An Evaluation of the Impact of Evidence Grouping on Certainty Grading when Performing a GRADE Assessment Without a Meta-Analysis

Penelope Cervelo Bouzo[1], Gillian Sibbring[2], Lucia Giles[3]

[1]Prime Market Access Ltd, London, UK, [2]Prime Market Access Ltd, Knutsford, UK, [3]Prime Market Access Ltd, Oxford, UK

Presented at the Professional Society for Health Economics and Outcomes Research (ISPOR) Europe Annual Meeting, 12-15 November 2023, Copenhagen, Denmark

PT13

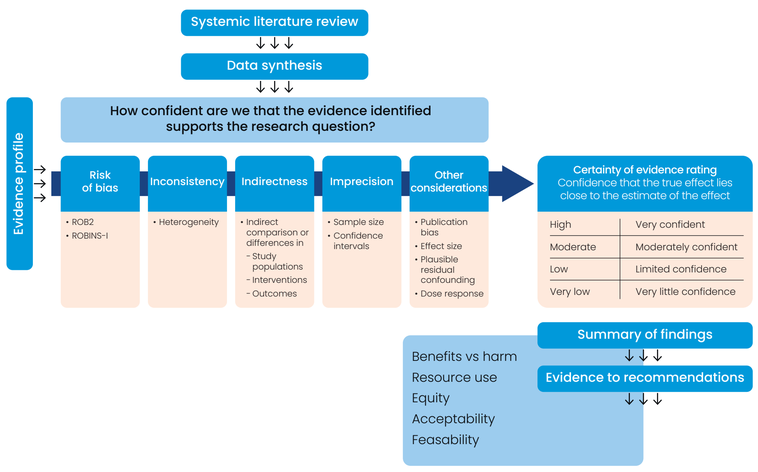
Background

- The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach provides a reproducible, transparent framework for rating the certainty of a body of evidence identified systematically.[1]
- Meta-analysis is often used to inform GRADE assessment but may not always be feasible, especially when the evidence is heterogenous.[2]
- To analyse the impact of evidence grouping on certainty rating in the absence of a meta-analysis, we conducted two GRADE assessments on the same body of evidence grouped differently, and compared the differences in certainty ratings.

Methods

- Embase and MEDLINE databases were searched *via* the Elsevier platform, using indexed (Emtree) terms, to identify randomised controlled trials (RCTs) comparing patient-important outcomes in patients with chronic obstructive pulmonary disease (COPD) receiving triple versus dual inhaled therapy.
- Eligibility criteria were defined using the Population, Intervention, Comparison, Outcomes and Study framework informed by the 2023 Gold Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines.[3]
  - The search was limited to human studies published between 11 April 2015 and 11 April 2023. Conference abstracts, proceedings and reviews were excluded.
  - Screening was conducted using web-based systematic review software and was aided by continuous artificial intelligence reordering/reprioritisation of records (DistillerSR. version 2.35; https://www.distillersr.com/).
- Owing to the methodological focus of this study, only a single patient-reported outcome, rate of disease exacerbation, was included in the GRADE assessment, with data only extracted for this measure. Selection of this outcome was informed by the 2023 GOLD guidelines.[3]
  - Risk of bias was assessed using the Cochrane Risk of Bias 2 (ROB2) tool for RCTs[4] as recommended in the GRADE handbook.[5]
- The GRADE framework[1] was applied to RCTs reporting annual rate of exacerbation in patients with COPD receiving triple or dual therapy (**Figure 1**).
  - One assessment considered all evidence identified on the endpoint of interest, whereas the second assessment grouped the evidence by measurement criteria; the evidence certainty rating was compared between assessments.

Figure 1. Diagram of the GRADE assessment process



Results

- A PRISMA diagram shows the flow of identified evidence in **Figure 2**.
- Data were extracted from five primary publications of five RCTs that reported annual rate of COPD exacerbations identified in the literature review.
- The risk of bias was considered as 'high risk' for three RCTs and the remaining two RCTs had 'some concerns'.
- All studies identified reported a statistically significant lower annual rate of exacerbations with triple versus dual therapy ($P<0.05$).
- When all identified evidence (Studies I-V) was assessed as a single group (**Table 1A**), the certainty of the body of evidence was rated 'very low' as a result of evidence downgrading in the 'risk of bias', 'imprecision' and 'indirectness' domains.
- When the evidence was grouped by measurement criteria (**Table 1B**), the evidence measured using the same criteria (Studies III-V) had a 'moderate' certainty rating, with evidence downgraded in the 'risk of bias' domain only.
- The evidence for which the measurement criteria for the outcome were not reported (Studies I and II), was rated as 'very low' certainty, owing to downgrading in the 'risk of bias', 'indirectness' and 'imprecision' domains.

Figure 2. PRISMA diagram showing the flow of identified evidence

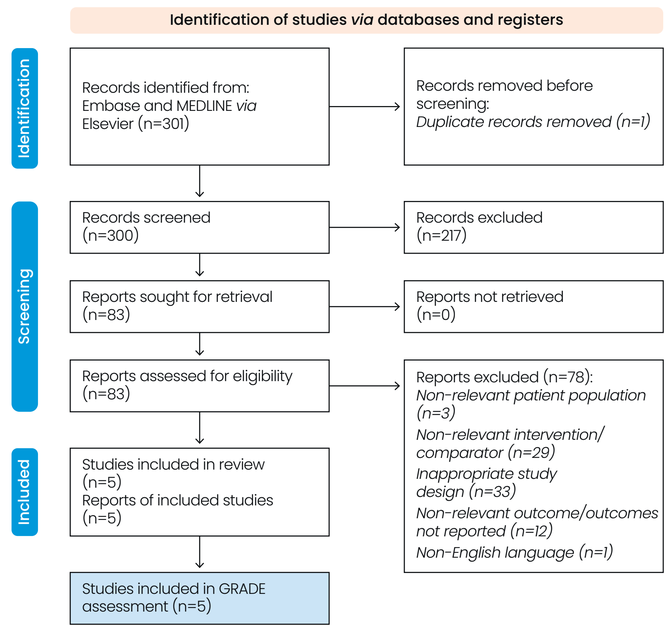


Table 1. Assessment of the certainty of evidence for the use of triple versus dual therapy on annual rate of exacerbation

A. Annual rate of exacerbation with triple vs dual therapy (all evidence)

	Quality assessment						Summary of findings
Outcome	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence per outcome and importance	Effect size for triple vs dual therapy
Critical outcome: annual rate of moderate-to-severe COPD exacerbations, assessed using multiple measurement criteria (follow-up time, 24 and 52 weeks)							
Study I, N=8509 *Study II, N=1902* *Study III, N=10 355* *Study IV, N=1532* *Study V, N=8588*	–2 Very serious risk of bias (methods of handling missing outcome data not reported and concerns around the results selected for reporting)	0 No serious concerns	–1 Serious indirectness (methods of outcome measurement not fully reported in all studies; follow-up time differed between studies)	–1 Serious imprecision (no power calculation or sample size reported)	0 No serious concerns	●○○○ Very low	All five studies reported a significant reduction in the rate of exacerbations with triple therapy vs dual therapy

B. Annual rate of exacerbation with triple vs dual therapy (grouped by measurement criteria)

	Quality assessment						Summary of findings
Outcome	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence per outcome and importance	Effect size for triple vs dual therapy
Critical outcome: annual rate of moderate-to-severe COPD exacerbations, assessed using the Anthonesian criteria (follow-up time, 52 weeks							
Study III, N=10 355 *Study IV, N=1532* *Study V, N=8588*	–1 Serious risk of bias (methods of handling missing outcome data not reported and concerns around the results selected for reporting)	0 No serious concerns	0 No serious concerns	0 No serious concerns	0 No serious concerns	●●●○ Moderate	All three studies reported a statistically significant reduction ($P<0.05$) in the rate of moderate or severe exacerbations with triple therapy vs dual therapy
Critical outcome: annual rate of moderate-to-severe COPD exacerbations, measurement criteria not reported (follow-up time, 24 and 52 weeks)							
Study I, N=8509 *Study II, N=1902*	–1 Serious risk of bias (methods of handling missing outcome data not reported and concerns around the results selected for reporting)	0 No serious concerns	–1 Serious indirectness (methods of outcome measurement not reported; follow-up time was 24 and 52 weeks)	–1 Serious imprecision (no power calculation or sample size reported)	0 No serious concerns	●○○○ Very low	Both studies reported a significant reduction ($P<0.001$) in the rate of exacerbations with triple therapy vs dual therapy

Discussion

- The difference in certainty ratings between the two assessments was driven by 'serious' concerns in the 'indirectness' and 'imprecision' domains, introduced by lack of reporting on outcome measurement and sample-size calculations in Studies I and II.
- When all five studies (Studies I–V) were combined (**Table 1A**), downgrading in the 'indirectness' and 'imprecision' domains resulted in a substantial reduction in the certainty of evidence, whereas when the studies were grouped by measurement criteria (**Table 1B**), Studies I and II had a 'very low' certainty rating and Studies III–V had a 'moderate' certainty rating.
- Irrespective of evidence grouping, the evidence was downgraded in the 'risk of bias' domain, marked by consistent downgrading for lack of reporting on the methods used to handle missing data and concerns around the results selected for reporting.

Limitations

- Owing to the methodological focus of this study, searching of MEDLINE and Embase was conducted in the Elsevier platform only, meaning that not all relevant publications may have been captured.
- Screening, data extraction and evidence grading were predominantly conducted by a single reviewer, which does not fully align with the recommendations made in the Cochrane handbook.[6]
- Patient-important outcomes considered in the GRADE assessment should be pre-defined prior to literature searching and not driven by data availability, which was not the case in this study owing to the methodological focus.
- Additionally, it is advisable to include an expert panel when selecting and rating patient-important outcomes as part of the GRADE assessment process,[5] as well as when considering clinical implications.

Conclusions and wider implications

- We illustrate how grouping of RCT evidence, for example by outcome measurement criteria, can affect certainty ratings as per the GRADE framework.
- This is especially pertinent when considering the growing body of literature reporting real-world evidence, and the numerous sources of heterogeneity associated with different study designs, characteristics of included patients, outcomes and measures reported.
- Finally, we show how the quality of reporting can affect the confidence in the evidence assessed. Again, this is of relevance when considering the need for evidence-based decision-making to drive access to innovative technology for patients.[7]

INFORMATION PANEL

Why conduct a GRADE assessment?

- GRADE assessment is used to inform decision-makers on the reliability of the evidence they are using to inform their decisions.[2] This is of importance irrespective of the feasibility to perform a meta-analysis.

Did you know?

- GRADE is the most widely adopted tool for evidence grading and for making clinical practice recommendations, with endorsements from >100 organisations worldwide.[1]

Situations in which quantitative data synthesis (such as meta-analysis) may not be feasible

- Where studies measure outcomes differently or outcomes are not standardised, for example in rare disease or real-world evidence.[2]
- When relevant data are not available or cannot be calculated for all included studies.[6]
- When there is substantial bias in the evidence.[6]
- When there are time constraints,[2] for example if healthcare recommendations are required rapidly in a changing disease landscape, such as during the coronavirus disease 2019 pandemic.

References

1. BMJ Best Practice. What is GRADE? Available at: https://bestpractice.bmj.com/info/toolkit/learn-ebm/what-is-grade/. Accessed 30 August 2023.
2. Murad MH, et al. *Evid Based Med.* 2017;22:85–87.
3. Global Initiative for Chronic Obstructive Lung Disease, GOLD. Global strategy for prevention, diagnosis and management of COPD: 2023 Report. Available at: https://goldcopd.org/2023-gold-report-2/. Accessed October 2023.
4. Sterne JAC, et al. *BMJ.* 2019;366:l4898.
5. Schünemann H, et al. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group. 2013. Available at: https://training.cochrane.org/resource/grade-handbook. Accessed 30 August 2023.
6. Higgins JPT, et al. *Cochrane handbook for systematic reviews of interventions* version 6.3 (updated February 2022). 2022. Available at: www.training.cochrane.org/handbook. Accessed 30 August 2023.
7. Goh, SH, Walsh, K. BMJ Best Practice. Clinical decision support: improving access to improve patient care. Available at: https://bestpractice.bmj.com/info/clinical-decision-support-improving-access-to-improve-patient-care/. Accessed 4 October 2023.

Acknowledgements and disclosures

This work was sponsored by Prime Market Access Ltd. All authors are employees of Prime.
We would like to thank Sofie Norregaard for her support with conducting the systematic review.