Journal of the Epidemiology Foundation of India
Vol 1 (1), Dec 2023, pp. 05-08

REVIEW ARTICLE

# A brief overview of the GRADE approach for rating strength of evidence and practice recommendations

**Manya Prasad[1], Pradeep Aggarwal[2]**

[1] Department of Clinical Research & Epidemiology, Institute of Liver and Biliary Sciences, New Delhi

[2]Department of Community & Family Medicine, All India Institute of Medical Sciences Rishikesh, Uttarakhand

**CORRESPONDING AUTHOR**

Dr Manya Prasad, Assistant Professor, Department of Clinical Research & Epidemiology, Institute of Liver and Biliary Sciences, New Delhi - 110070

Email: manya.2311@gmail.com

**CITATION**

Prasad M. A brief overview of the GRADE approach for rating strength of evidence and practice recommendations. Journal of the Epidemiology Foundation of India. 2023;1(1):05-08.

https://doi.org/10.56450/JEFI.2023.v1i01.002

**ARTICLE CYCLE**

Received: 18/12/2023; Accepted: 20/12/2023; Published: 31/12/2023



## INTRODUCTION

The medical fraternity is faced with an avalanche of medical literature. While the current best evidence should inform practice guidelines and recommendations, indeed, much of the data available is not of the highest standard, and suggestions based on sub-par evidence may result in undesirable outcomes. Moreover, many context-specific considerations, such as costs and patient preferences may be key concerns in formulating recommendations. A methodical methodology to evaluate the effectiveness of suggestions and the standard of information is necessary to address these difficulties. The "GRADE (Grading of Recommendations, Assessment, Development, and Evaluation)" offers such a framework for any given clinical question and offers advantages over traditional approaches to guideline development. Crucial elements for ensuring the trustworthiness of the guideline development process, such as recruiting an appropriate panel, excluding conflicts of interest, systematically reviewing the best evidence, rating certainty in evidence, and incorporating patient values and preferences, are rendered due importance in the process.

GRADE evaluates the entire body of available information on a clinical question, subsequently informing guideline development. The body of scientific data is usually presented as a meta-analysis and systematic review. Since its initial publication in 2004 (1), the framework has been widely recognized and implemented by more than 110 organizations globally. Notable amongst these are the "World Health Organization", "Cochrane Collaboration", "Centre for Disease Control", "National Health Service", and many medical associations around the world. Additionally, GRADE is included in thousands of clinical practice guidelines included in online texts like Dynamed and UpToDate.

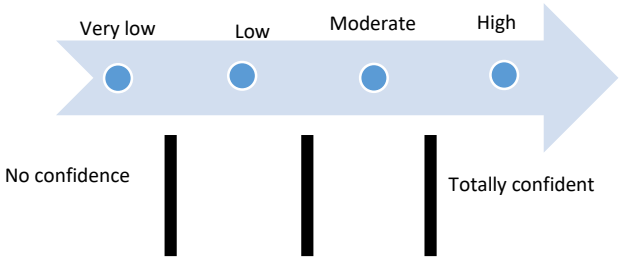
## PROCESS

GRADE provides a method for assessing and rating the quality of the evidence while accounting for several variables, including study design (observational or randomized controlled trial), bias risk, discrepancy, ambiguity, errors, and publishing bias. This

systematic and transparent framework aims to consider all methodological issues that may threaten the validity of conclusions for the body of evidence (2).

Evidence-based certainty is thought of as existing on a spectrum that extends from extremely high confidence to no confidence at all. The GRADE system divides certainty into four categories along this continuum: high, average, low, and extremely low. (Figure 1). In this context, the terms 'certainty', 'confidence', or 'strength' indicate the same concept.

**Figure 1. The confidence in evidence is on a continuum**

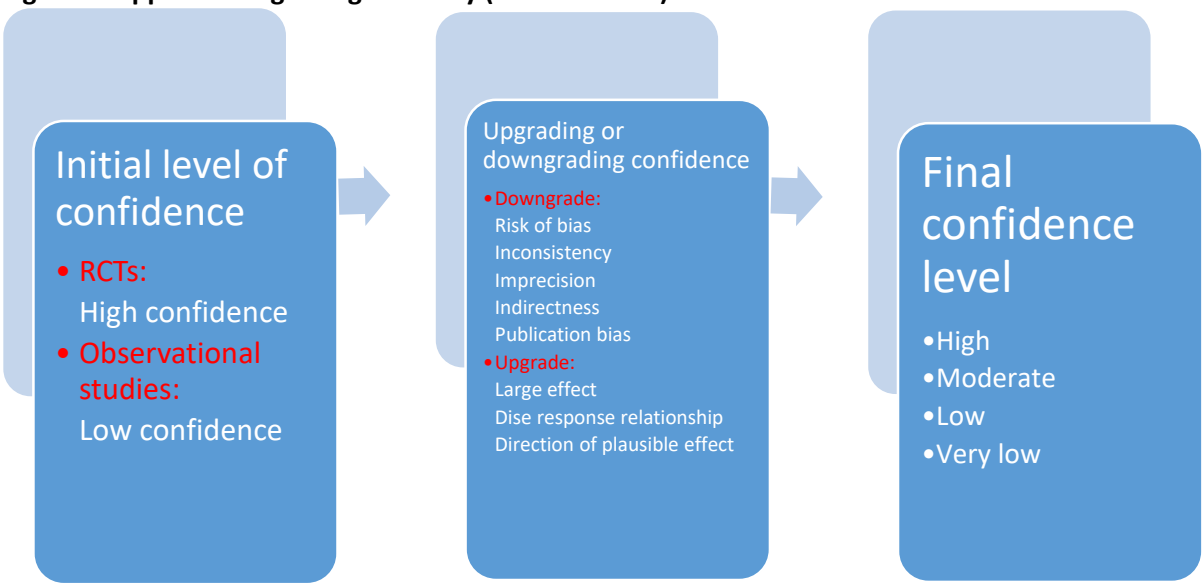



Evidence of high quality or certainty comes from carefully carried out research with reliable findings and little chance of error. Conversely, very low certainty signifies insufficient evidence to support firm conclusions, arising from studies with high risk of bias, inconsistent results, few events or sample, or other concerns (3).

The following are the key issues that determine rating up or rating down of the body of evidence for each outcome in the body of evidence, i.e., a systematic review: (Figure 2):

**1. Randomized or non-randomized study Design**: Because of their strict methodology, randomized controlled trials (RCTs) are regarded as being of greater value. Therefore, within the framework, RCTs start with high confidence by default. However, several factors could cause estimates from RCTs to be rated with less certainty or reliability. Conversely, observational studies have a lower degree of confidence at first but could be raised in specific circumstances. (4).

**Figure 2. Approach to grading certainty (or confidence) in estimates**




**2. Downgrading Evidence:**

Several variables can affect how certain evidence is rated. A summary is given below:

**2.1. Risk of Bias**: Refers to the possibility of systematic errors producing skewed outcomes. For RCTs, considerations include "random sequence generation", "allocation concealment", "blinding", completeness of outcome data, analysis by intention to treat, and other sources of bias, such as stopping early for benefit and selective outcome reporting.

**2.2. Inconsistency**: Refers to dissimilarity in results across different studies, also known as heterogeneity. Consistent findings across

multiple studies increase confidence in the evidence. Assessment involves visual inspection of forest plots or statistical methods like the I-squared statistic, where higher I2 indicates more concern for inconsistency. An I-squared statistic value of more than 50% raises some concern about inconsistency, though other cut-offs are reported to guide classification of degrees of inconsistency (5).

**2.3. Indirectness**: Evaluates how well the existing data directly answers the relevant clinical question while considering variations in the research population, the intervention, the comparison, or the results. In terms of study population, characteristics such as age, comorbidities or degrees of severity may be different. Concerning intervention, dose/duration/route of drug may differ. Similarly, various standards of care and reported surrogate outcomes present challenges that may warrant concerns for indirectness. The presence of indirect comparisons is another example, in terms of the availability of evidence on Drug A and B each versus placebo, but the comparison of interest being Drug A versus Drug B.

**2.4. Imprecision**: Imprecision reflects the uncertainty associated with the estimated effect. Wide confidence intervals indicate imprecision. Recommendations are downgraded for imprecision if a change in the upper or lower boundary of the confidence interval would alter the recommendation. It is advised to compute the summary (or optimal) data size and assess the suitability of sample size and events to assess imprecision in a situation where CIs are adequately narrow, effects are high, but both sample size and number of incidents are modest.

**2.5. Publication Bias**: Occurs when research findings that are of statistical significance or favourable have a higher chance of being published. To effectively assess the possibility of publication bias, a significant number of research must be conducted using a variety of methodologies, each of which has limitations. Some examples are funnel plot (minimum 10 studies needed), Beggs rank correlation, Eggers test (minimum 20 studies), and Doi plot (minimum 5 studies). However, a thorough examination of these tools is outside the purview of this work. (6-9).

**3. Upgrading Evidence**

A significant amount of effect observed in observational studies is one factor that can cause the evidence to be rated higher. Even observational studies can yield high-quality data when a strikingly big effect is observed in a short amount of time. Examples include the use of Epinephrine for anaphylactic shock and Frusemide for pulmonary edema. The existence of a dose-response association and a pattern of plausible bias are two more elements that elevate the quality of the evidence.

**4. Going from evidence to recommendations**

GRADE acknowledges the critical relationship between the degree of evidence certainty and the potency of suggestions.

Higher certainty evidence justifies more robust recommendations and may entail becoming quality of care criteria. Conversely, less certain evidence could result in suggestions that are less strong or limited, and these are situations where patient values and preferences have a bigger role to play, entailing due expression of uncertainty in interaction with patients.

**5. Evidence-to-decision framework**

Several key considerations play a pivotal role in determining the direction and strength of recommendation. These considerations are encapsulated within the Evidence-to-Decision framework, taking into account factors such as the balance of advantages and risks, beliefs and choices, cost-efficiency, feasibility, equity, etc. (10). The evidence-to-decision framework takes into account the evidence synthesis, and puts it in context of other key considerations that influence recommendations.

Detailed guidance on upgrading and downgrading evidence is available online, along with a freely accessible software for constructing GRADE profiles and summary of findings tables, GRADEPro (https://www.gradepro.org/) (12).

## CONCLUSION

In summary, the GRADE tool represents an important application of clinical epidemiology in medicine and represents a systematic structure for evaluating the efficacy of evidence and recommendations. The methodology continues to evolve, adding aspects for diagnosis questions and refining techniques for evaluating imprecision. It has a noticeable influence on the creation and application of guidelines, encouraging openness, coherence, and efficient dissemination of data and suggestions.

## REFERENCES


1. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004 Jun 19;328(7454):1490.
2. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-926.
3. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401-406.
4. Schünemann HJ, Cuello C, Akl EA, et al. 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. J Clin Epidemiol. 2019;111:105-114.
5. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. J Clin Epidemiol. 2011;64(12):1294-1302.
6. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011 Jul 22;343:d4002.
7. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ. 2006 Sep 16;333(7568):597-600.
8. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol. 2000 Nov;53(11):1119-29.
9. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994 Dec;50(4):1088-101.
10. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016;353:i2089.
11. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. J Clin Epidemiol. 2017;81:101-110.
12. GRADE Working Group. GRADE Handbook. Updated October 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html


.