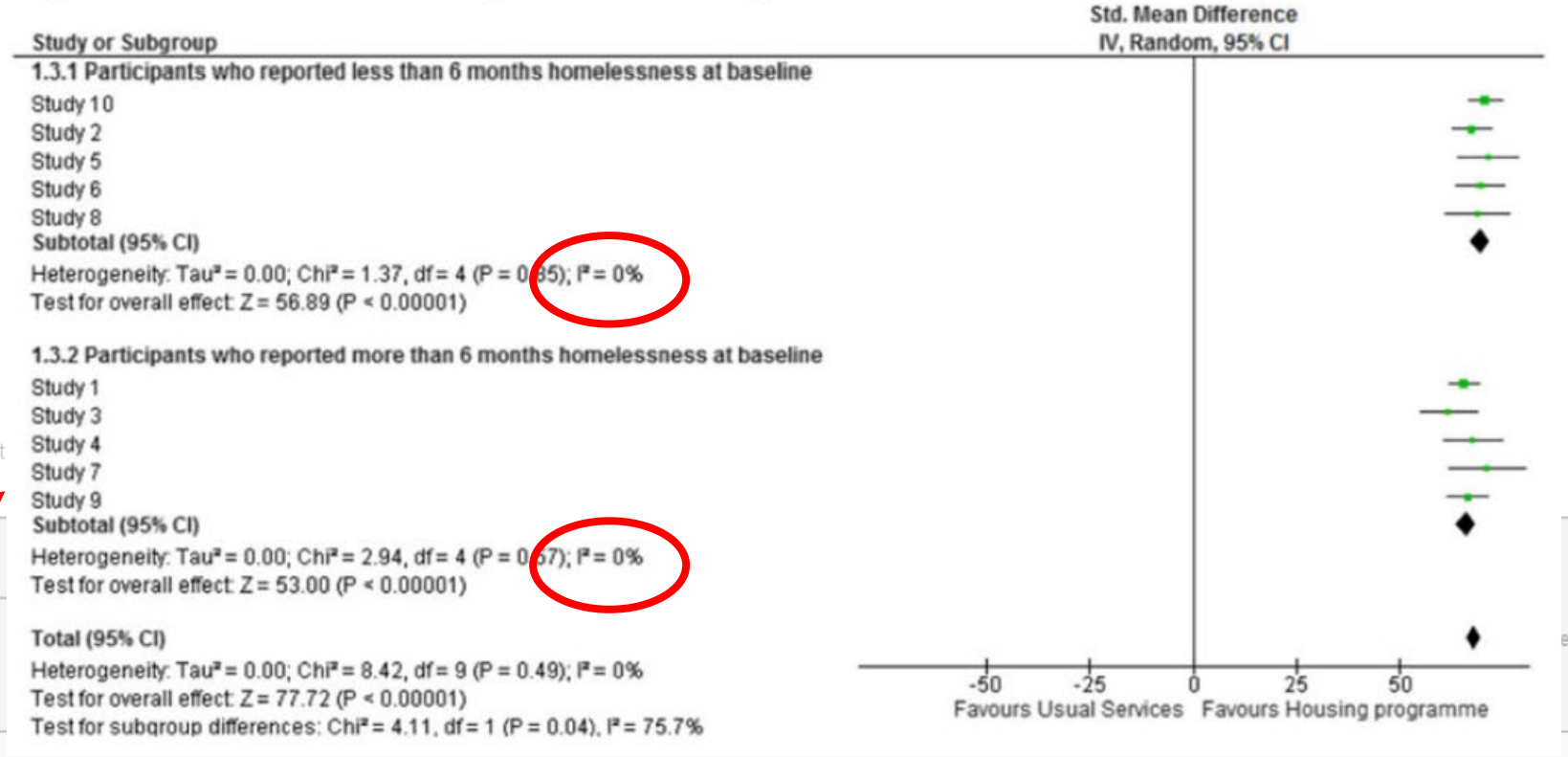


Table 8.1.2. Comparison 1.A.2 – GRADE management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

	Quality assessment											
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Large effect	Plausible confounding	Number of participants: High intensity case management	Number of participants: Low intensity case management	Effect	Quality
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none			204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Indirectness

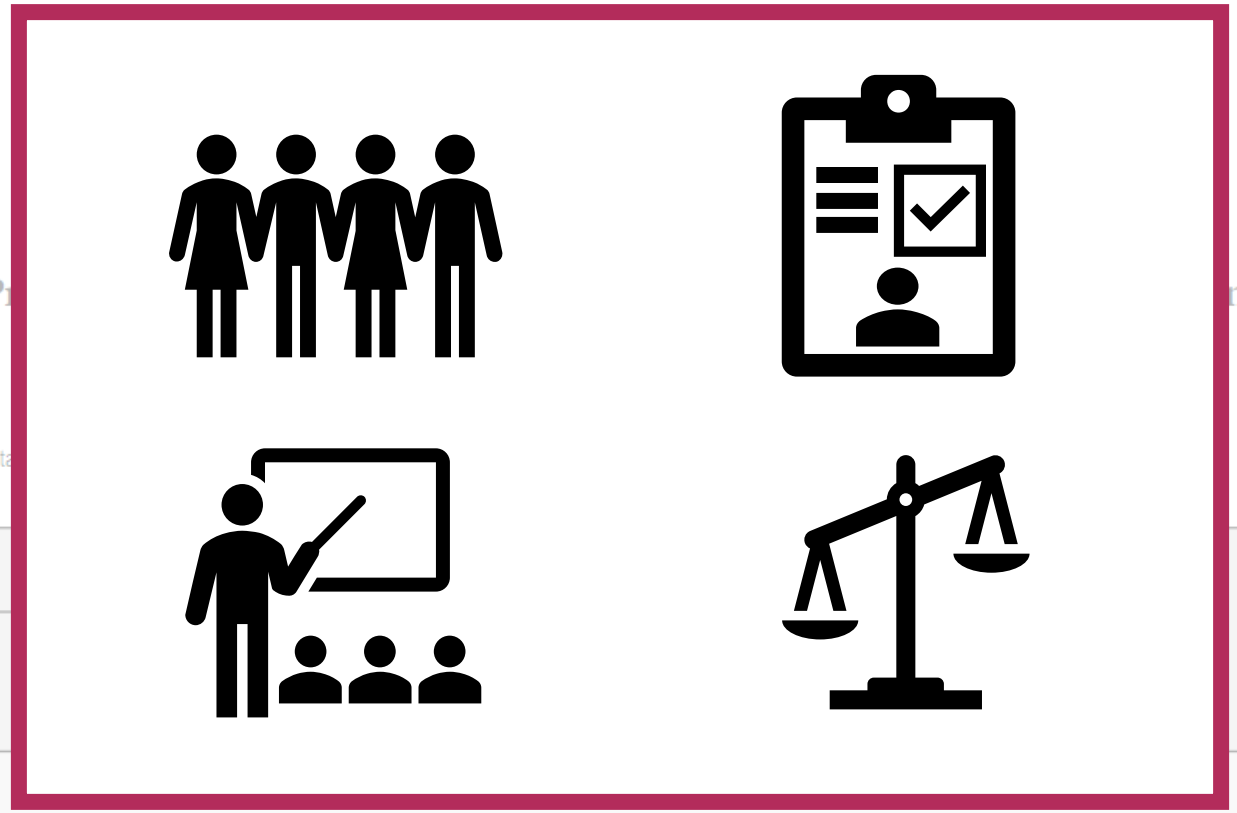


Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Pr… …nsity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving housing sta…
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment												
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision							Importance
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Imprecision

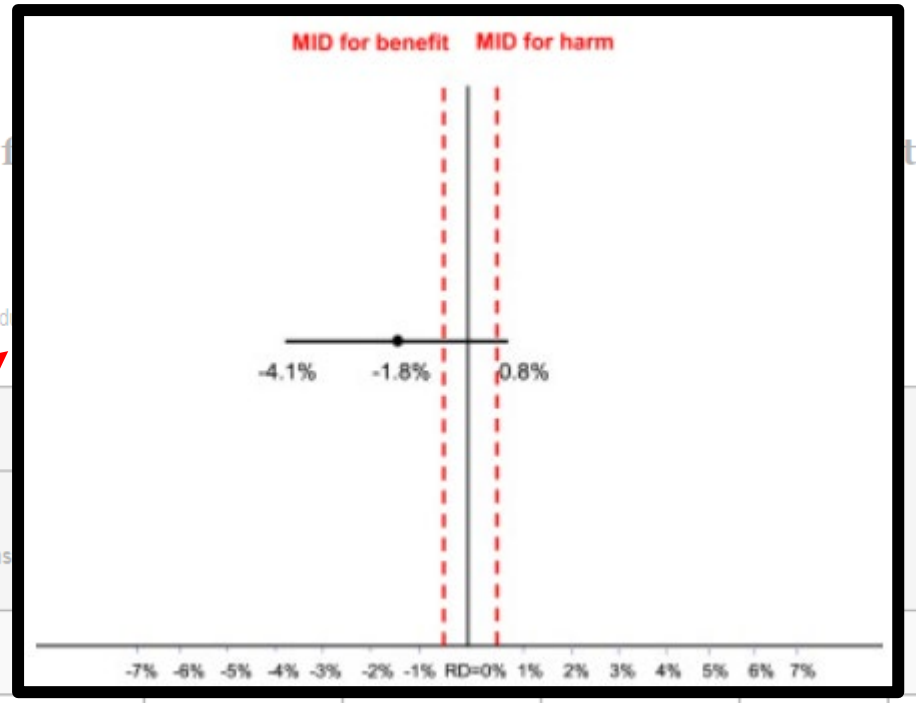


Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Profile f … to low intensity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving housing stability and red
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

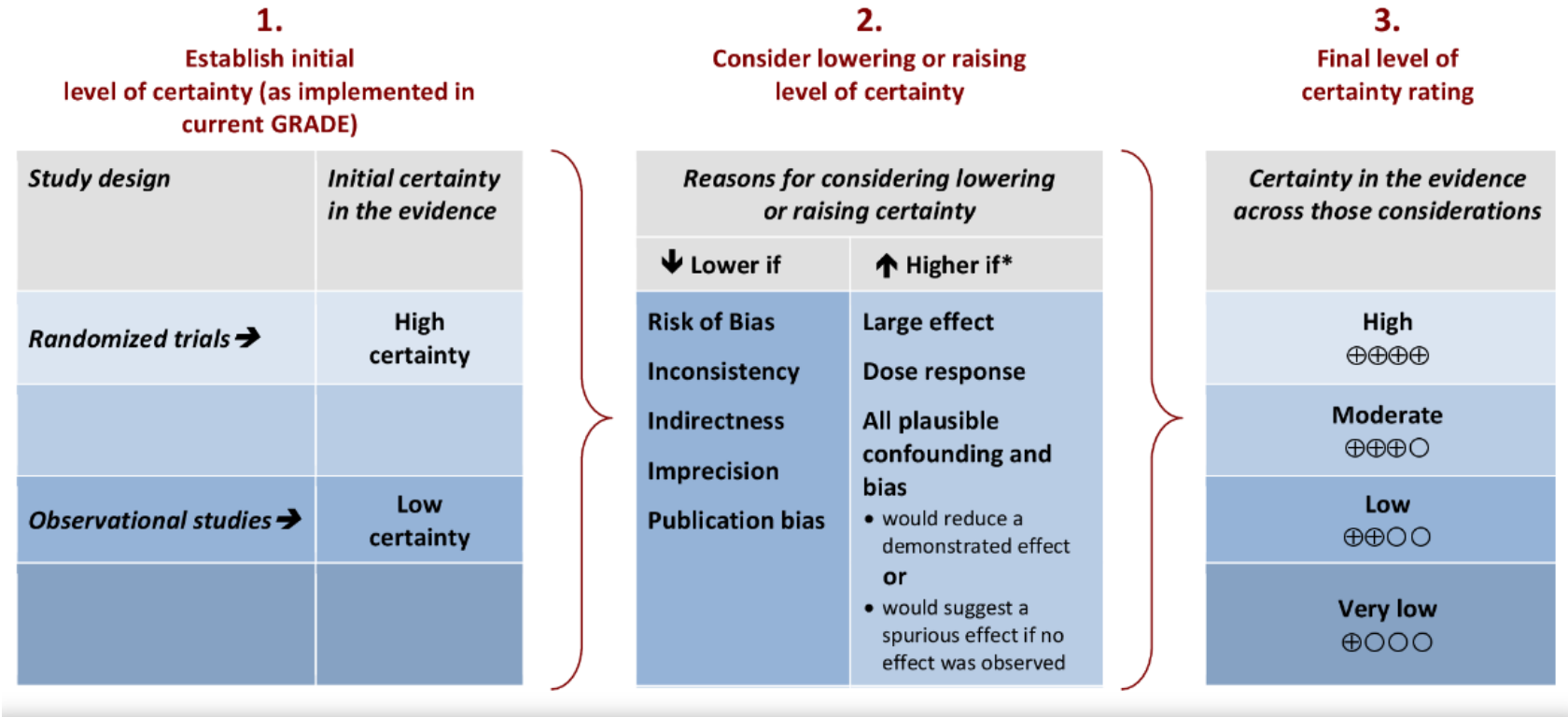
Quality assessment												
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other cons					Quality	Importance
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⊕⊕◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

Schünemann, H. J et al. (2022). GRADE guidance 35: update on rating imprecision for assessing contextualized certainty of evidence and making decisions. *Journal of Clinical Epidemiology*, *150*, 225-242.

GRADE overall assessment



1. Establish initial level of certainty (as implemented in current GRADE)		2. Consider lowering or raising level of certainty		3. Final level of certainty rating
Study design	*Initial certainty in the evidence*	*Reasons for considering lowering or raising certainty*		*Certainty in the evidence across those considerations*
		↓ Lower if	↑ Higher if*	
Randomized trials ➔	High certainty	Risk of Bias Inconsistency Indirectness Imprecision Publication bias	Large effect Dose response All plausible confounding and bias • would reduce a demonstrated effect or • would suggest a spurious effect if no effect was observed	High ⊕⊕⊕⊕
				Moderate ⊕⊕⊕◯
Observational studies ➔	Low certainty			Low ⊕⊕◯◯
				Very low ⊕◯◯◯

Table 1. Use of GRADE not considering ROBINS-I and similar tools: According to GRADE, certainty, quality, strength of the evidence or the confidence in the estimate of effect, is determined for each outcome based on a... Expand

Published in 2018

GRADE Working Group (2019). GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. Clinical Epidemiology 105-114.

H. Schünemann, E. Akl, R. Mustafa, G. Gartlehner, R. Kunz, G. Guyatt

Transparency…

Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Profile for high intensity case management compared to low intensity case management

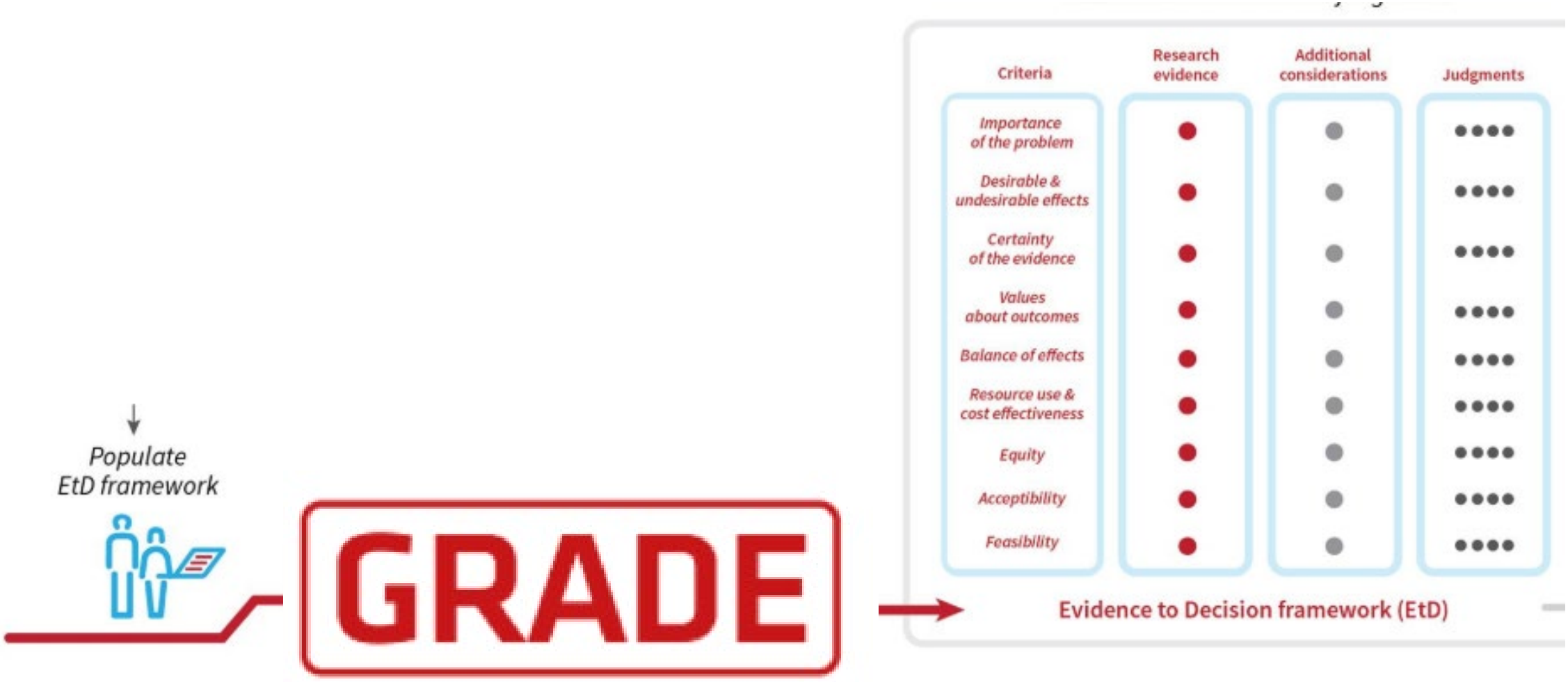
Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving housing stability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⊕⊕◯◯ LOW	

CI: Confidence interval; **SMD:** Standardised mean difference

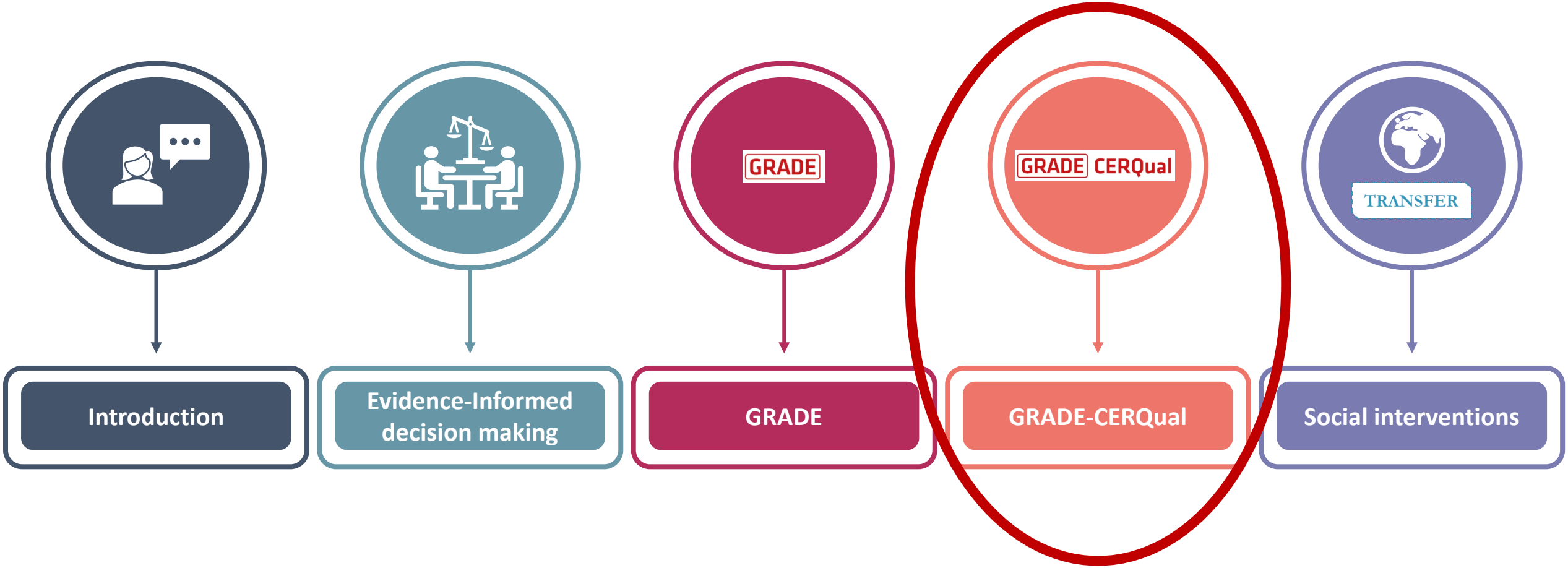
1. Risk of detection bias in one study. Inadequate reporting of methods in both studies.
2. Wide confidence intervals which include benefits and harms.
3. Inconsistency between results from the pooled analysis (two studies) and the third study that could not be included in the pooled analysis (Morse 1997).The third study reported that participants in the intervention group reported more days in stable housing than the control group (F=3.54, df=2, 129, p<0.032)

GRADE in the Evidence-to-Decision framework



These diagrams on the GRADE Evidence-to-Decision frameworks were developed by the DECIDE Project and the GRADE Working Group

Overview


GRADE
GRADE CERQual
TRANSFER
Introduction
Evidence-Informed decision making
GRADE
GRADE-CERQual
Social interventions

NIPH
Norwegian Institute of Public Health

Assessing confidence in the evidence from reviews of qualitative research

GRADE-CERQual

What is qualitative evidence?

Qualitative research aims to describe the social world; understand people's views, experiences and motivations; and often to explain the social world by developing hypotheses, theories or models

Common methods for qualitative research:

- Focus groups
- Individual, semi-structured interviews
- (Participant) observation
- Document analysis

Systematic reviews of qualitative research (or "qualitative evidence syntheses") identify and synthesize these types of studies. These syntheses are becoming increasingly popular

How do qualitative evidence syntheses differ from reviews of effectiveness?

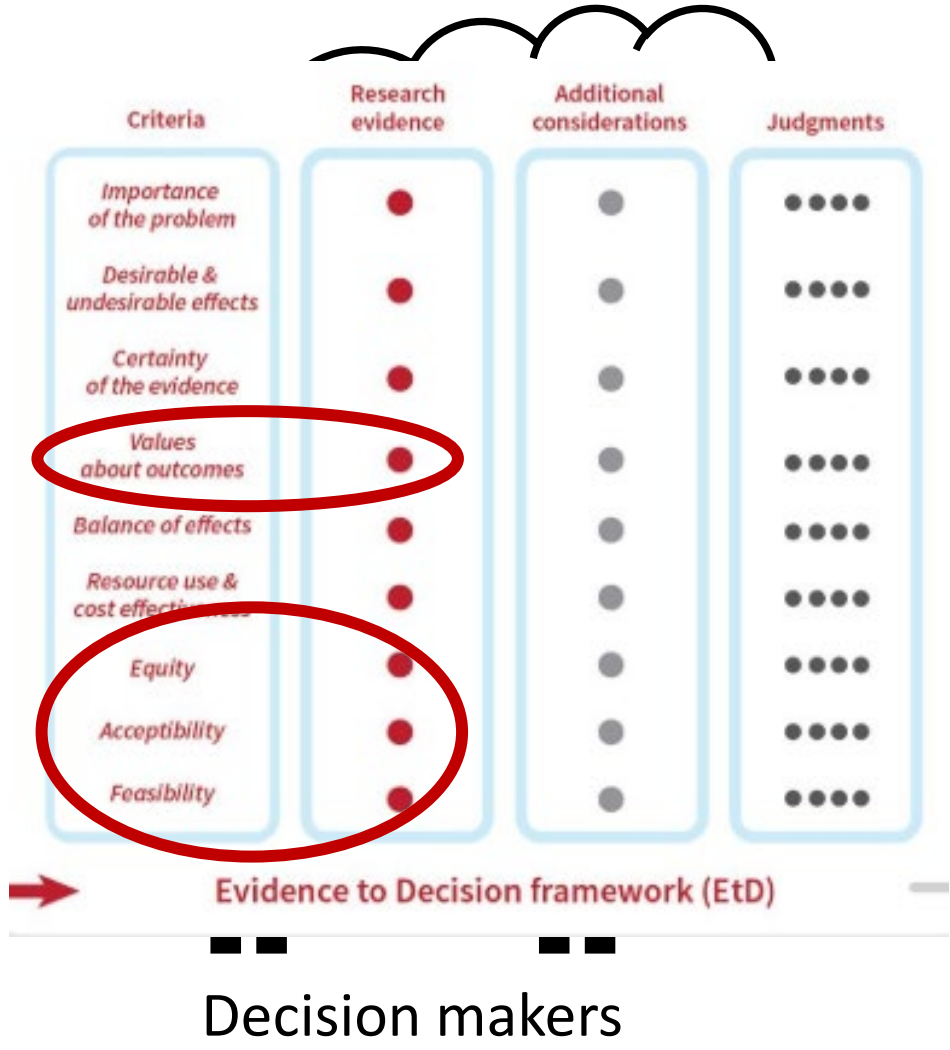
The main structure is broadly similar

We carry out systematic searches for relevant qualitative studies → We assess the quality of and extract data from the studies that are included → We synthesise this data

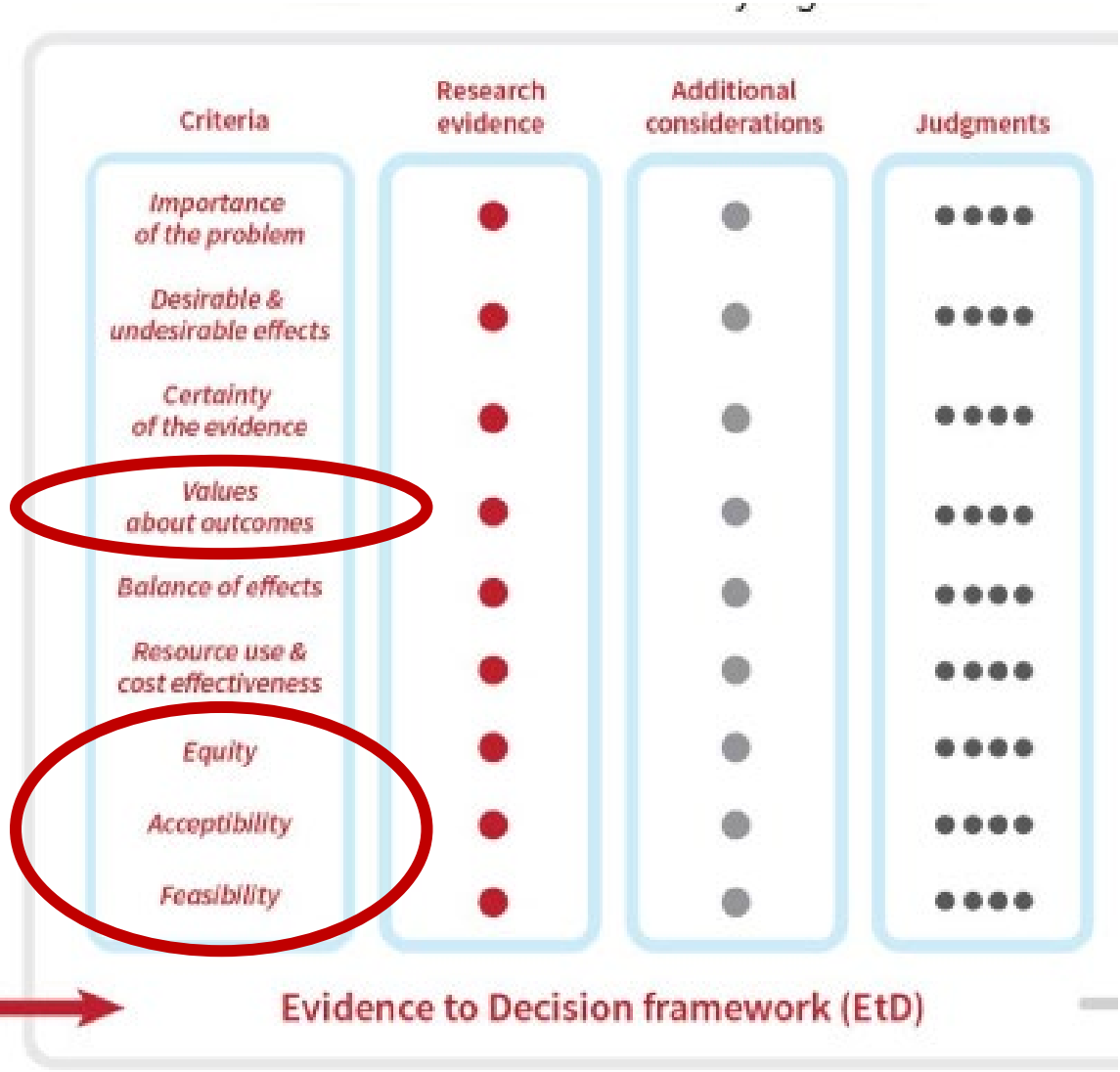
But follows principles appropriate for qualitative research

Qualitative research in decision making…

- Systematic reviews of qualitative research (also called qualitative evidence syntheses) are increasingly common
- Increasingly used in guideline or policy development processes
- Decision makers need methods to assess how much confidence to place in findings from these reviews
- Decision makers likely to make these judgements anyway –helpful to provide a systematic and transparent way for doing this



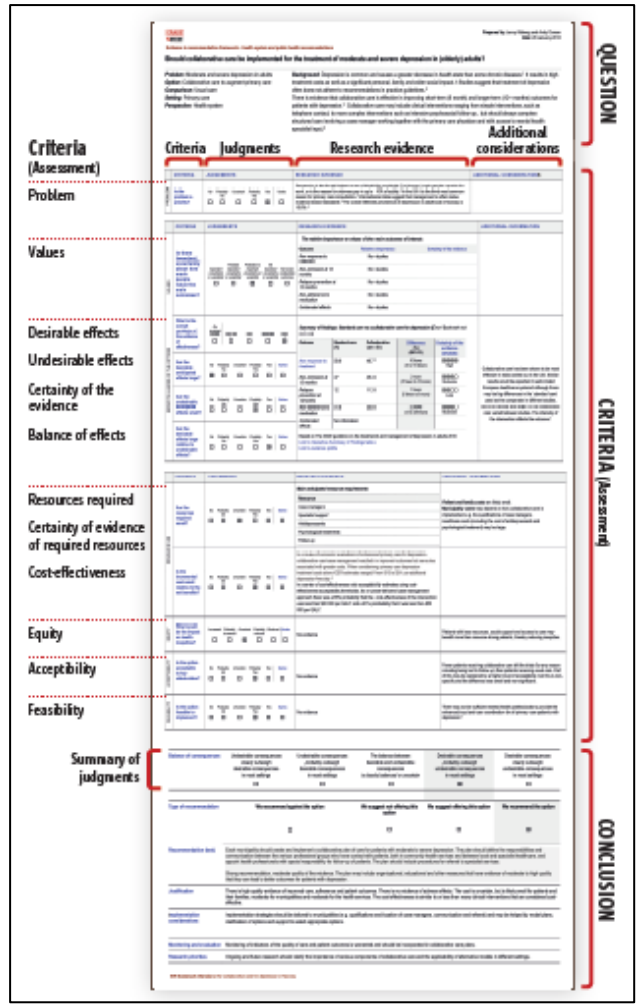
GRADE-CERQual in the Evidence-to-Decision framework



These diagrams on the GRADE Evidence to Decision frameworks were developed by the DECIDE Project and the GRADE Working Group

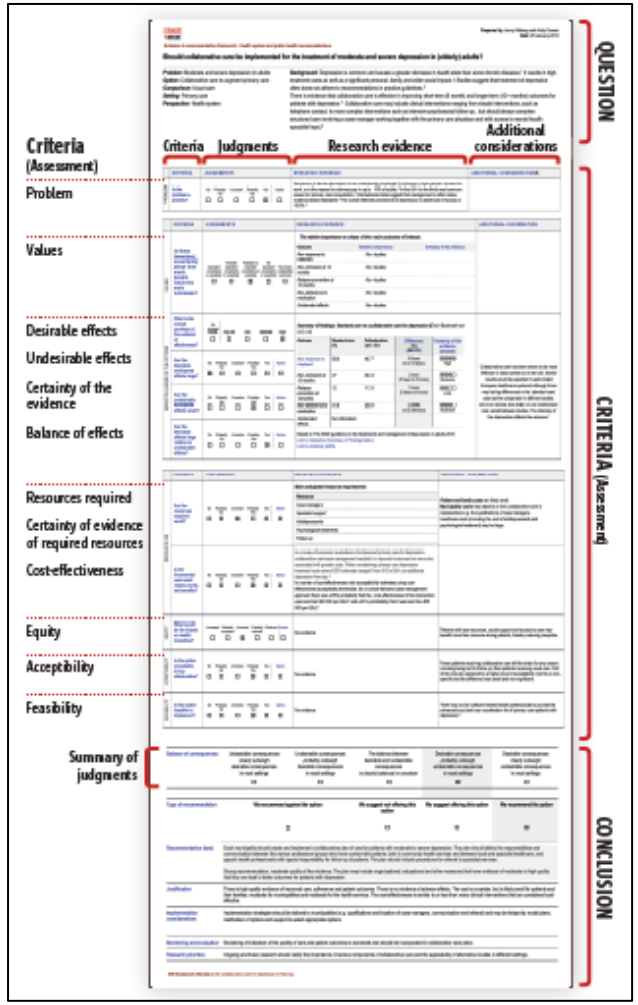
GRADE Evidence-to-Decision Frameworks and qualitative evidence

How people value the outcomes	Consider whether there are differences in, or uncertainties about, how stakeholders value the outcomes
Acceptability of the intervention	Consider the extent to which an intervention is considered to be reasonable, satisfactory or adequate to relevant stakeholders
Feasibility of the intervention	Consider the extent to which an intervention is capable of being accomplished or implemented

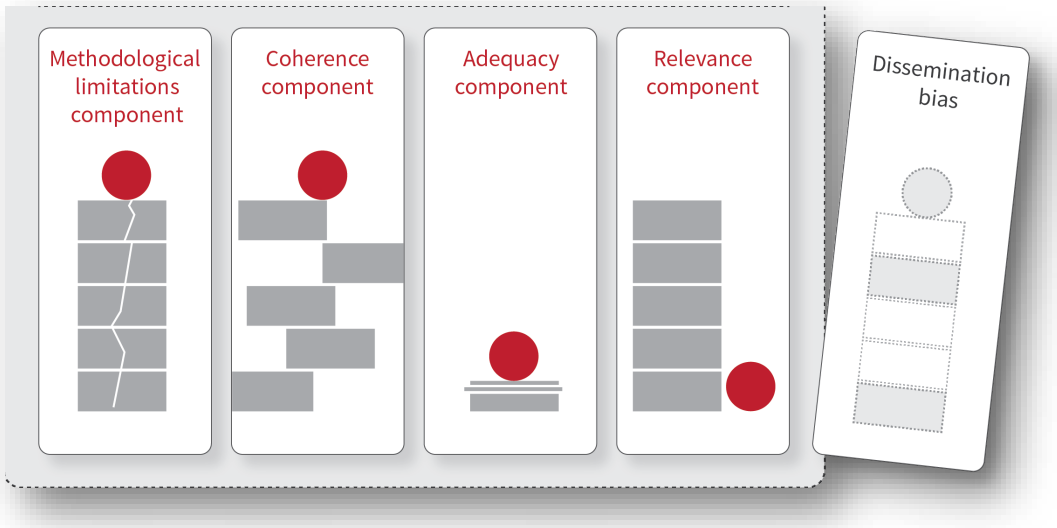


GRADE Evidence-to-Decision Frameworks and qualitative evidence

Gender, equity and human rights	…the extent to which certain groups are likely to benefit more or less than others from the intervention in ways that could be corrected, for instance because of their place of residence, ethnicity, gender or sex and so on …the extent to which the intervention may impact on stakeholders' universal rights as individuals or groups, or lead to discriminatory practices or unjust power relations.
Implementation considerations	What factors, referred to in the evidence above, should national or local decision makers consider when planning to implement an intervention? Could include both barriers and facilitators to implementing an intervention and how these play out across different groups and contexts.



GRADE-CERQual approach



CERQual: Confidence in the Evidence from Reviews of Qualitative Research

- GRADE-CERQual aims to transparently assess and describe how much confidence to place in findings from qualitative evidence syntheses
- CERQual is part of the range of approaches for assessing confidence in evidence developed by the GRADE Working Group
- A key tool for facilitating the use of qualitative evidence in decision making processes

Why did we develop GRADE-CERQual?

- Systematic reviews of qualitative research (also called qualitative evidence syntheses) becoming increasingly common
- Also increasingly being used in guideline or policy development processes
- Users need methods to assess how much confidence to place in findings from these reviews
- Users likely to make these judgements anyway –helpful to provide a systematic and transparent way for doing this

How was GRADE-CERQual developed?

- Researchers with backgrounds in qualitative research and systematic reviews
- Broad consultation with wide group of stakeholders

Needed an approach that:

- Could be applied to typical types of qualitative evidence syntheses
- Was easy to use
- Allowed judgements to be reported transparently
- Allowed the judgements to be understood

The GRADE-CERQual approach aims to:

- Transparently assess and describe how much confidence to place in findings from qualitative evidence syntheses

GRADE-CERQual is not a tool for:

- Assessing how well an individual qualitative study was conducted
- Assessing how well a systematic review of qualitative studies was conducted
- Assessing quantitative studies of quality of care

What do we mean by ’confidence in the evidence’?

The extent to which a review finding is a reasonable representation of the phenomenon of interest

- i.e. the phenomenon of interest is unlikely to be substantially different from the research finding

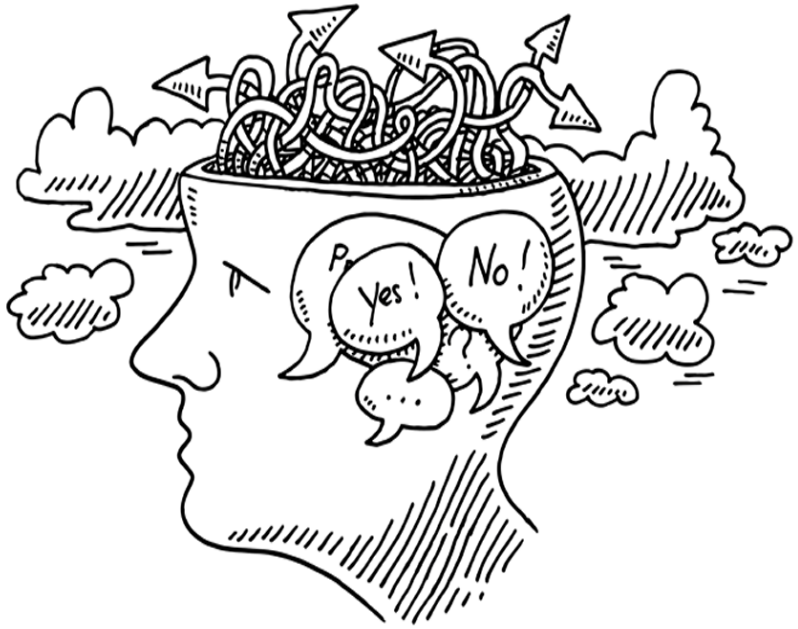
GRADE-CERQual is applied to *individual* synthesis findings

- In the context of a qualitative evidence synthesis, a finding is…:

> …an analytic output that describes a phenomenon or an aspect of a phenomenon

- Findings from qualitative evidence syntheses typically presented as:
  - Themes, categories or theories
  - As both descriptive or more interpretive findings

Why assess confidence in qualitative evidence?



- Users of evidence tend to make judgements implicitly about how trustworthy evidence or information is
- Implicit bias, based on implicit attitudes and stereotypes, may drive these judgements (Greenwald et al. 2006)
- It may be therefore helpful to provide a systematic and transparent way of assessing confidence in evidence

Relationship to GRADE

- GRADE-CERQual is part of the GRADE Working Group

- GRADE-CERQual shares the same aim as the GRADE tool used to assess the certainty of evidence of *effectiveness*

- However, GRADE-CERQual is grounded in the principles of qualitative research

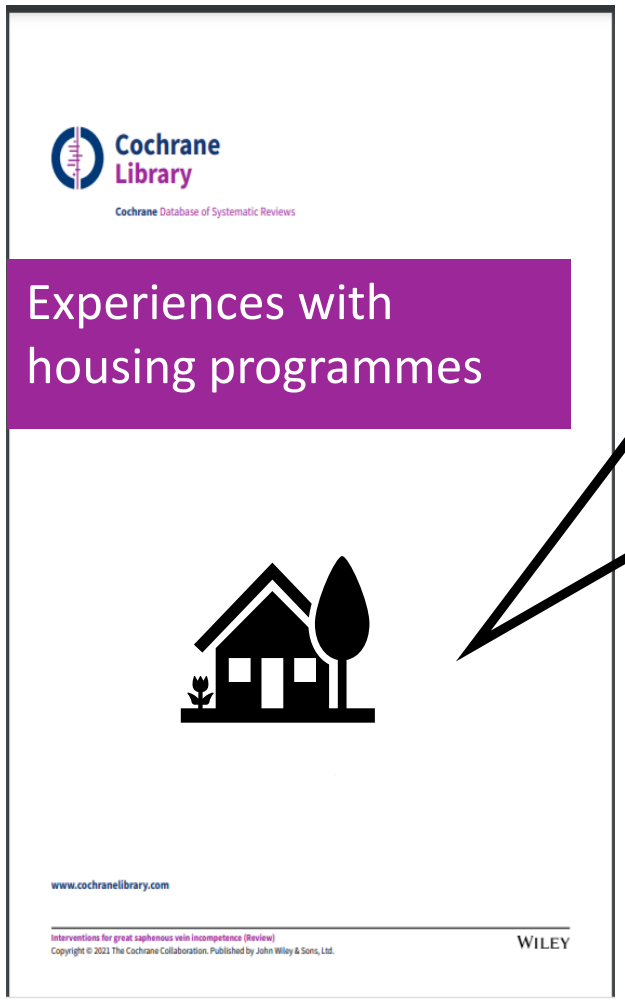
Mapping GRADE and CERQual - commonalities

GRADE criteria	GRADE-CERQual domains
Risk of bias	***Methodological limitations*** of the individual studies contributing to a review finding
Inconsistency	***Coherence*** of the review finding
Indirectness	***Relevance*** to the review question of the individual studies contributing to a review finding
Imprecision	***Adequacy*** of data contributing to a review finding
Publication bias	Dissemination bias

What skills do you need to apply CERQual?

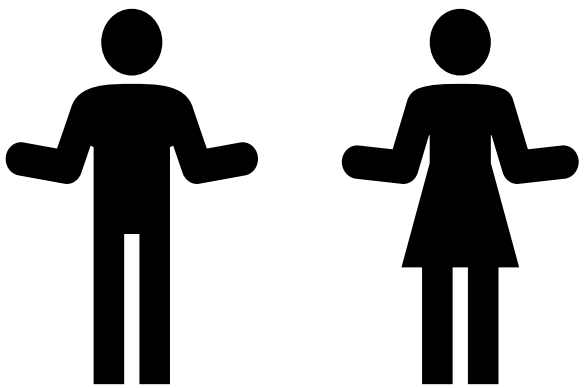
- An understanding of systematic review methodology
- An understanding of the principles of qualitative research

GRADE-CERQual in practice



Homeless people prefer housing programmes that allow them to choose their own housing.

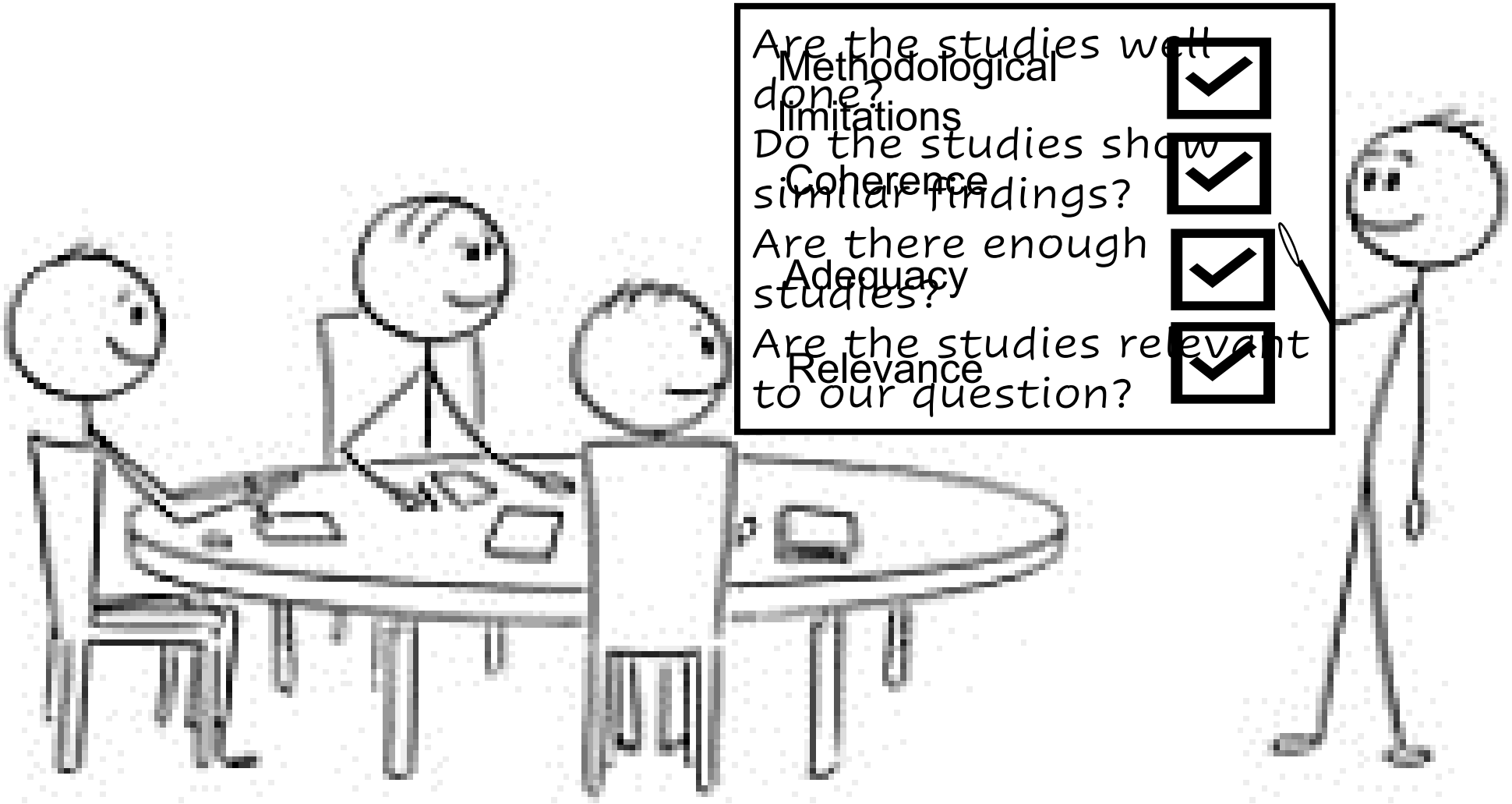
How much should we believe this finding?



Decision makers

GRADE CERQual

GRADE-CERQual in practice

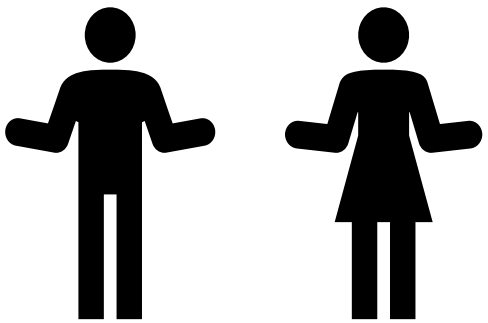


Review team

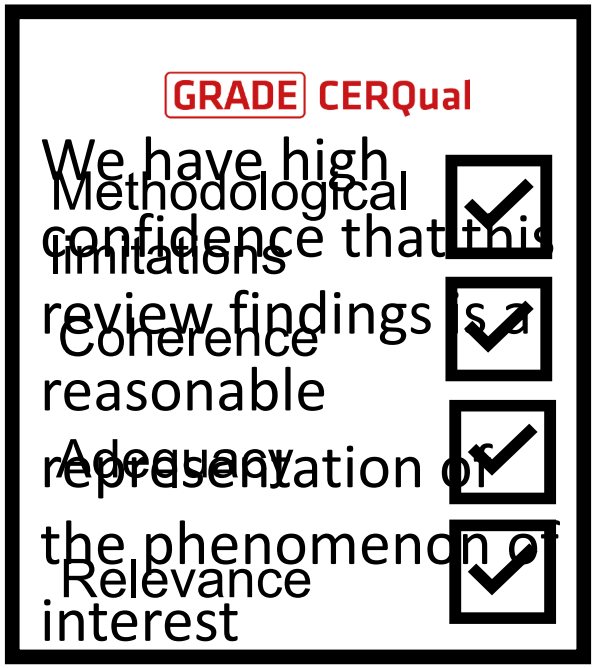
GRADE CERQual

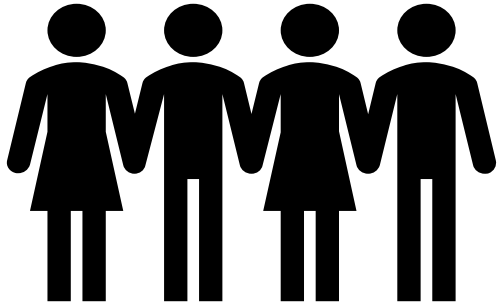
GRADE-CERQual in practice

How much should we believe this finding?



Decision makers

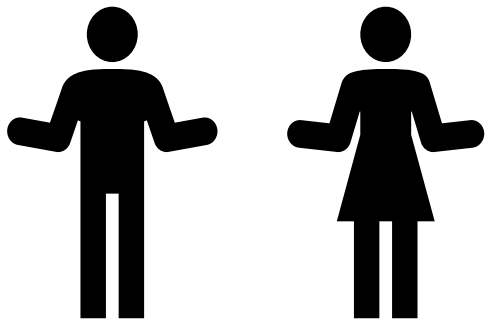




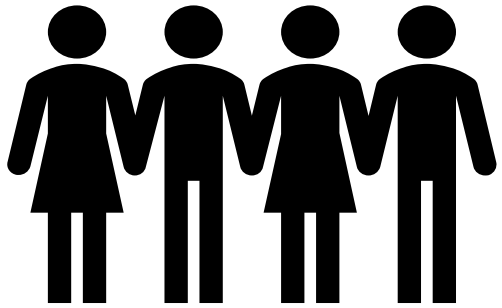
Review team

GRADE-CERQual in practice

Thank you. We now know that we can have high confidence in the finding that *homeless people prefer housing programmes that allow them to choose their own housing*.



Decision makers



Review team

GRADE CERQual

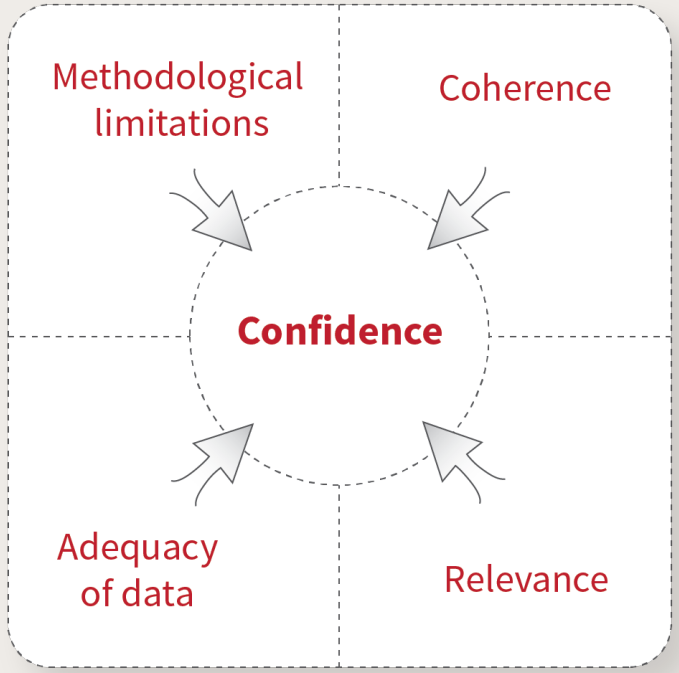
GRADE-CERQual approach (2)

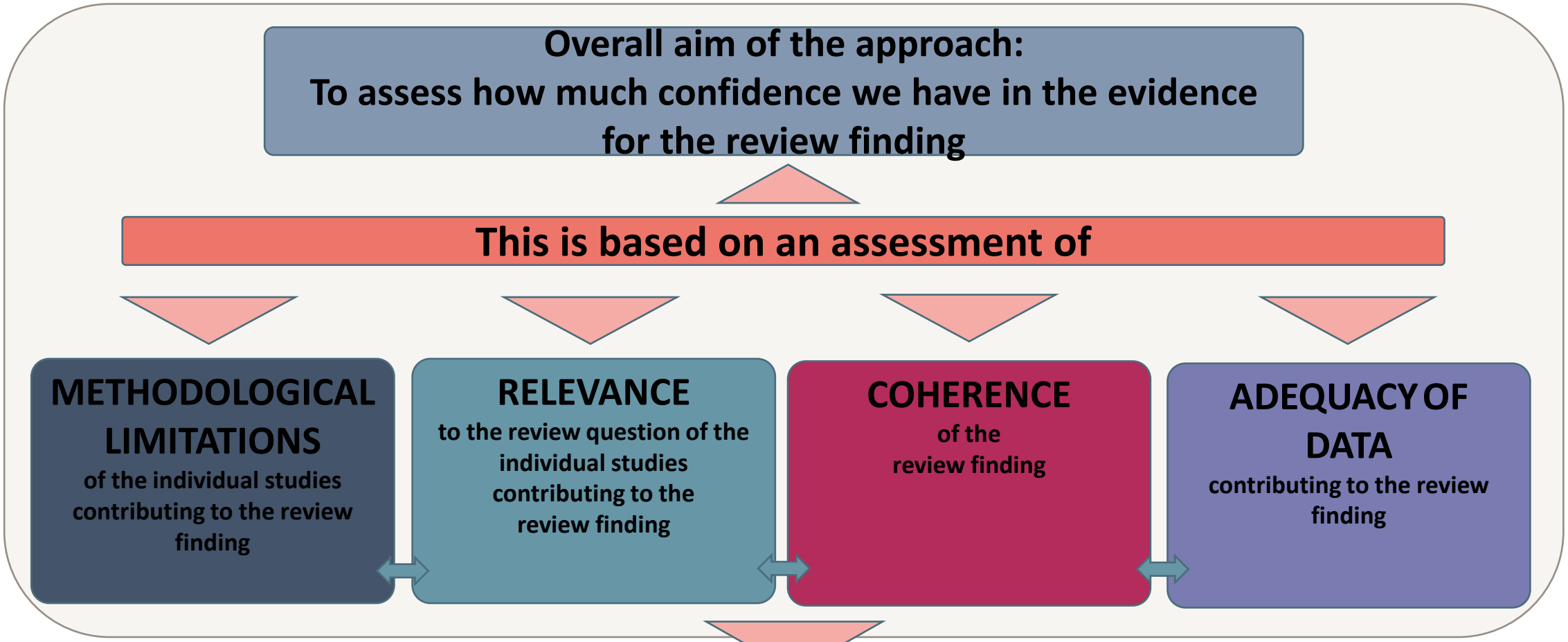
Confidence in the evidence: the extent to which a synthesis finding is a reasonable representation of the phenomenon of interest

- i.e. the phenomenon of interest is unlikely to be substantially different from the research finding

A CERQual assessment of confidence is based on four components

The approach is applied to each analytic output of a synthesis (e.g., a theme or category) that describes a phenomenon or an aspect of a phenomenon





Summary of Qualitative findings table

Objective:			
Perspective:			
Included programmes:			
Review finding	**Overall assessment of confidence in the re finding**	**E-CERQual assessment**	**Studies contributing data to the review finding**
Review finding 1	XXX	XXX	XXX
Review finding 2	XXX	XXX	XXX

Aims of the CERQual Summary of Qualitative Findings (SoQF) table

- To encourage review authors to consider carefully **what constitutes a finding** in the context of their review, and to express these findings clearly
- To provide a **structured summary** of the review findings and the information contributing to the CERQual assessment for each finding
- To help ensure that review author judgements underlying CERQual assessments are **transparent**
- To facilitate the **understanding and use of synthesis findings**, including the uptake of findings into guidelines and other processes, by end users

CERQual *Summary of Qualitative Findings tables*

Objective: To identify, appraise, and synthesise qualitative research evidence on the barriers and facilitators to the implementation of lay health worker programmes for maternal and child health[#]

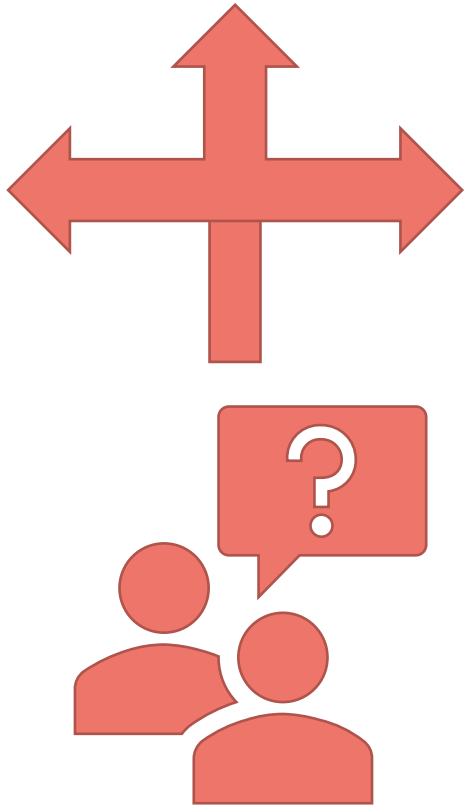
Perspective: Experiences and attitudes of stakeholders about lay health worker programmes in any country

Included programmes: Programmes that were delivered in a primary or community health care setting, that intend to improve maternal or child health, and that had used any type of lay health worker, including community health workers, village health workers, birth attendants, peer counsellors, nutrition workers, and home visitors

Review Finding	CERQual Assessment of Confidence in the Evidence	Explanation of CERQual Assessment	Studies Contributing to the Review Finding
While regular salaries were not part of many programmes, other monetary and nonmonetary incentives, including payment to cover out-of-pocket expenses and "work tools" such as bicycles, uniforms, or identity badges, were greatly appreciated by lay health workers.	Moderate	This finding was graded as moderate confidence because of minor concerns regarding methodological limitations, relevance, coherence, and adequacy.	Studies 2; 5; 11; 12; 22; 29
Some unsalaried lay health workers expressed a strong wish for regular payment.	Low	This finding was graded as low confidence because of moderate concerns regarding relevance and substantial concerns regarding adequacy of data.	Studies 5; 13

GRADE CERQual

Qualitative evidence in decision making



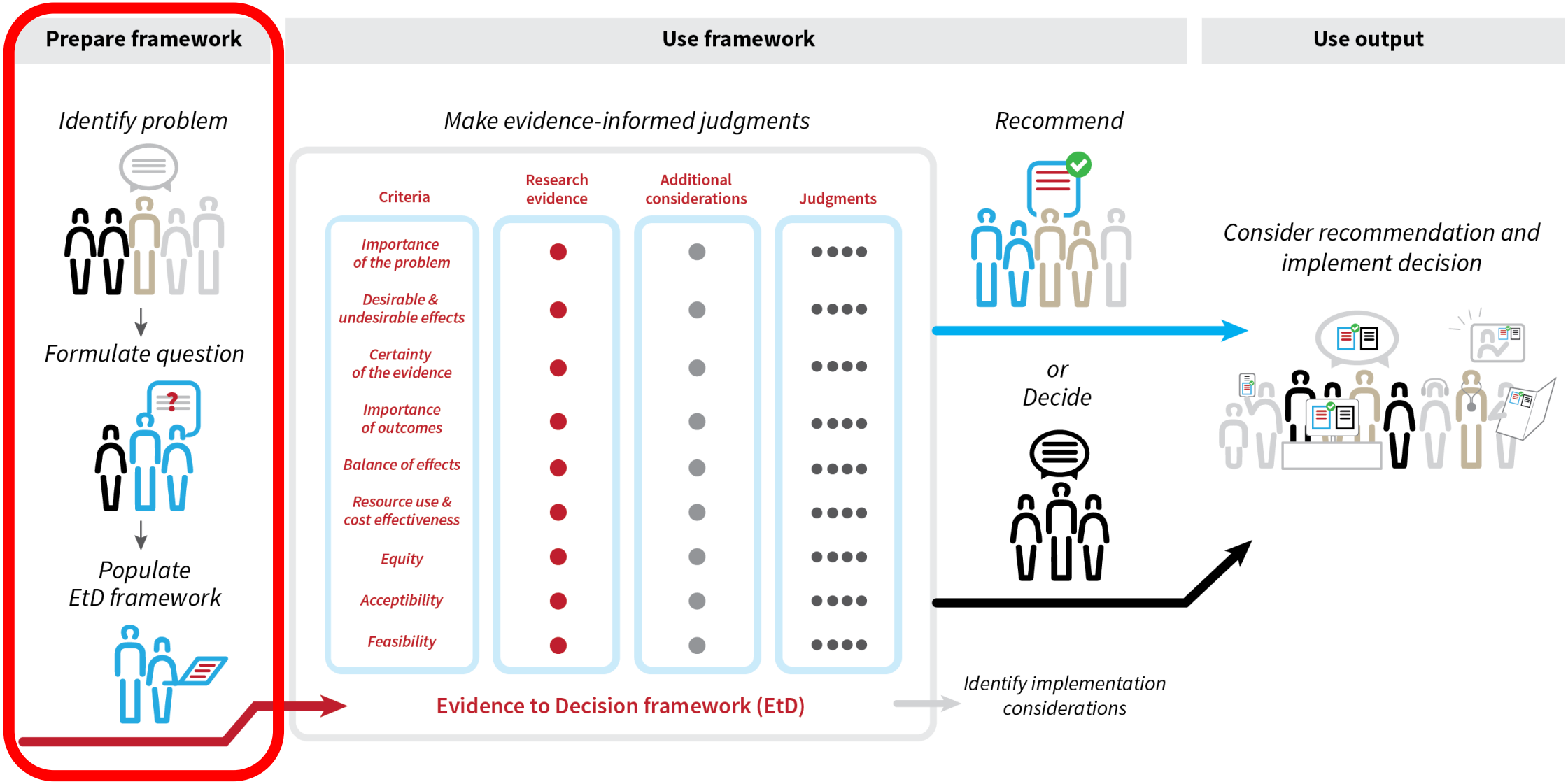
Example

Using qualitative evidence syntheses in decision-making

World Health Organization guidelines on digital health:

Evidence-to-decision framework used to assess different types of evidence

GRADE Evidence-to-decision process



One question: Should the WHO recommend targeted client communication via mobile phone?

Giving targeted information to specific groups of patients and the public by mobile phone, for example:

- health promotion messages
- reminders about health services
- test results

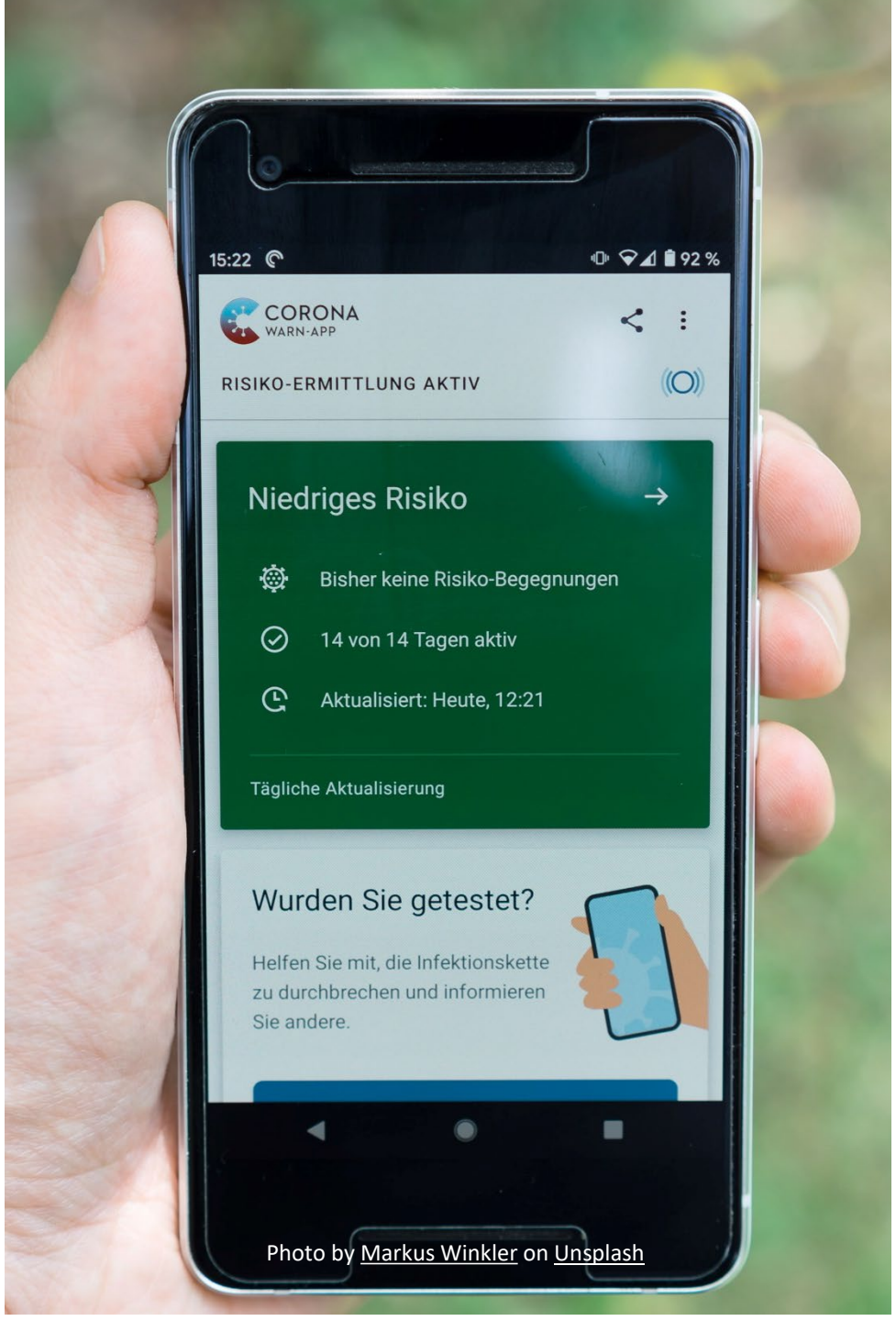


Photo by Markus Winkler on Unsplash

The WHO focused on:

- Adolescents and sexual/reproductive health
- Adults and sexual/reproductive health
- Pregnant women and postpartum women
- Parents of children under 5

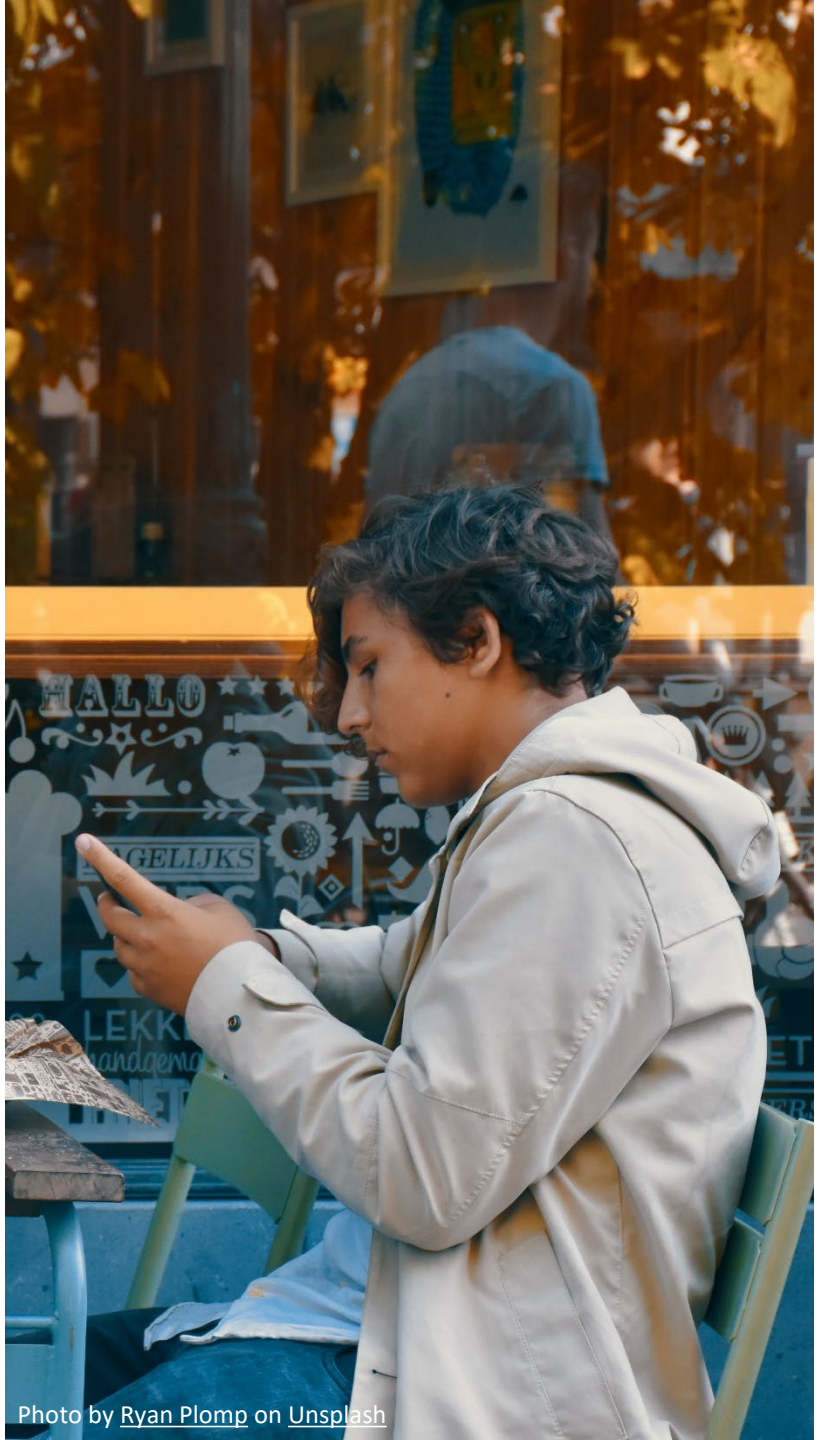
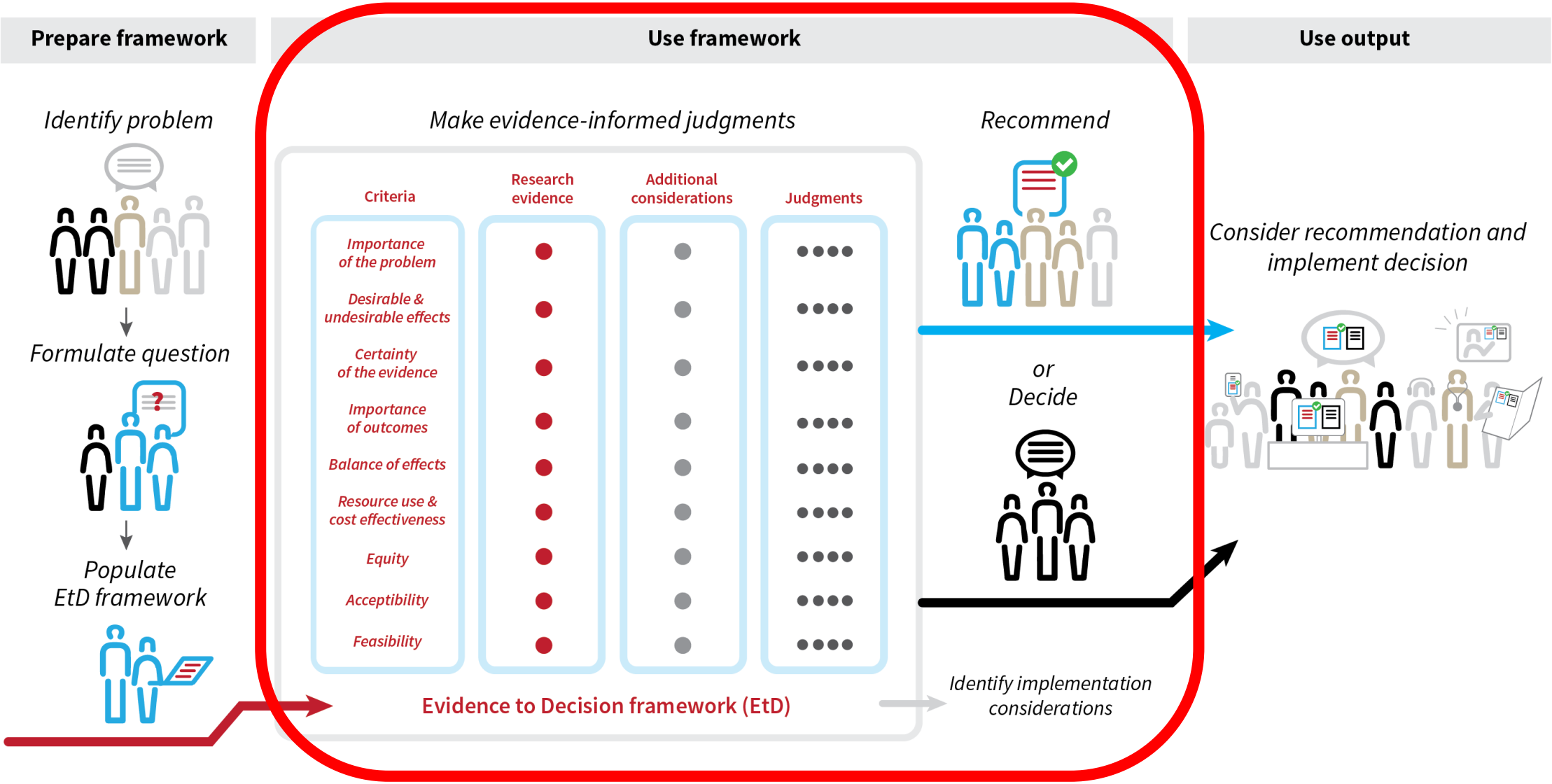


Photo by Ryan Plomp on Unsplash

GRADE Evidence-to-decision process



Targeted client communication via mobile phone: what effect does it have..

…on health service utilisation, health status, health behaviour?

Cochrane Review of effectiveness (Palmer et al 2020):

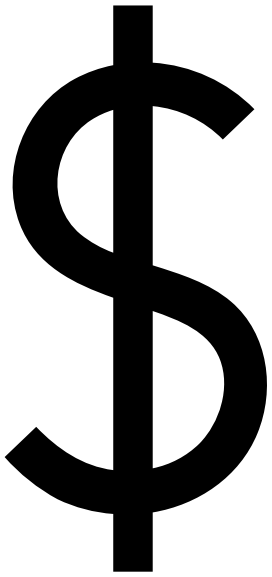
- **Many gaps in the evidence** or low/very low certainty evidence

- **Existing evidence shows mixed effects**: probably some benefits for some outcomes; may make little or no difference to others

Targeted client communication via mobile phone: Cost / resource use

No systematic review prepared. Information based on expert opinion:

- Large start-up costs and large recurring costs



Targeted client communication via mobile phone: do people find it acceptable?

Cochrane Qualitative Evidence Synthesis (Ames et al 2018):

- Many clients **positive** to these services (moderate confidence):
  - Provides them with support and connectedness
  - Feels like someone is interested in their situation and cares about them
  - Gives a sense of direction, reassurance

Targeted client communication via mobile phone: do people find it acceptable?

- …however, clients who are dealing with health conditions that are often stigmatised or very personal (e.g. HIV, family planning and abortion care) **worry that their confidential health information will be disclosed** (high confidence)

Targeted client communication via mobile phone: do people find it acceptable?

People's perceptions and experiences influenced by messages':

- Cost (*high confidence*)
- Content (*moderate confidence*)
- Frequency, timing (*moderate confidence*)
- Sender (*moderate confidence*)
- Length and language (*low confidence*)
- Tone (*low confidence*)

Targeted client communication via mobile phone: is it feasible?

Cochrane Qualitative Evidence Synthesis (Ames 2018):

- Problems in many settings with network connectivity, access to electricity, system integration and device usability (*high confidence*)

Rosenbaum

Targeted client communication via mobile phone: what is the impact on equity?

- Are certain groups of people likely to be **systematically** disadvantaged in relation to these services?

Cochrane Qualitative Evidence Synthesis (Ames 2018):

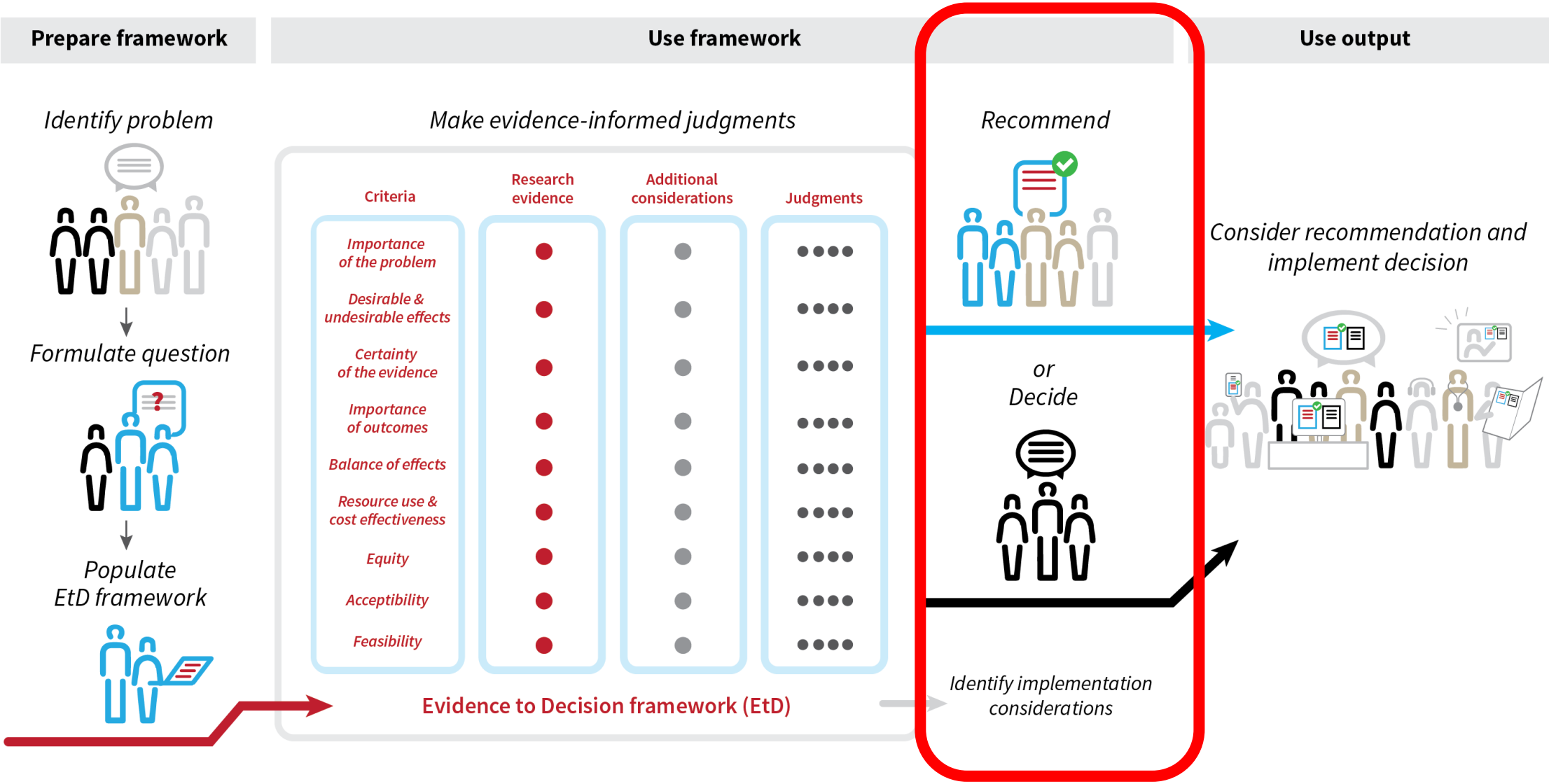
- These services may be particularly helpful to clients **with caring or work responsibilities, clients who live far from health facilities and clients with few funds** (low confidence)

Targeted client communication via mobile phone: what is the impact on equity?

…access to these services may be particularly difficult for:

- People with **poor access to network** or **electricity** (*high confidence*)
- People with **stigmatised health conditions** (concern about confidentiality issues) (*high confidence*)
- People who speak **minority languages** or who have **low literacy skills or low digital literacy skills** (*moderate confidence*)
- People with **poor access to mobile phones**, particularly **women and adolescents**, who may have to share or borrow a phone or who have access to phones controlled by others (*moderate confidence*)

GRADE Evidence-to-decision process



Making the recommendation

The panel assessed the evidence:

- Effectiveness unclear / mixed
- Large costs
- Widespread acceptability, but important conditions / exceptions
- Feasibility challenges
- Equity implications mixed

Making the recommendation

Should policy makers implement targeted client communication via mobile phone for adults, adolescents, pregnant women and parents to communicate about sexual and reproductive health?

- **Recommend?**
- **Recommend with certain conditions?**
- **Recommend against?**

Targeted client communication via mobile phone: what did the panel recommend?

Conditional recommendation: The intervention was recommended under the condition that potential concerns about sensitive content and data confidentiality can be addressed.

Implementation considerations: Implementers should:

- ensure access to network connectivity and electricity
- ensure that the content, format and delivery of information meets the needs of different target groups
- involve stakeholders in the design of the programme
- Etc

GRADE CERQual

Qualitative evidence syntheses (QES) in decision making

WHO example: QES provided evidence about:

- acceptability, feasibility and equity issues
- implementation considerations

QES can also provide evidence about questions, interventions and outcomes that matter to people

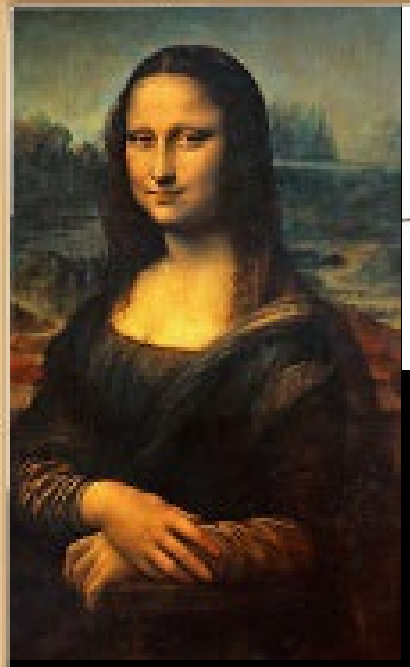
GRADE CERQual

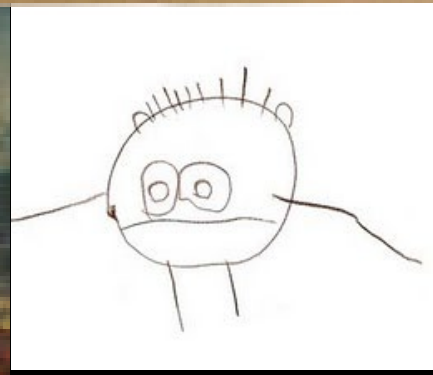
GRADE-CERQual – making assessments of the 4 components

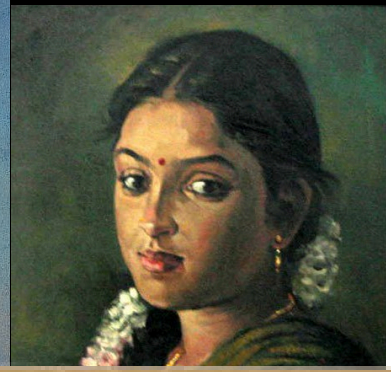

CERQual made easy

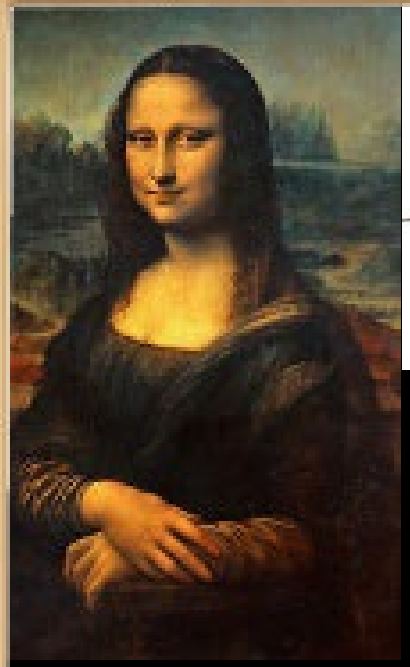
Example (based on Ames 2018):

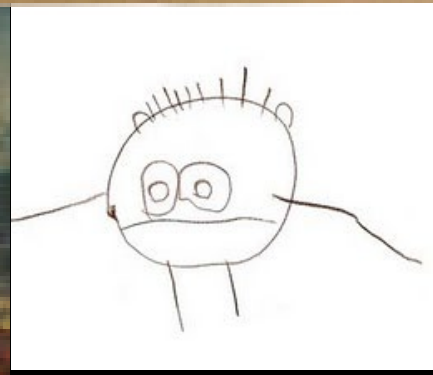
“Many pregnant women are positive to receiving information and support from a peer support group through text messages. They see it as a source of support and connectedness, they feel like someone is interested in their situation and cares about them, and it gives them a sense of direction and reassurance.”

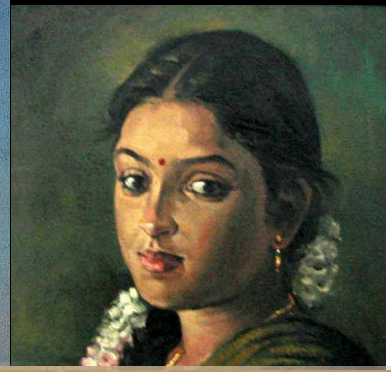










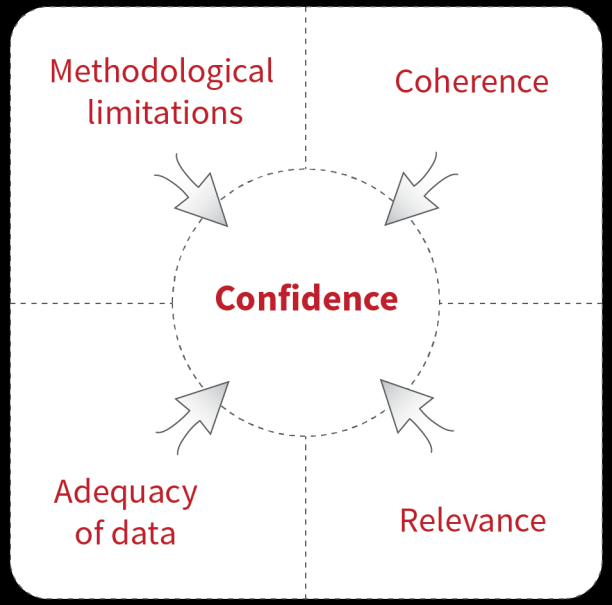


For each CERQual component, you need to identify your concerns and whether these are:

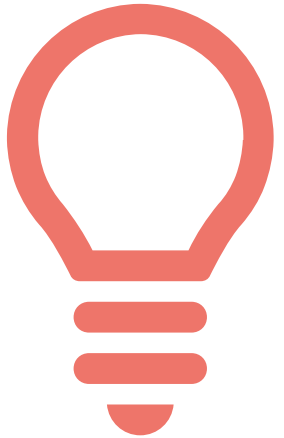
- No or very minor concerns
- Minor concerns
- Moderate concerns
- Serious concerns

After assessing all four components an overall assessment is made, expressed as either:

- High confidence
- Moderate confidence
- Low confidence
- Very low confidence

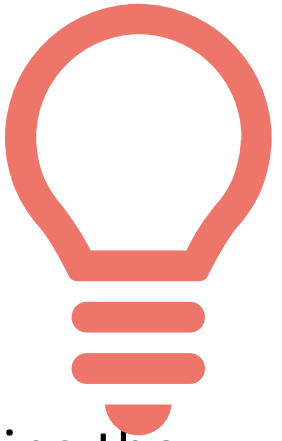


Remember…



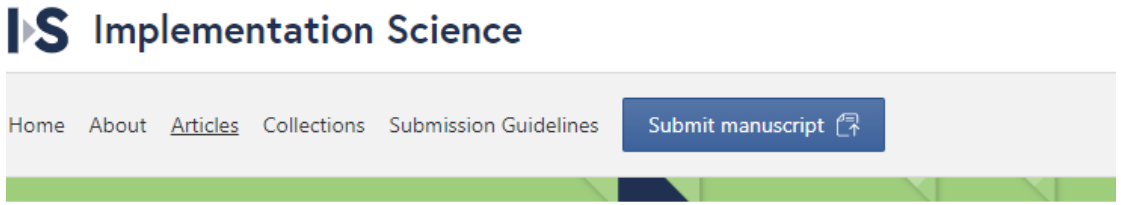
- GRADE-CERQual is an approach for assessing how much confidence can be placed in individual findings from a systematic review of qualitative studies
- The approach is intended to help decision makers use evidence from systematic review of qualitative studies in their decision making processes
- A GRADE-CERQual assessment involves judgements. By being systematic and transparent, we hope to make the thinking behind the judgements explicit to users

Remember…



- GRADE-CERQual assessments are best done by the team who is conducting the systematic review of qualitative studies, as they will be familiar with the data
- The review team should have skills and experience in applying GRADE-CERQual

Additional resources



Introduction | Open Access | Published: 25 January 2018

Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series

Simon Lewin, Andrew Booth, Claire Glenton, Heather Munthe-Kaas, Arash Rashidian, Megan Wainwright, Meghan A. Bohren, Özge Tunçalp, Christopher J. Colvin, Ruth Garside, Benedicte Carlsen, Etienne V. Langlois & Jane Noyes

Implementation Science **13**, Article number: 2 (2018) | Cite this article

44k Accesses | **389** Citations | **38** Altmetric | Metrics

Research | Open Access | Published: 08 August 2019

Qualitative Evidence Synthesis (QES) for Guidelines: Paper 1 – Using qualitative evidence synthesis to inform guideline scope and develop qualitative findings statements

Soo Downe, Kenneth W. Finlayson, Theresa A. Lawrie, Simon A. Lewin, Claire Glenton, Sarah Rosenbaum, María Barreix & Özge Tunçalp

Health Research Policy and Systems **17**, Article number: 76 (2019) | Cite this article

Research | Open Access | Published: 08 August 2019

Qualitative Evidence Synthesis (QES) for Guidelines: Paper 2 – Using qualitative evidence synthesis findings to inform evidence-to-decision frameworks and recommendations

Simon Lewin, Claire Glenton, Theresa A. Lawrie, Soo Downe, Kenneth W. Finlayson, Sarah Rosenbaum, María Barreix & Özge Tunçalp

Health Research Policy and Systems **17**, Article number: 75 (2019) | Cite this article

Research | Open Access | Published: 08 August 2019

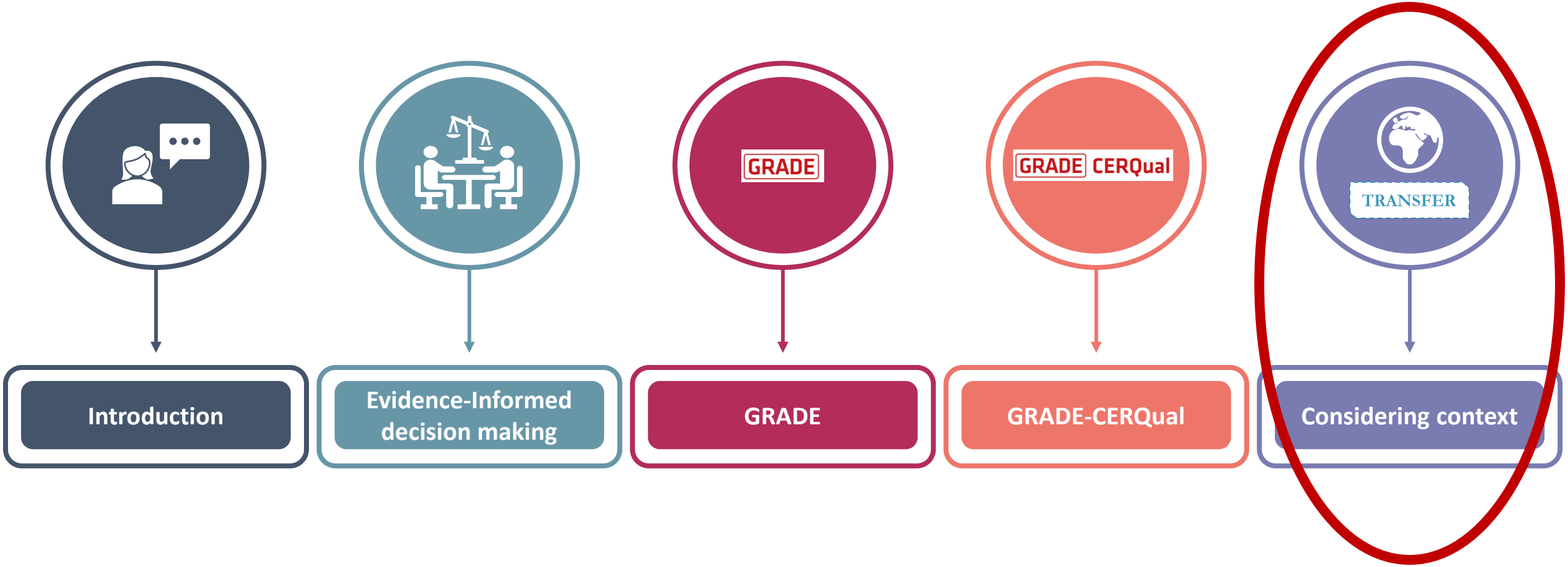
Qualitative Evidence Synthesis (QES) for Guidelines: Paper 3 – Using qualitative evidence syntheses to develop implementation considerations and inform implementation processes

Claire Glenton, Simon Lewin, Theresa A. Lawrie, María Barreix, Soo Downe, Kenneth W. Finlayson, Tigest Tamrat, Sarah Rosenbaum & Özge Tunçalp

Health Research Policy and Systems **17**, Article number: 74 (2019) | Cite this article

World Health Organization hrp

Overview

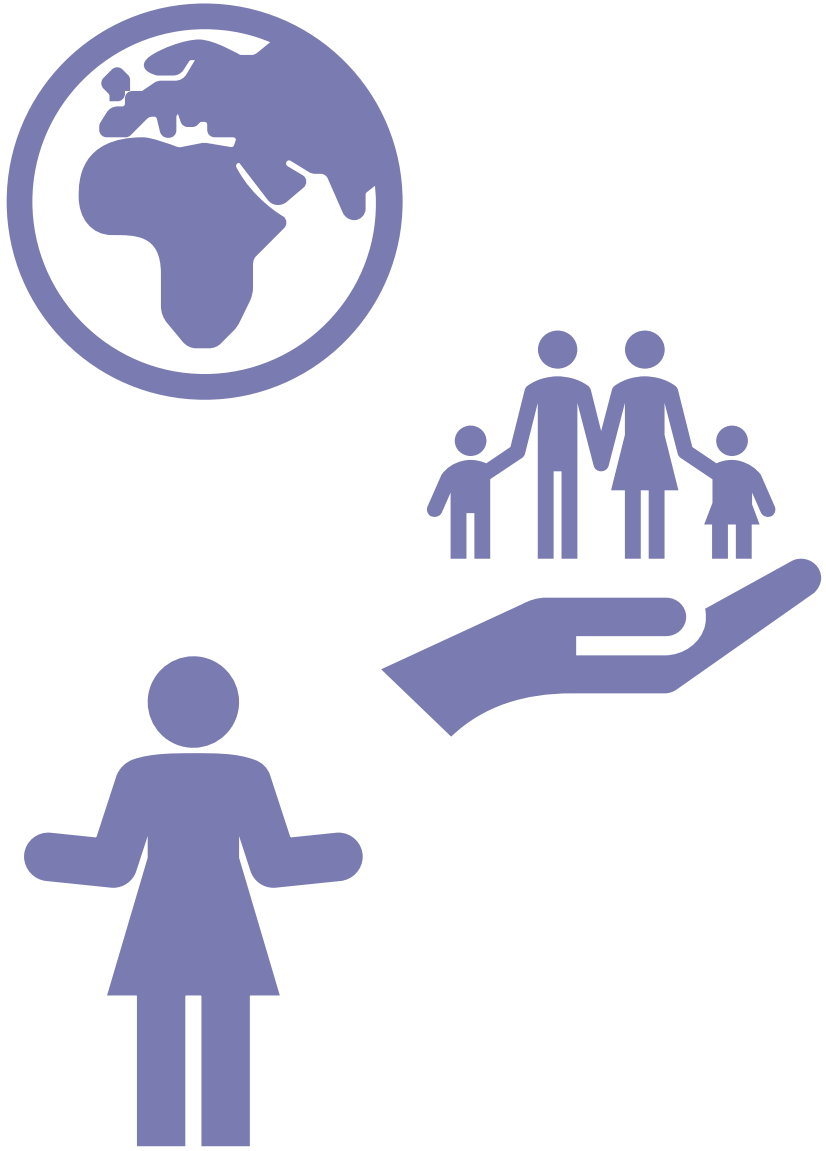

GRADE
GRADE CERQual
TRANSFER
Introduction
Evidence-Informed decision making
GRADE
GRADE-CERQual
Considering context

NIPH
Norwegian Institute of Public Health

Considering context

Why consider context?

- Decision makers may be more likely to use the findings from systematic reviews when the findings were the result of collaborations between decision makers and researchers
- Discussing how the results are relevant for the decision makers, and which factors are important for contextualizing the evidence may improve uptake of evidence



Tricco A, Cardoso R, Thomas S, Motiwala S, Sullivan S, Kealey M, Hemmelgarn B, Ouimet M, Hillmer M, Perrier L *et al*: **Barriers and facilitators to uptake of systematic reviews by policy makers and health care managers: a scoping review**. *Implementation Science* 2016, **11**(4).

Need to consider when making a GRADE assessment

- ✔ 1. Differences in population (applicability)
- ✔ 2. Differences in interventions (applicability)
- ✔ 3. Differences in outcomes measures (surrogate outcomes)
- ✔ 4. Indirect Comparisons (A v B = A v C + B v C)

Need to consider when making a GRADE-CERQual assessment

GRADE CERQual

Relevance

- ✔ Direct relevance
- ✔ Indirect relevance
- ✔ Partial relevance

How to consider context?

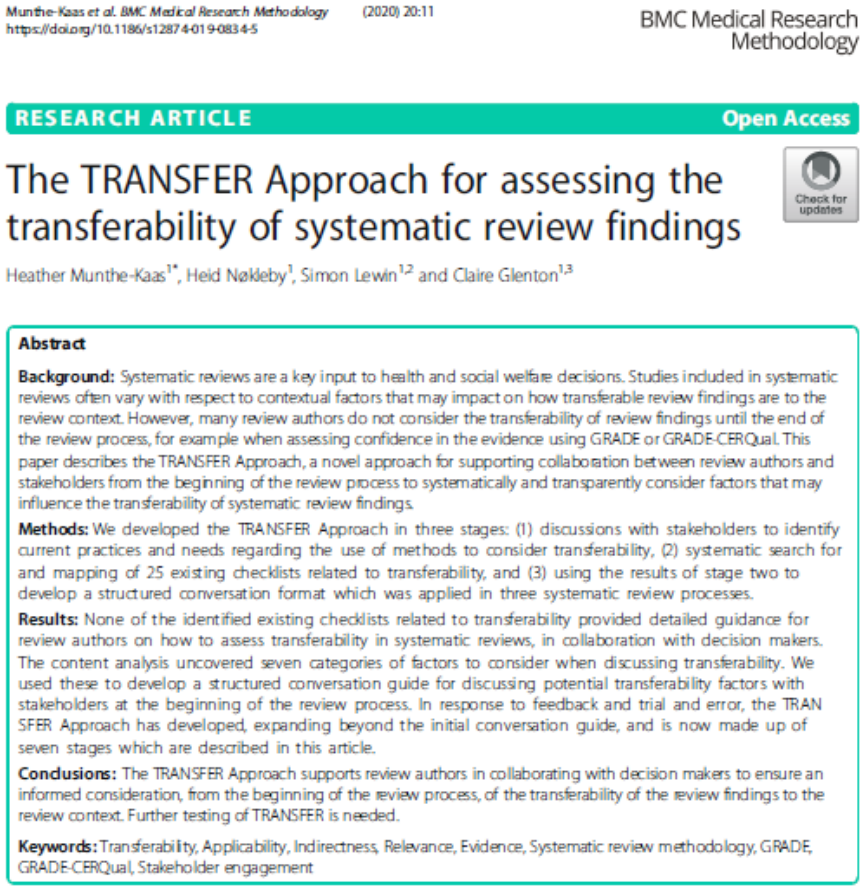
- In collaboration with relevant stakeholders and experts
- Systematically
- Transparently
- From the beginning of the systematic review process

TRANSFER approach

Guidance for review authors on how to:

1. Improve collaboration with decision makers to

2. Systematically and transparently consider and assess transferability of review findings to the review context



Munthe-Kaas et al. BMC Medical Research Methodology (2020) 20:11
https://doi.org/10.1186/s12874-019-0834-5

BMC Medical Research Methodology

RESEARCH ARTICLE Open Access

The TRANSFER Approach for assessing the transferability of systematic review findings

Heather Munthe-Kaas[1*], Heid Nøkleby[1], Simon Lewin[1,2] and Claire Glenton[1,3]

Abstract

Background: Systematic reviews are a key input to health and social welfare decisions. Studies included in systematic reviews often vary with respect to contextual factors that may impact on how transferable review findings are to the review context. However, many review authors do not consider the transferability of review findings until the end of the review process, for example when assessing confidence in the evidence using GRADE or GRADE-CERQual. This paper describes the TRANSFER Approach, a novel approach for supporting collaboration between review authors and stakeholders from the beginning of the review process to systematically and transparently consider factors that may influence the transferability of systematic review findings.

Methods: We developed the TRANSFER Approach in three stages: (1) discussions with stakeholders to identify current practices and needs regarding the use of methods to consider transferability, (2) systematic search for and mapping of 25 existing checklists related to transferability, and (3) using the results of stage two to develop a structured conversation format which was applied in three systematic review processes.

Results: None of the identified existing checklists related to transferability provided detailed guidance for review authors on how to assess transferability in systematic reviews, in collaboration with decision makers. The content analysis uncovered seven categories of factors to consider when discussing transferability. We used these to develop a structured conversation guide for discussing potential transferability factors with stakeholders at the beginning of the review process. In response to feedback and trial and error, the TRANSFER Approach has developed, expanding beyond the initial conversation guide, and is now made up of seven stages which are described in this article.

Conclusions: The TRANSFER Approach supports review authors in collaborating with decision makers to ensure an informed consideration, from the beginning of the review process, of the transferability of the review findings to the review context. Further testing of TRANSFER is needed.

Keywords: Transferability, Applicability, Indirectness, Relevance, Evidence, Systematic review methodology, GRADE, GRADE-CERQual, Stakeholder engagement

Why consider context?

Identify **need** for a systematic review

Collaborate with decision makers to refine PICO

Identify transferability factors

Define characteristics of review context related to transferability factors

Define **inclusion and exclusion criteria**

Develop **search** strategy for relevant databases, grey literat

Screen titles/abstracts and full text for inclusion

Assess **methodological strengths and limitations** of included

Extract relevant data from included studies

Synthesize data: Meta-analyse data, or narrative review

GRADE/-CERQual certainty of evidence (indirectness / relevance)

What factors could influence transferability of the review findings?

- Stakeholders receive the systematic review findings

- They are dissatisfied with review
  - Review question does not exactly what they were interested in
  - Doubtful that the findings will apply to their setting

Why consider context?

What factors could influence transferability of the review findings?

Identify **need** for a systematic review

Collaborate with decision makers to refine PICO

Identify transferability factors

Define characteristics of review context related to transferability factors

Define **inclusion and exclusion criteria**

Develop **search** strategy for relevant databases, grey literature

Screen titles/abstracts and full text for inclusion

Assess **methodological strengths and limitations** of included studies

Extract relevant data from included studies

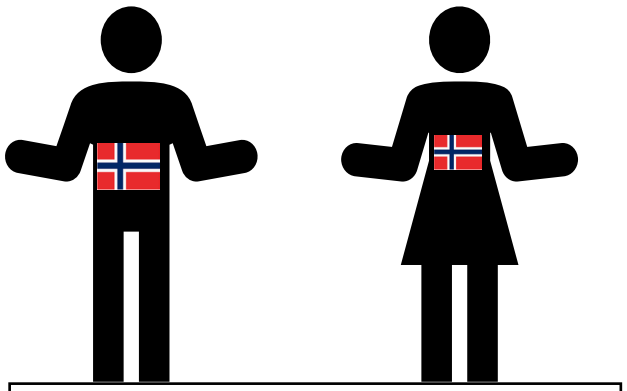
Synthesize data: Meta-analyse data, or narrative review

GRADE/-CERQual certainty of evidence (indirectness / relevance)

TRANSFER in practice

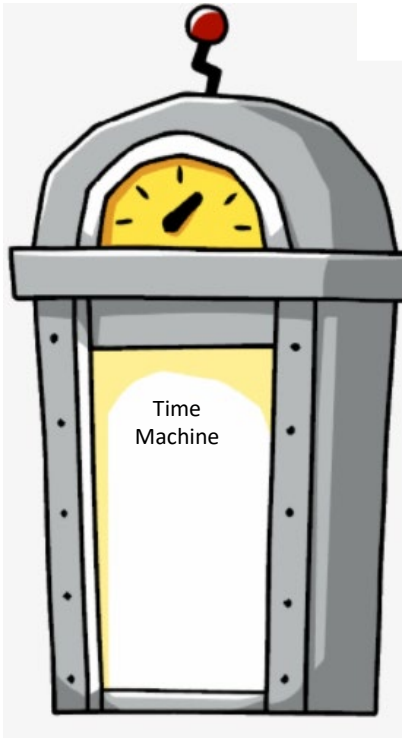
Homeless people prefer housing programmes that allow them to choose their own housing

Does this finding apply to our setting?



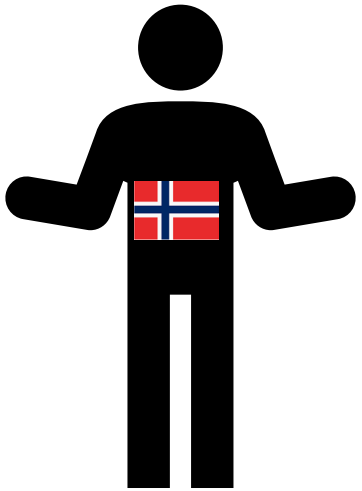
End of review process

Let’s consider context before we conduct the systematic review…



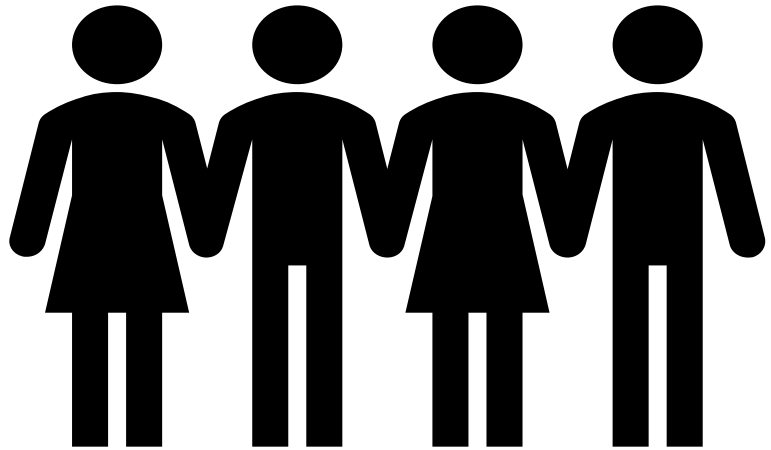
Nicepng.com

We are wondering about housing programmes for homeless people in Norway…



Decision makers

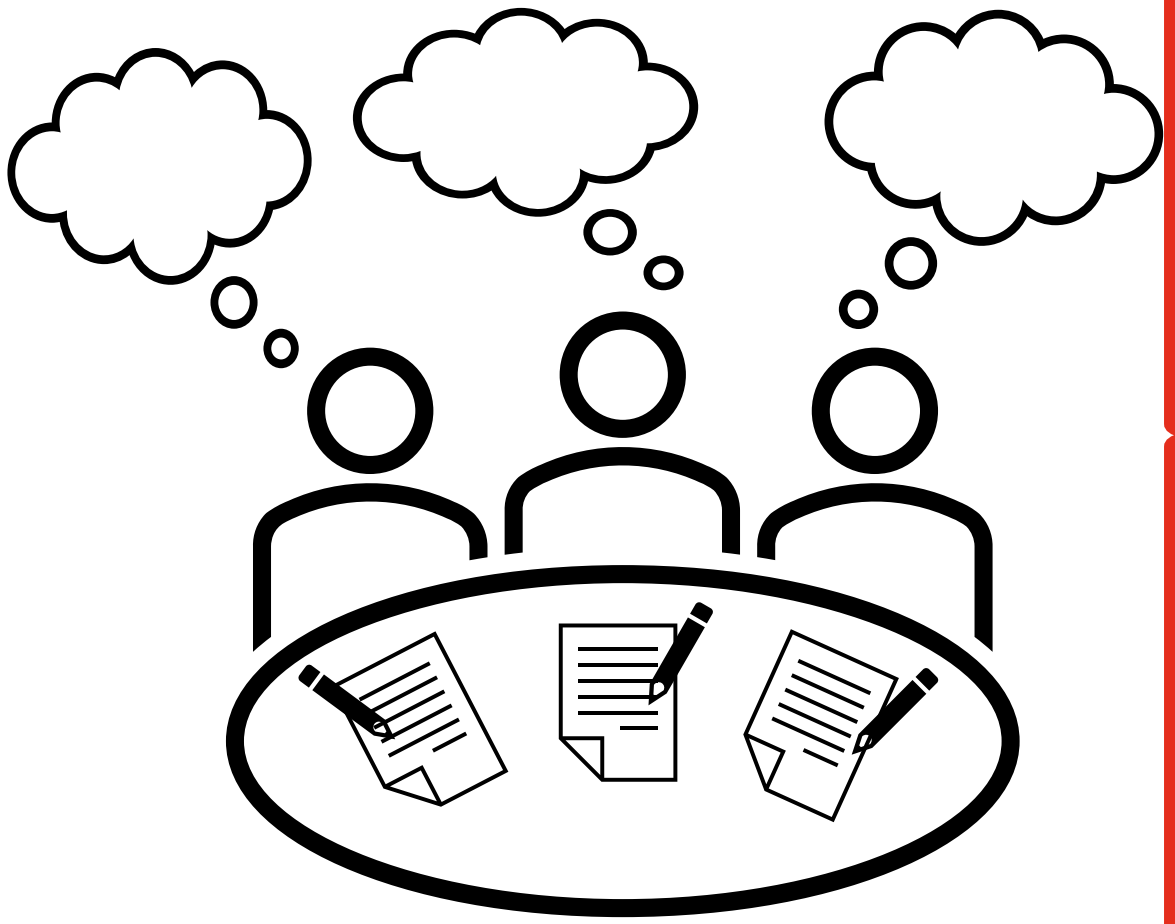
We can help you!



Review team

Beginning of review process

TRANSFER in practice



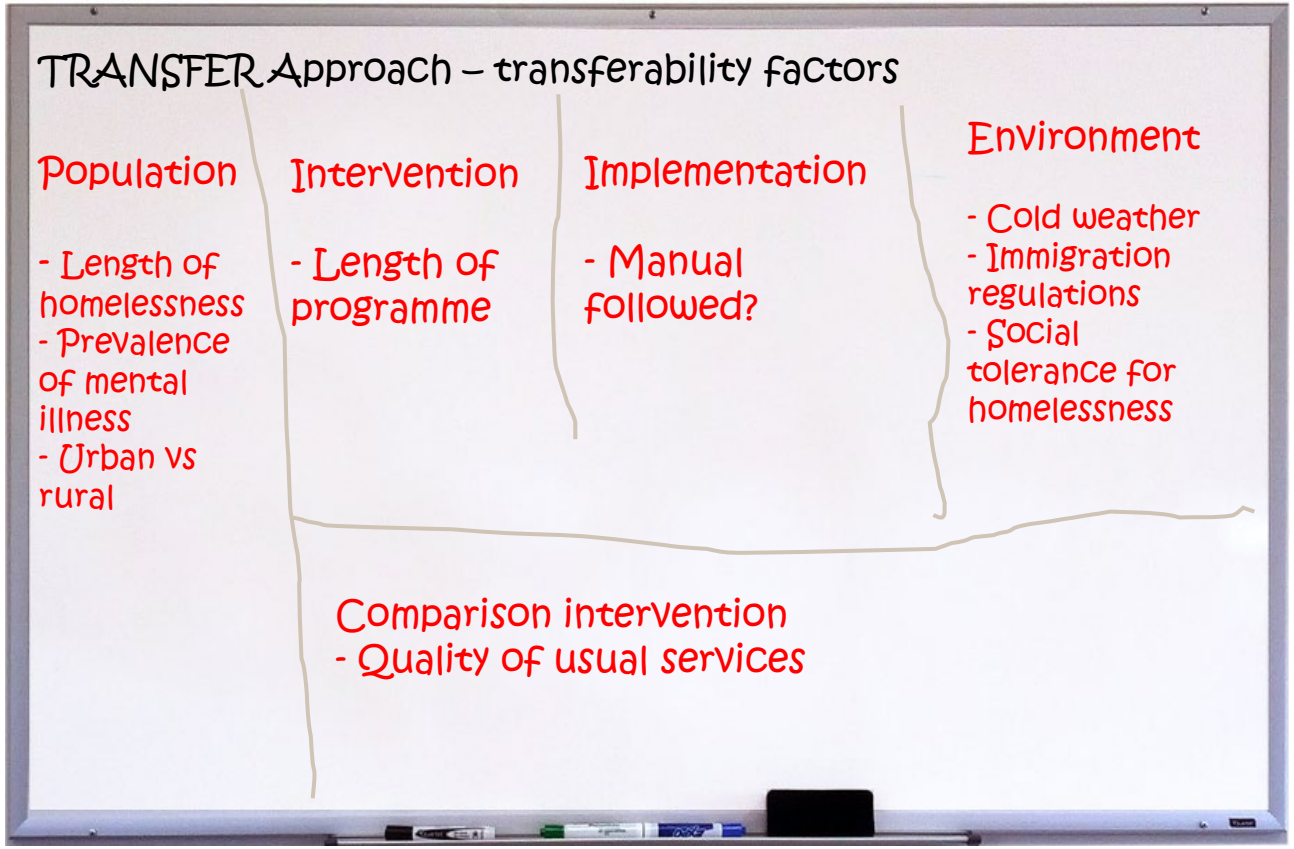
Stakeholders and review authors

TRANSFER Factor	Would you be concerned if data comes from contexts where...	Example	Notes
Environmental context			
Temporal context	... the data was collected at a different point in time?	e.g., studies conducted before 2000	
Geopolitical context	... the geographical, political or economic context is different?	e.g., studies conducted in post-conflict settings, settings where there is famine, high income settings, democratic settings, settings with colder/warmer temperatures, rural or urban settings	
Health or welfare system context	...the health or welfare system is arranged differently?	e.g., free versus fee-based primary health care, comprehensive vs. limited family welfare services	
Local professional/Expert opinion	... local professional/expert opinions are different?	e.g., experts are explicitly in favor or/against the intervention	
Community acceptability	... the local community has a different level/degree of acceptability for the intervention or the condition being addressed by the intervention?	e.g. religious reasons, ethical reasons, other social reasons	
Existence of alternative and/or co-existing interventions	... participants are exposed to alternative or supplemental interventions while participating in the intervention under examination?	e.g. contexts where all parents of small children are provided with free family counselling at the same time as they participate in a study where the intervention is online counselling for families with small children	
Participants			
Participant characteristics	...participants are different with respect to demographic characteristics, level of education, etc.?	e.g., studies on participants older/younger than those in your context, contexts with a different gender ratio,	
Participant compliance	...participants are different with respect to how well they follow instructions?	e.g., studies on pedestrian interventions to improve traffic safety in contexts where people are more/less likely to follow traffic rules	
Availability of personal support for participants	...participants have different access to personal support networks?	e.g., contexts where families live close by vs. individualistic cultures	
Characteristics of illness / condition and comorbidities	...participants' condition or illness and comorbidities are different?	e.g., studies on premenstrual symptoms from Asian cultures versus western cultures where research suggests a difference in how women experience these conditions	
Participant acceptability and preferences	...participants level of acceptability and/or preferences regarding interventions/treatment, etc. are different?	e.g., studies of colon cancer screening interventions for men from contexts where they prefer to be called into/make their own annual appointments	
Participant need for / access to information	...participants have a different need for/access to/expectation of information?	e.g., studies from contexts where participants have a greater expectation of receiving comprehensive and detailed information regarding their treatment/intervention	
Intervention			
Details related to the intervention	... the intervention components/stages/phases/elements are routinely/consistently differ from your context?	These issues may be covered in while defining the review question and covered under inclusion/exclusion criteria in some cases.	
	...the intervention has a different duration, frequency, intensity?	These issues may be covered in while defining the review question and covered under inclusion/exclusion criteria in some cases.	
	...the intervention is delivered in a different setting?	These issues may be covered in while defining the review question and covered under inclusion/exclusion criteria in some cases.	
	...the availability and/or characteristics of materials/manuals for delivering the intervention is different?	These issues may be covered in while defining the review question and covered under inclusion/exclusion criteria in some cases.	
	...the intervention is delivered differently than it would be in a "real life setting"?	e.g. laboratory/efficacy studies	
	...the intervention has been tailored?	These issues may be covered in while defining the review question and covered under inclusion/exclusion criteria in some cases.	
	...the intervention is not delivered according to how it should be (i.e. implantation fidelity)?	e.g., the study authors describe clear deviations from how the intervention is intended to be developed (checklists such as TIDier could be helpful here)	
Category / status of the intervention	... the intervention is categorized differently?	e.g. policy, practice, programme, guideline	
Implementation of the intervention	...the intervention is delivered by service providers who differ from those in your setting?	e.g., number of service providers, characteristics of service providers, such as training or skill level or type/status of service providers' position, their compliance with implementation directions, any other factors that may influence their motivation to implement the intervention, such as religious beliefs, cultural background or support from leadership/colleagues?	
	...the intervention is implemented by an organization that differs from those that would be expected to implement the intervention in your setting?	e.g., type of organization, size/structure, culture, policies, service and financing systems, interagency working relationships, available/allocated resources, communication/endorsement of intervention, evolution/sustainability of intervention	
Comparison intervention			
	...the quality or comprehensiveness of the comparison intervention is different?	This is likely to be important for the transferability of most interventions	
	..."usual services" is different with respect to quality, comprehensiveness or content?	This is likely to be important for the transferability of most interventions	
Outcomes			
	...the way an outcome is defined or measured is different, including length and intensity of follow-up?	e.g., culturally different scales to measure quality of life, long-term versus short-term follow-up	
	...the way an outcome is prioritized (by clients/patients) is different?	e.g., patient-important outcomes	

TRANSFER conversation guide

TRANSFER factor	Would you be concerned if data come from contexts where…	Example
Temporal context	…the data was collected at a different point in time?	e.g., studies conducted before 2000
Geopolitical context	…the geographical, political or economic context is different?	e.g., studies conducted in post-conflict settings, settings where there is

TRANSFER in practice



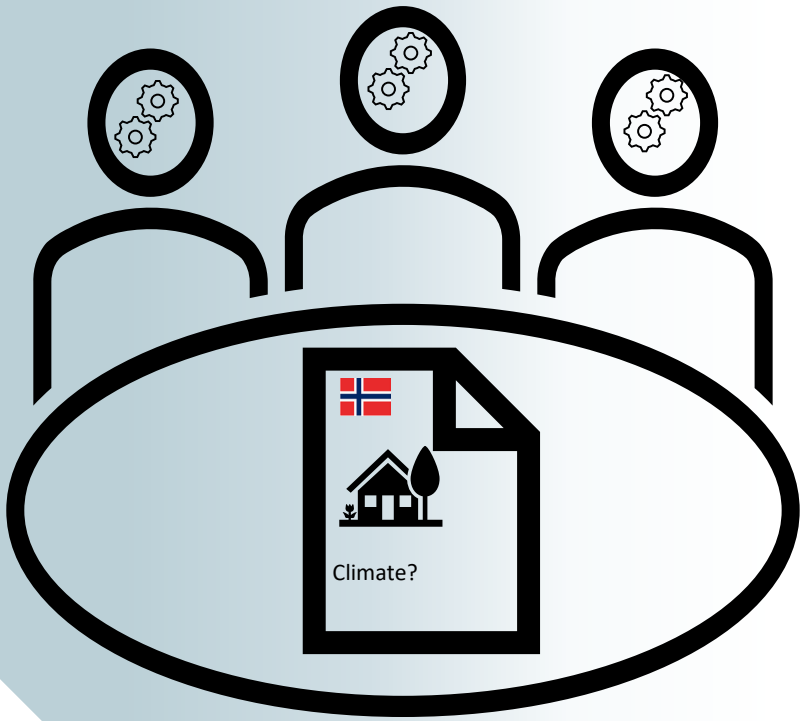
Stakeholders and review authors

TRANSFER in practice

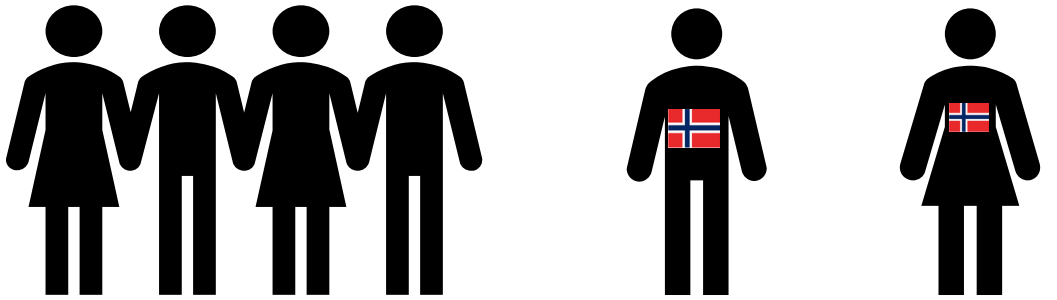
Stakeholders and review authors

TRANSFER in practice

Review team



We don’t have serious concerns about the transferability of the finding to the Norwegian context. Data comes from studies that were done in cold countries.



Review team

Decision makers

Supporting a GRADE assessment of indirectness?

Table 8.1.2. Comparison 1.A.2 – GRADE Evidence Profile for high intensity case management compared to low intensity case management

Author(s): Heather Munthe-Kaas, Rigmor Berg
Date: 11.11.2016
Question: High intensity case management compared to low intensity case management for improving housing stability and reducing homelessness
Setting: USA
Bibliography: Essock 2006; Drake 1998; Morse 1997

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	high intensity case management	low intensity case management	Relative (95% CI)	Absolute (95% CI)		
Mean number of days spent in stable housing (follow up: 36 months; assessed with: self-report)												
2	randomised trials	serious [1]	Serious[3]	not serious	not serious [2]	none	204	197	-	SMD **0.1 SD higher** (0.1 lower to 0.29 higher)	⨁⨁◯◯ LOW	

intervention group reported more days

Review finding: Housing programmes lead to fewer days spent homeless compared to usual services			
TRANSFER factors	**Assessment**	**Explanation**	**Supporting studies**
Length of homelessness of participants	No concerns	The studies represented a range of participants with length of homelessness at baseline rangeing from 1 month to more than 4 years. All of the studies showed the same direction of effect.	1-10
Climate	Minor concerns	The studies only partially represented the review context (cold climates). We are unsure if the finding is tranfserable to settings with warm or temperate climates.	1-10
Overall assessment	Minor concerns	There are no substantial differences between the included studies and the review context with respect to length of homelessness. However, the review finding is only based on evidence from cold climate settings, and we do not have any evidence available regarding how the intervention may work in warm settings.	1-10

Supporting a GRADE assessment of indirectness?

Housing programmes compared to usual services for reducing homelessness and improving housing stability in Norway

Patient or population: Adults who are homeless
Setting: USA, Canada, Denmark, Australia
Intervention: Housing programmes
Comparison: Usual services

Quality assessment							Summary of findings				
							Study event rates (%)			Anticipated absolute effects	
Nº of participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Usual services	Housing First	Relative effect (95% CI)	Risk with Usual services	Risk difference with Housing First
Number of days spent in stable housing (12 months)											
3027 (10 RCTs)	Not serious	Not serious[1]	Serious[2]	Not serious	none	⨁⨁⨁◯ MODERATE	1502	1525	-	-	SMD 20.24 **days more** (15.11 to 25.37)

CI: Confidence interval; **MD:** Mean difference

1. Large inconsistency, however a priori hypotheses related to length of homelessness and quality of usual services can explain heterogeneity.
2. No concerns regarding differences between studies and review context with respect to length of homelessness. Minor concerns regarding differences between studies and review context related to climate. Only cold climates represented in the data.

Supporting a GRADE-CERQual assessment of relevance?

Finding	Summary of review finding	Studies contributing to the review finding	Methodological limitations	Coherence	Relevance	Adequacy	CERQual assessment (confidence in the findings)	Explanation of CERQual assessment
Factors affecting experience of being homeless								
1	Participants who receive housing programmes experience less stress and are more positive to long term opportunities	Study a, study b, study c, study d, study e, study f, study g, study h, study i, study j	Minor concerns regarding methodological limitations due to issues with reflexivity	Minor concerns regarding coherence	Minor concerns regarding relevance	No concerns regarding adequacy	Moderate confidence	Due to minor concerns regarding methodological limitations, coherence, and relevancy.

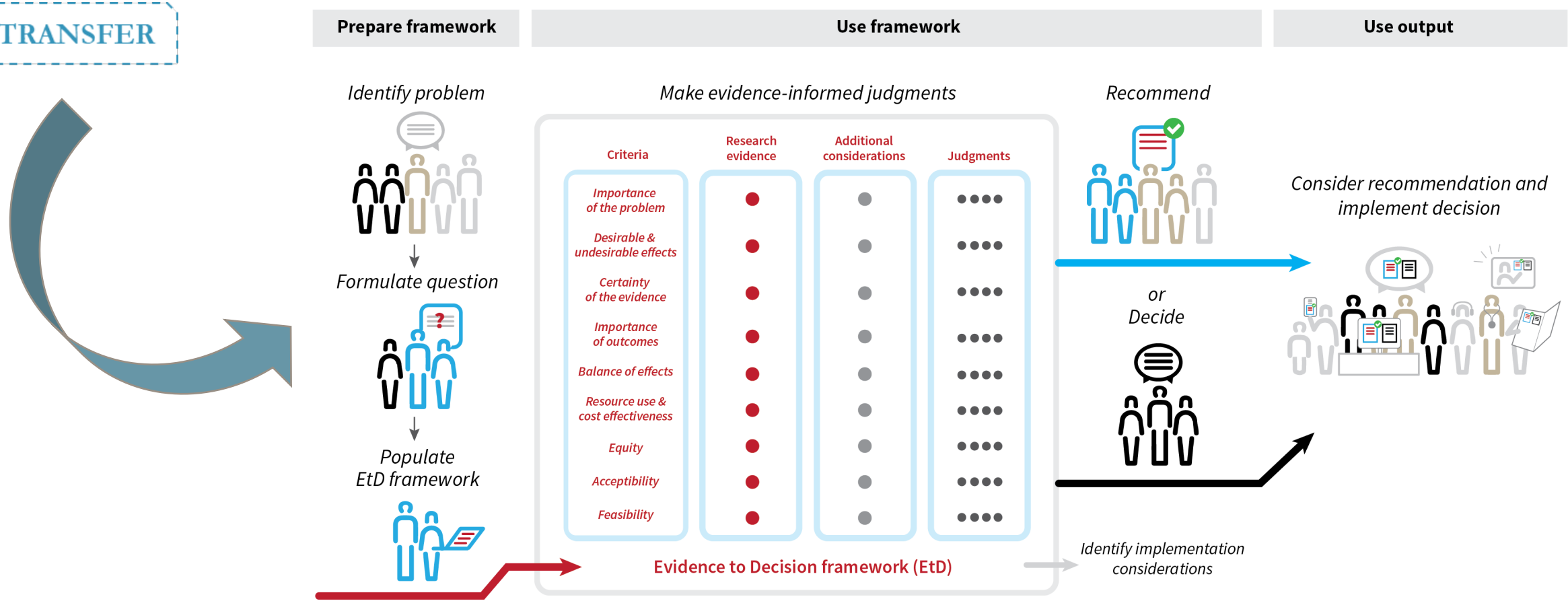
Review finding: Participants who receive housing programmes experience less stress and are more positive to long term opportunities			
TRANSFER factors	**Assessment**	**Explanation**	**Supporting studies**
x	No concerns		1-10
Y	No concerns		1-10
Z	Minor concerns		1-10
Overall assessment	Minor concerns		1-10

GRADE

GRADE CERQual

TRANSFER

GRADE Evidence-to-decision process



In summary…

GRADE

GRADE CERQual

TRANSFER

- Systematic reviews of social interventions can be challenging
- We need to assess and communicate our certainty and confidence in findings from systematic reviews
- We need to consider context in systematic reviews
- We need to work with stakeholders throughout the systematic review to ensure a systematic, transparent process with useful and relevant results

Thanks!

NIPH
Norwegian Institute of Public Health

Thanks to:

Sarah Rosenbaum (NIPH) and DECIDE for allowing me to adapt some of her slides

Collaborators in the GRADE-CERQual Project Group and the TRANSFER working group

Heather.munthe-kaas@fhi.no