Updated GRADE guidance for imprecision rating using the minimally contextualized approach

Linan Zeng, Gordon Guyatt

September 13, 2022 Cochrane Training

GRADE prior guideline on imprecision rating

ELSEVIER

Journal of Clinical Epidemiology 64 (2011) 1283–1293

Journal of Clinical Epidemiology

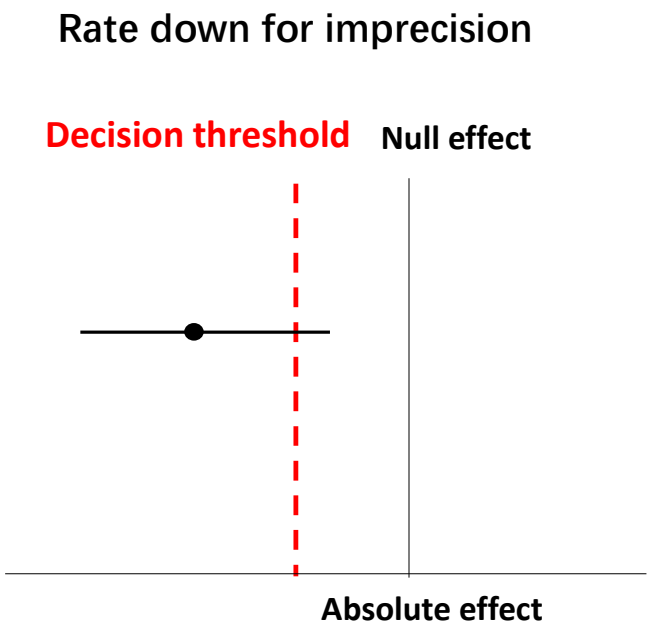
GRADE guidelines 6. Rating the quality of evidence—imprecision

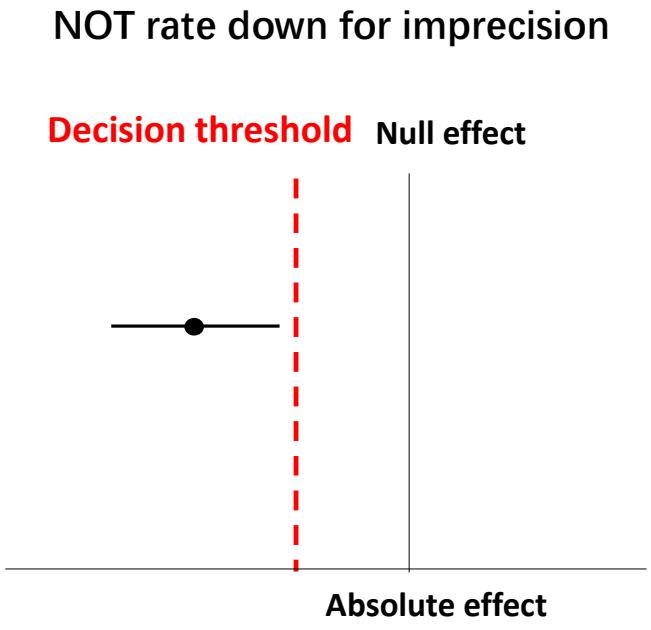
Gordon H. Guyatt[a,b,*], Andrew D. Oxman[c], Regina Kunz[d,e], Jan Brozek[a], Pablo Alonso-Coello[f], David Rind[g], PJ Devereaux[a], Victor M. Montori[h], Bo Freyschuss[i], Gunn Vist[c], Roman Jaeschke[b], John W. Williams Jr.[j], Mohammad Hassan Murad[h], David Sinclair[k], Yngve Falck-Ytter[l], Joerg Meerpohl[m,n], Craig Whittington[o], Kristian Thorlund[a], Jeff Andrews[p], Holger J. Schünemann[a,b]

GRADE prior guideline on imprecision rating

Guidelines

- Primary criterion: whether the CI crosses the decision threshold





GRADE prior guideline on imprecision rating

Systematic reviews

- Primary criterion: OIS approach
- When the effect is large and sample size is modest, check if the optimal information size (OIS) is met.

  If not, rate down

GRADE re-clarification of the construct of certainty of evidence

Journal of Clinical Epidemiology 87 (2017) 4–13

GRADE UPDATE OF PAPERS

The GRADE Working Group clarifies the construct of certainty of evidence

Monica Hultcrantz[a,b,*], David Rind[c,d], Elie A. Akl[e,f], Shaun Treweek[g], Reem A. Mustafa[e,h], Alfonso Iorio[e,i], Brian S. Alper[j,k], Joerg J. Meerpohl[l,m], M Hassan Murad[n], Mohammed T. Ansari[o], Srinivasa Vittal Katikireddi[p], Pernilla Östlund[a,q], Sofia Tranæus[a,q,r], Robin Christensen[s], Gerald Gartlehner[t,u], Jan Brozek[e,i], Ariel Izcovich[v], Holger Schünemann[e,i], Gordon Guyatt[e,i]

Journal of Clinical Epidemiology 137 (2021) 163–175

ORIGINAL ARTICLE

GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings

Linan Zeng[a,b,*], Romina Brignardello-Petersen[b], Monica Hultcrantz[c], Reed A.C. Siemieniuk[b], Nancy Santesso[b], Gregory Traversy[d], Ariel Izcovich[e], Behnam Sadeghirad[b,f], Paul E. Alexander[b], Tahira Devji[b], Bram Rochwerg[b,g], Mohammad H. Murad[h], Rebecca Morgan[b], Robin Christensen[i,j], Holger J. Schünemann[b,g], Gordon H. Guyatt[b,g]

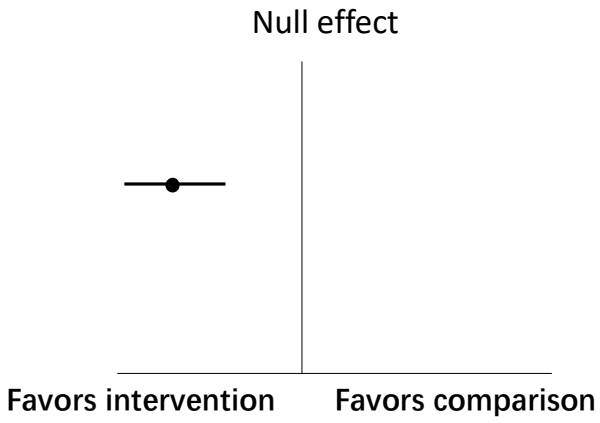
In either guidelines or systematic reviews, when we rate the certainty of evidence, we are assessing our confidence where the point effects lies relative to particular threshold(s) of interest.

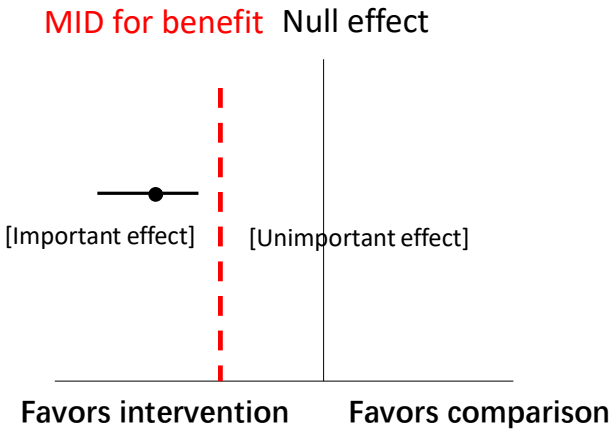
Additional insights regarding imprecision rating

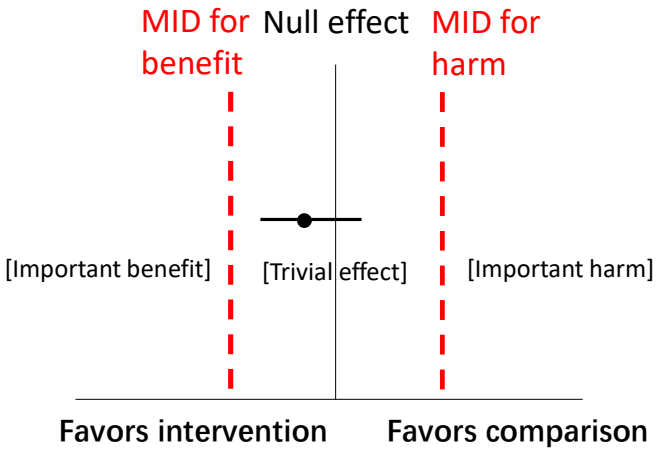
1) In systematic reviews, we are much more likely to use the approach that relies on thresholds and CIs (hereafter that "CI approach") than optimal information size (OIS) to judge imprecision.
2) We should consider rating down more than one level when the CI appreciably crosses the threshold(s) of interest.

Threshold of interest, target of certainty of evidence rating in minimally contextualized approach

	(a)	(b)	(c)
Threshold of interest	Null effect	Minimally important difference (MID)	Range of trivial or no effect
Target of certainty rating	An effect is truly present or not.	An important effect is truly present or not.	A trivial or no effect is truly present or not.







Example 1

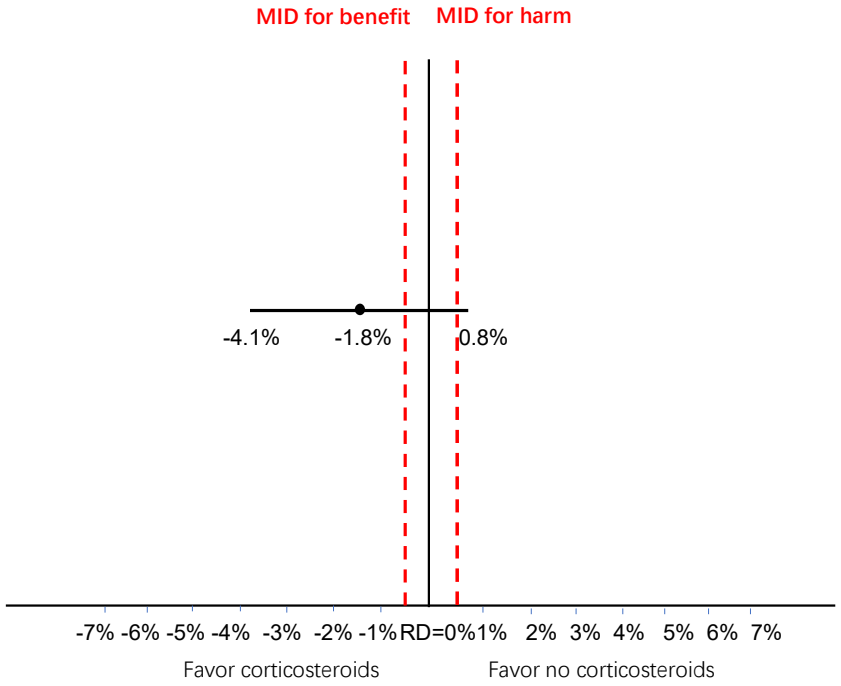
P: patients with sepsis
I/C: corticosteroids versus no corticosteroids
O: death (short-term)

Threshold of interest:
MID for benefit at a reduction of 0.5%

Target of certainty rating:
Corticosteroids have an important reduction in death.

Imprecision judgment:
As the CI includes important harm (i.e. an important increase in death), the authors should likely consider rating down two levels for imprecision.

Plain language summary:
Corticosteroids "probably" have an important benefit (Rating down one)
Corticosteroids "may" have an important benefit (Rating down two)



Circumstances when one should consider rating down two levels for imprecision based on CI approach

1) When rating the certainty that there is a true important benefit, the point estimate reflects **an important benefit**, and the boundary of the CI least favorable to the intervention includes the possibility of harm, particularly **important harm**. (Example 1)

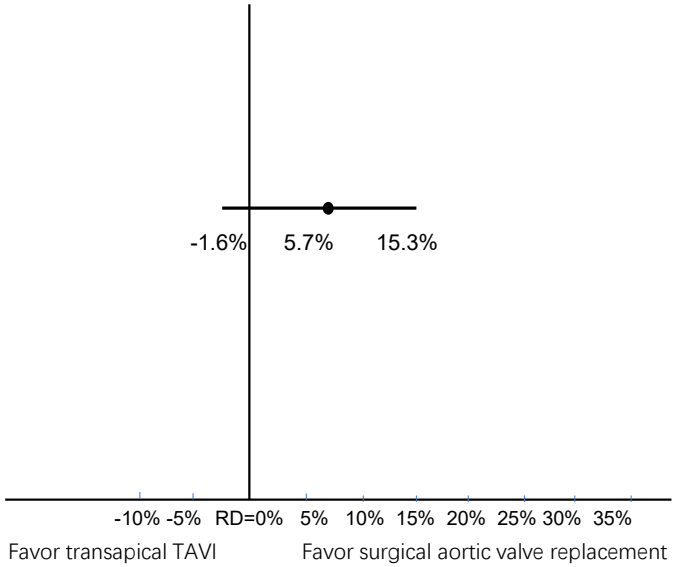
Example 2

P: patients with severe aortic stenosis at low and intermediate risk of intra- or perioperative death
I/C: transcatheter aortic valve implantation (TAVI) versus surgical aortic valve replacement
O: death

Target of certainty rating:
Transapical TAVI has an important increase in death.

Imprecision judgment:
As the CI includes an important benefits, the authors should likely consider rating down two levels for imprecision.

Plain language summary:
Transapical TAVI "may" have an important harm.



Circumstances when one should consider rating down two levels for imprecision based on CI approach

2) When rating the certainty that there is a true important harm, the point estimate reflects **an important harm**, and the boundary of the CI most favorable to the intervention includes benefit, particularly **important benefit**. (Example 2)

Example 3

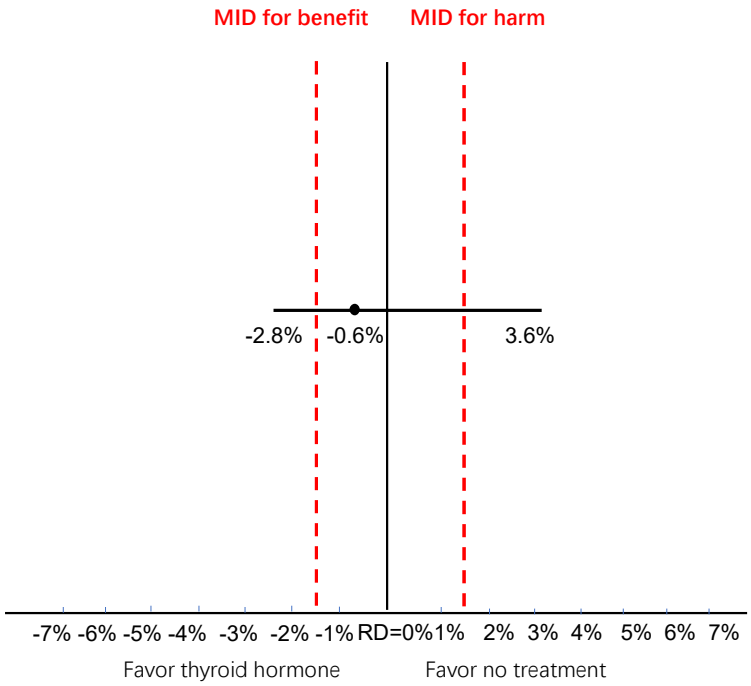
P: patients with subclinical hypothyroidism
I/C: thyroid hormone versus no treatment
O: cardiovascular events

Threshold of interest:
MID for benefit at a reduction of 1.5%

Target of certainty rating:
Thyroid hormone has a trivial or no effect on cardiovascular events.

Imprecision judgment:
As the CI includes both important benefits and important harm, the authors should likely consider rating down two levels for imprecision.

Plain language summary:
Thyroid hormone "may" have trivial or no effect on benefit and harm.



Circumstances when one should consider rating down two levels for imprecision based on CI approach

3) When rating the certainty that the true effect is trivial or no benefit or harm, the point estimate is consistent with **a trivial effect** and the CI includes both **important benefit and important harm**. (Example 3)

Example 4

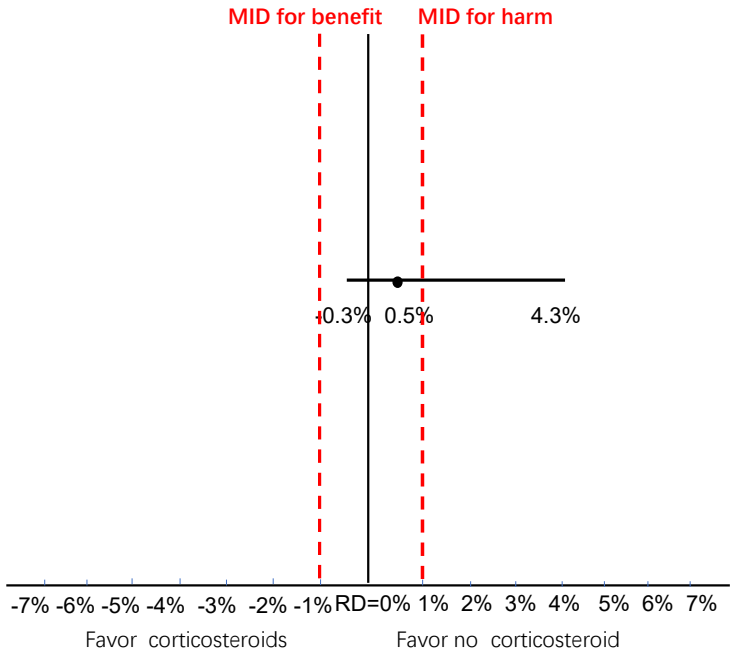
P: patients with sepsis
I/C: corticosteroids versus no corticosteroids
O: strokes

Threshold of interest:
MID for benefit at a reduction of 1.0%

Target of certainty rating:
Corticosteroids have trivial or no effect on stroke.

Imprecision judgment:
As the CI includes an important harm, the more important of the outcome/ the larger the magnitude of effect, the more likely the authors would rating down two levels for imprecision.

Plain language summary:
Corticosteroids "may" have trivial or no effect on benefit



Circumstances when one should consider rating down two levels for imprecision based on CI approach

4) When rating the certainty that the true effect is trivial or no benefit or harm, the point estimate is consistent with **a trivial effect**, and the CI includes substantial (possibly large) **important harm**. (Example 4)

Example 5

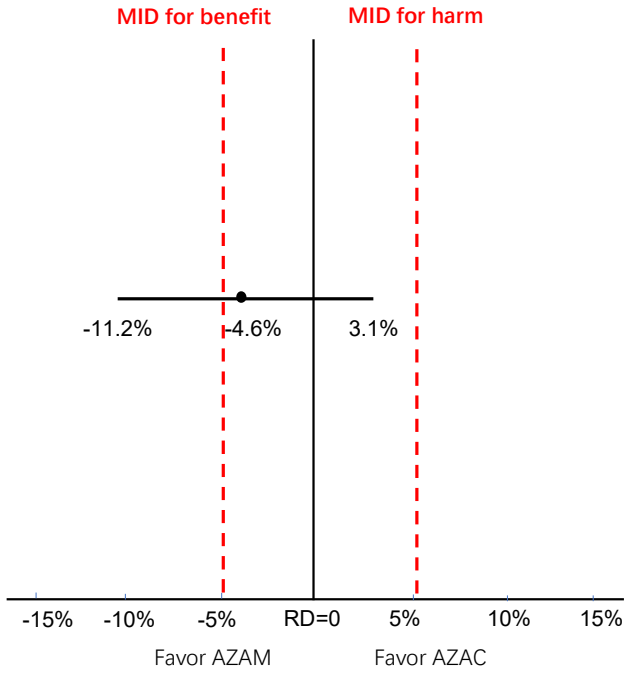
P: patients with acute myeloid leukemia (AML)
I/C: azacitidine monotherapy (AZAM) versus azacytidine combination (AZAC)
O: thrombocytopenia

Threshold of interest:
MID for benefit at a reduction of 5%

Target of certainty rating:
AZAM has a trivial or no effect on thrombocytopenia.

Imprecision judgment:
As the CI includes important benefit, the more important of the outcome/the larger the magnitude of effect, the more likely authors would rate down two for imprecision.

Plain language summary:
Corticosteroids "may" have trivial or no effect on benefit (Rating down two)



Circumstances when one should consider rating down two levels for imprecision based on CI approach

5) When rating the certainty that the true effect is a trivial or no benefit or harm, the point estimate is consistent with a **trivial effect**, and the CI includes substantial (possibly large) **important benefit**. (Example 5)

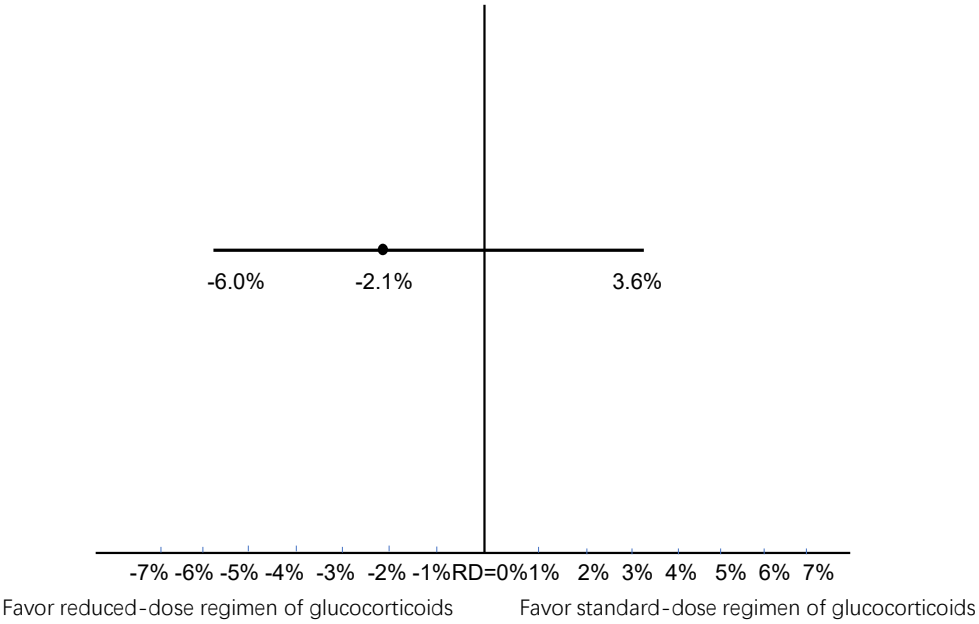
Example 6

P: patients with vasculitis
I/C: reduced-dose regimen of glucocorticoids versus standard-dose regimen of glucocorticoids
O: mortality

Target of certainty rating:
The reduced-dose regimen of glucocorticoids reduces mortality.

Imprecision judgment:
Although the point estimate suggests a benefit, an important harm remains plausible.

Plain language summary:
The reduced-dose regimen of glucocorticoids "may" have a benefit in reducing mortality.



Circumstances when one should consider rating down two levels for imprecision based on CI approach

6) When rating the certainty of non-zero benefit, the point estimate suggests **benefit**, and the CI includes **important harm**. (Example 6)

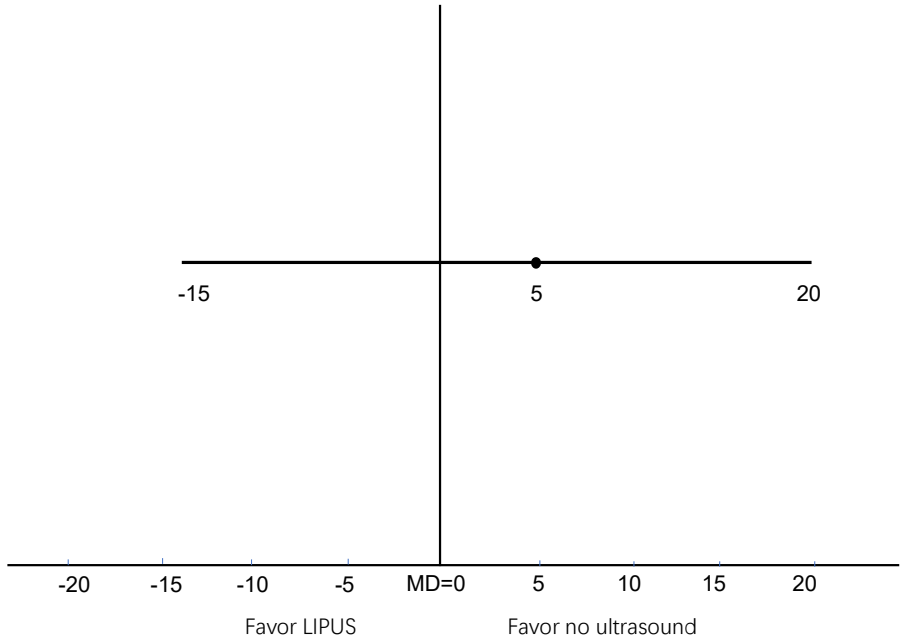
Example 7

P: patients with any type of fracture
I/C: low intensity pulsed ultrasound (LIPUS) versus no ultrasound
O: return back to work

Target of certainty rating:
LIPUS increases the days off required before returning to work.

Imprecision judgment:
CI includes an important benefit.

Plain language summary:
LIPUS "may" increases the days off required before returning to work.



Circumstances when one should consider rating down two levels for imprecision based on CI approach

7) When rating the certainty of non-zero harm, the point estimate suggests **harm**, and the CI includes **important benefit**. (Example 7)

Circumstances when one should consider rating down two levels for imprecision based on CI approach

1) When rating the certainty that there is a true important benefit, the point estimate reflects an important benefit, and the boundary of the CI least favorable to the intervention includes the possibility of harm, particularly important harm. (Example 1)

2) When rating the certainty that there is a true important harm, the point estimate reflects an important harm, and the boundary of the CI most favorable to the intervention includes the possibility of benefit, particularly important benefit. (Example 2)

3) When rating the certainty that the true effect is trivial or no benefit or harm, the point estimate is consistent with a trivial effect and the CI includes the possibility of both important benefit and important harm. (Example 3)

4) When rating the certainty that the true effect is trivial or no benefit or harm, the point estimate is consistent with a trivial effect, and the CI includes the possibility of substantial (possibly large) important harm. (Example 4)

5) When rating the certainty that the true effect is trivial or no benefit or harm, the point estimate is consistent with a trivial effect, and the CI includes the possibility of substantial (possibly large) important benefit. (Example 5)

6) When rating the certainty of non-zero benefit, the point estimate suggests benefit, and the CI includes the possibility of important harm. (Example 6)

7) When rating the certainty of non-zero harm, the point estimate suggests harm, and the CI includes the possibility of important benefit. (Example 7)

When one should check optimal information size (OIS)

- When the CI does not cross threshold(s) of interest and the effect is large (RRR over 30%) and implausible, GRADE suggests considering whether the OIS is met

- When GRADE suggests rating down two levels for imprecision based on OIS calculation ?

- Note: OIS approach focuses on the relative estimate of effect.

Circumstances when one should consider rating down two levels for imprecision based on OIS calculation

1) For dichotomous outcomes, when the ratio of the upper to the lower boundary of the CI is more than 2.5 for odds ratio (odds ratio, OR) or 3 for risk ratio (risk ratio, RR).

ELSEVIER

Journal of Clinical Epidemiology
Volume 139, November 2021, Pages 49-56

Original article

GRADE guidelines 33: Addressing imprecision in a network meta-analysis

Romina Brignardello-Petersen [a], Gordon H. Guyatt [a], Reem A. Mustafa [a, b], Derek K. Chu [a], Monica Hultcrantz [c], Holger J. Schünemann [a, d], George Tomlinson [e]

Example 8

P: patients with chronic suppurative otitis media (1 RCT, 40 patients)
I: topical antibiotics (15/20, 75%)
C: no treatment (8/20, 40%)
O: resolution of ear discharge

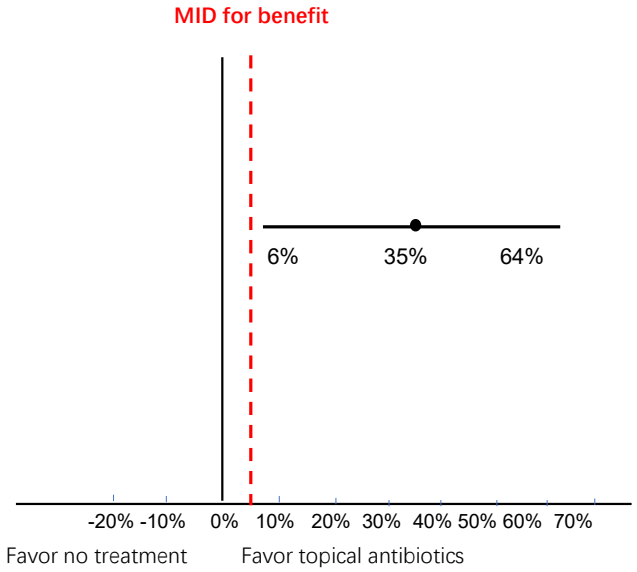
Relative effect: 88% increase (RR 1.88, 95% 1.04 to 3.39)

Threshold of interest:
MID for benefit at a reduction of 5%.

Target of certainty rating:
Topical antibiotics have an important benefit.

Ratio of the upper to the lower boundary of CI around RR:
3.26 (3.39/1.04) > 3

Imprecision judgement: rating down two levels



Circumstances when one should consider rating down two levels for imprecision based on OIS calculation

2) When total sample size of a meta-analysis is smaller than 30%-50% of OIS (i.e. smaller than 30%-50% of 800).

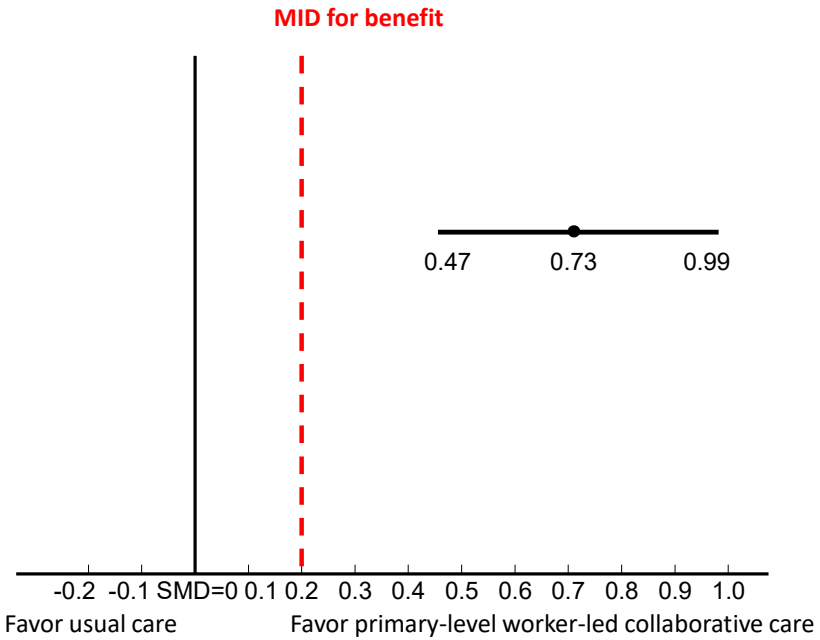
Example 9

P: people living with mental disorders and distress (2 RCTs, 249 patients)
I: primary-level worker interventions versus usual care
O: quality of life

Threshold of interest:
MID for benefit at a reduction of 0.2 standard deviations

Target of certainty rating:
Primary-level worker interventions had an important increase in the quality of life.

Check if OIS is met:
More conservative: Sample size < 50% of OIS (249 vs. 400)
Less conservative: Sample size > 30% of OIS (249 vs. 240)



Circumstances when one should consider rating down two levels for imprecision based on OIS calculation using a minimally contextualized approach

- For dichotomous outcomes, when the ratio of the upper to the lower boundary of the CI is more than 2.5 for odds ratio (odds ratio, OR) or 3 for risk ratio (risk ratio, RR). (Example 8)
- For continuous outcomes, when the total sample size of a meta-analysis is smaller than 30%-50% of the OIS. (Example 9)

An exception: when the baseline risk is low, GRADE suggests being more conservative in rating down for imprecision

- When the baseline risk is very low (rare event), any changes (even big change) in the relative estimates of effect would result in small changes in absolute estimates of effect.
- Focusing on the CI around absolute effect would lead one to reject rating down more than one level for imprecision.
- What a very low baseline risk is depends on the importance of outcome.

Example 10

P: patients with intermittent claudication (2 RCTs, 300 patients)
I: cilostazol (16/150, 10.6%)
C: placebo (7/150, 4.6%)
O: abnormal stools (adverse event)

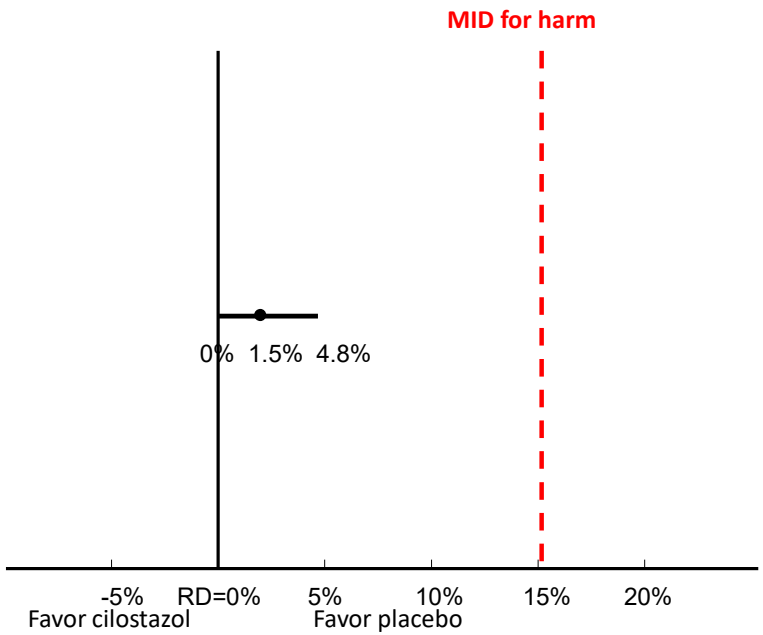
Relative effect: OR:2.44, 95% CI 0.97 to 6.11; RR 2.29, 95% CI 0.97 to 5.40

If baseline risk: 1%, RD: 1.5%, 95%CI 0% to 4.8%

Threshold of interest:
MID for harm at an increase of 15%

Target of certainty rating:
Cilostazol had a trivial effect on abnormal stools.

Ratio of the upper to the lower boundary of CI around OR:
6.3 (6.11/0.97) > 2.5



Simulations of estimated risk differences with the change of odds ratio at three different baseline risks

Cilostazol group		Control group		OR (95% CI)	RD when BS is 1% (95% CI)	RD when BS is 3% (95% CI)	RD when BS is 5% (95% CI)
No. of patients experiencing event	No. of patients not experiencing event	No. of patients experiencing event	No. of patients not experiencing event				
16	134	7	143	2.44 (0.97, 6.11)	1.5% (-0.026%, 4.8%)	4.0% (-0.078%, 1.3%)	6.3% (-0.12%, 19%)
16	134	6	144	2.87 (1.09, 7.53)	1.8% (0.088%, 6.1%)	5.1% (0.26%, 15%)	8.1% (0.42%, 23%)
16	134	5	145	3.46 (1.23, 9.71)	2.4% (0.23%, 7.9%)	6.7% (0.67%, 20%)	10% (1.1%, 28%)
16	134	4	146	4.36 (1.42, 13.36)	3.2% (0.41%, 10%)	8.9% (1.2%, 26%)	14% (1.9%, 36%)

BS: baseline risk; CI: confidence interval; RD: risk difference; OR: odds ratio

When GRADE suggests rating down three levels for imprecision

- Authors might consider rating down three levels for imprecision using a minimally contextualized approach.

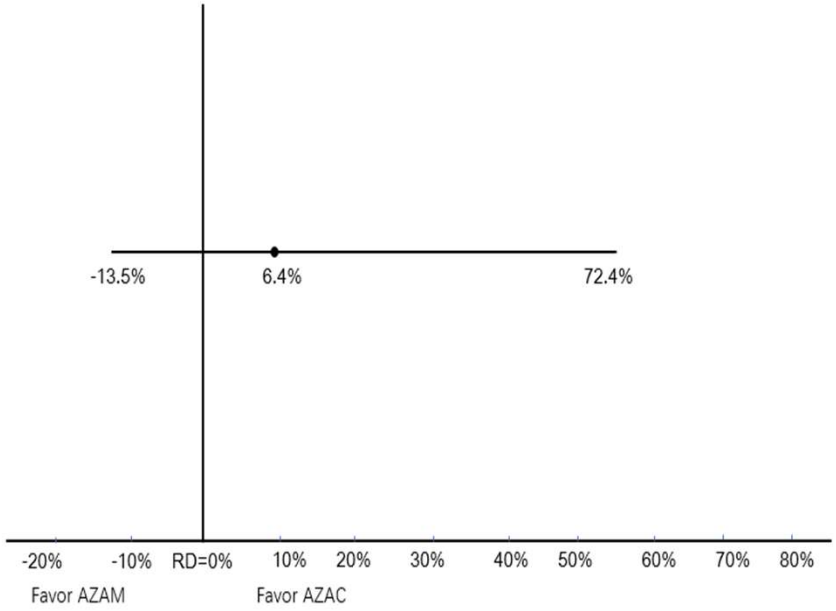
Example 11

P: patients with acute myeloid leukemia (AML)
I: azacitidine monotherapy (AZAM)
C: azacitidine combination (AZAC)
O: septic shock

Threshold of interest:
Null effect threshold (i.e. RD=0%)

Imprecision: Rating down three levels

Target of certainty rating: no need to decide



Summary

- Inform systematic review and guideline authors who choose to use the CI approach to rating imprecision within a minimally contextualized framework.
- Emphasize the usefulness of the CI approach, reserving OIS calculation to situations of implausibly large treatment effects.
- Focus on the circumstances in which GRADE users will seriously consider rating down two levels for imprecision.