

Tiotropium versus placebo in chronic obstructive pulmonary disease: meta-analysis of trough FEV1 and non-fatal serious adverse events

Draft, 17 October 2026

Methods

Studies and outcomes

Four randomized trial reports comparing tiotropium with placebo in chronic obstructive pulmonary disease (COPD) were analysed: Bateman (2010), Freeman (2007), Magnussen (2008) and Tonnel (2008). The reports were selected from those included in the Cochrane review by Karner (2014); no new literature search or screening was performed. Bateman (2010) reports two randomized trials (NCT00168831 and NCT00168844) in one publication.

Two outcomes were analysed: change from baseline in trough forced expiratory volume in one second (FEV1), in mL, and the number of participants with at least one non-fatal serious adverse event (SAE). Outcome data were taken from the published reports, and the values used in the analyses were checked against a reference sheet of the values printed in the source articles.

Effect measures and synthesis

FEV1 was pooled as a common-effect inverse-variance mean difference (MD) in mL with a Wald 95% confidence interval (CI). Tonnel (2008) values were converted from L to mL by multiplying the contrast and its standard error (SE) by 1000. SAE were pooled as a common-effect Mantel–Haenszel odds ratio (OR) with its native 95% CI. Common-effect models were used as the main models, following Karner (2014). A protocol-specified DerSimonian–Laird random-effects model was fitted for FEV1 as a sensitivity analysis. Heterogeneity was described with I^2 and, for the random-effects model, with τ^2 and Q-profile confidence intervals. Analyses were performed in R 4.6.1 with the metafor package (version 5.2-1).

The SAE outcome counted participants with at least one SAE, each participant once. In Bateman (2010), the two tiotropium arms were combined (108 + 125 = 233 participants with at least one SAE among 670 + 667 = 1337 participants) and compared once with the shared placebo group (110 of 653 participants). In Tonnel (2008), 42 of 266 participants in the tiotropium group and 38 of 288 in the placebo group had at least one SAE. The printed SAE rows do not separately resolve non-fatal status; deaths were not subtracted, and missing counts were not imputed.

Reports not pooled

Bateman (2010) was presented narratively for FEV1 and excluded from the FEV1 pool only. Its Table 3(A) reports adjusted contrasts of 0.127 L (SE 0.013 L) and 0.150 L (SE 0.013 L) against a shared placebo; neither compatible arm-level mean changes with standard deviations nor a covariance for the two contrasts is available to combine them. The Karner (2014) value of 141 mL (SE 6 mL) was not substituted.

Freeman (2007) was presented narratively for FEV1. Its Table 2 reports median and interquartile-range responses and a Hodges–Lehmann shift estimate of 0.06 L (95% CI 0.01 to 0.10), not a mean

difference. A mean and SE were not imputed, and subgroup contrasts were not used as independent studies.

Freeman (2007) and Magnussen (2008) did not enter the SAE pool because usable SAE counts were missing; missing counts were neither treated as zero nor imputed.

Certainty of evidence

Certainty of evidence was rated for each outcome with GRADE. Evidence from randomized trials starts at high certainty and can be rated down for risk of bias, inconsistency, indirectness, imprecision and publication bias.

Imprecision was assessed against the null: each pooled confidence interval was compared with no effect (an MD of 0 mL for FEV1 and an OR of 1 for SAE). This target concerns support for a non-zero effect, not clinical importance (Zeng 2022, GRADE Guidance 34; Guyatt 2025, Core GRADE 2). No minimally important difference or clinical benefit/harm threshold was prespecified, so certainty about a clinically important effect, or about little or no important difference, was not assessed. An information-size downgrade was not added to an interval-based downgrade. No baseline risk was prespecified, and absolute SAE effects were not computed.

Results

Included reports

Four reports were analysed: Bateman (2010), Freeman (2007), Magnussen (2008) and Tonnel (2008). Bateman (2010) combines two randomized trials in one report. Magnussen (2008) enrolled adults with COPD and concomitant asthma and was retained with that population limitation. The FEV1 analysis listed four reports and pooled two (Magnussen and Tonnel); the SAE analysis listed and pooled two report-level contributions (Bateman and Tonnel). These are report counts, not interchangeable counts of independent trials or unique participants.

Narrative FEV1 findings

Bateman (2010) reported two adjusted contrasts against the shared placebo: 0.127 L (SE 0.013 L) and 0.150 L (SE 0.013 L). Freeman (2007) reported a Hodges–Lehmann shift of 0.06 L (95% CI 0.01 to 0.10), rather than a mean difference. These findings are presented narratively and contribute no weight to the pooled FEV1 estimate.

Quantitative synthesis

Both primary models converged. Estimates are rounded to two decimal places.

Table 1. Pooled estimates.

Outcome	Estimate and 95% confidence interval	Listed reports	Pooled contributions
Trough FEV1 change	MD 98.74 mL (62.97 to 134.52)	4	2
Non-fatal serious adverse events	OR 1.08 (0.87 to 1.35)	2	2

Table 2. Trough FEV1 change: study-level estimates (common-effect inverse-variance model).

Report	MD, mL (95% CI)	Weight (%)
Magnussen (2008)	98.00 (52.92 to 143.08)	62.98
Tonnel (2008)	100.00 (41.20 to 158.80)	37.02
Pooled	98.74 (62.97 to 134.52)	100.00

Table 3. Non-fatal serious adverse events: study-level estimates (Mantel–Haenszel model).

Report	Participants with ≥1 SAE / total, tiotropium vs placebo	OR (95% CI)	Weight (%)
Bateman (2010)	233/1337 vs 110/653	1.04 (0.81 to 1.34)	79.89
Tonnel (2008)	42/266 vs 38/288	1.23 (0.77 to 1.98)	20.11
Pooled	423 participants with ≥1 SAE among 2544	1.08 (0.87 to 1.35)	100.00

Study-level 95% CIs are Wald intervals; the pooled SAE interval is the Mantel–Haenszel interval.

The DerSimonian–Laird random-effects sensitivity analysis for FEV1 gave the same MD of 98.74 mL (95% CI 62.97 to 134.52), with $\tau^2 = 0$ (95% CI 0 to 1322.02; boundary fit) and $I^2 = 0\%$ (95% CI 0 to 64.92%). In the common-effect models, I^2 was 0% for both FEV1 and SAE. The SAE interval includes the null value; this does not establish equivalence or absence of harm. Both syntheses contain only two report-level contributions, so heterogeneity estimates are uncertain.

Certainty of evidence

Table 4. GRADE certainty of evidence.

GRADE domain	Trough FEV1 change	Non-fatal serious adverse events
Starting certainty	High	High
Risk of bias	Serious	Serious
Inconsistency	Not serious	Not serious
Indirectness	Not serious	Serious
Imprecision	Not serious	Serious
Publication bias	Not assessable	Not assessable
Certainty	Moderate	Very low

Certainty was moderate for FEV1, rated down for risk of bias, and very low for SAE, rated down for risk of bias, outcome indirectness and imprecision. The FEV1 interval excluded an MD of 0 mL, whereas the SAE interval crossed an OR of 1. Information size was not added as a second penalty to an interval-based downgrade.

References

Included trial reports

Bateman (2010). Efficacy and safety of tiotropium Respimat® SMI in COPD in two 1-year randomized studies. *International Journal of COPD*.

Freeman (2007). Efficacy and safety of tiotropium in COPD patients in primary care – the SPiRiva Usual CarE (SPRUCE) study. *Respiratory Research*.

Magnussen (2008). Improvements with tiotropium in COPD patients with concomitant asthma. *Respiratory Medicine*.

Tonnel (2008). Effect of tiotropium on health-related quality of life as a primary efficacy endpoint in COPD. *International Journal of COPD*.

Methodological references

Guyatt (2025). Core GRADE 2: choosing the target of certainty rating and assessing imprecision. *BMJ*.

Karner (2014). Tiotropium versus placebo for chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews*.

Zeng (2022). GRADE Guidance 34: update on rating imprecision using a minimally contextualized approach. *Journal of Clinical Epidemiology*.