National Institute for Health and Care Excellence

Draft version

Chronic obstructive pulmonary disease in over 16s: diagnosis and management

[J] Evidence reviews for the length of corticosteroid use during exacerbations

NICE guideline NG115

Evidence reviews

February 2019

Draft for Consultation

This evidence review was developed by the NICE Guideline Updates Team

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:**https://www.nice.org.uk/terms-and-conditions#notice-of-rights**

Contents

Corticosteroid use during exacerbations

Review question

Are shorter durations of ≤ 7 days of corticosteroid treatment effective at treating acute exacerbations in people with COPD compared to longer treatments of > 7 days?

Introduction

It is important to ensure that corticosteroid courses are not prescribed for longer than necessary due to the known adverse events associated with corticosteroid use, including fluid retention, pneumonia, hypertension, diabetes mellitus, adrenal suppression and osteoporosis. If there is an opportunity to shorten corticosteroid treatment without losing effectiveness this should be pursued in the interests of patient safety and quality of life. The NICE COPD guideline (NG115) currently recommends that patients with acute exacerbations of COPD should be treated with systemic corticosteroid treatment for 7 to 14 days. However, clinical practice has changed and courses of less than 7 days are now routinely used in the NHS. This review aims to investigate the evidence behind this change in practice and update the guideline accordingly. This review is based upon the 2018 Cochrane review "Different durations of corticosteroid therapy for exacerbations of chronic obstructive pulmonary disease" by Walters et al. (2018).

This review identified studies that fulfilled the conditions listed in Table 1, as specified in the protocol followed by the Cochrane Airways Group (Walters 2018). For full details of the review protocol, see appendix A.

PICO table

Table 1: PICO table for the duration of corticosteroid use during exacerbations

Population	Adults with an acute exacerbation of COPD. (The definition of an acute exacerbation could include any combination of an increase in breathlessness, sputum volume, sputum purulence, cough or wheeze.)
Interventions	Systemic corticosteroid (SCS) given for a period of seven or fewer days.
Comparator	Systemic corticosteroids given for longer than seven days.
Outcomes	• Treatment failure (for example, the need for additional treatment) • Relapse after treatment (e.g. treatment for new acute exacerbation, re-admission or hospitalisation for COPD) • Adverse drug effects • Mortality • Cardiac complications • Lung function (FEV1) • Length of hospital stay • Arterial blood gases • Breathlessness • Quality of life • Resource use and costs

For continuous outcomes: early response was measured on or before day seven of treatment, and end of treatment response measurements were made at the time point equivalent to the end of the longer treatment period.

Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A, and the methods section in appendix B.

Two of the subgroup analyses specified in the review protocol (inpatient versus outpatients and studies with people who had previously used corticosteroids versus corticosteroid naïve people) were not carried out for this review because the majority of included studies did not report data for the categories of interest in an accessible format. However, the majority of included studies could be divided into those administering corticosteroids orally or intravenously (IV) and so this subgroup analysis was conducted.

The GRADE tables only show the results of the subgroup analyses if a difference between subgroups was detected based on a subgroup p value < 0.05. In all other cases, only the pooled results are presented.

The search strategies used in this review are detailed in appendix C.

Declarations of interest were recorded according to NICE's 2014 conflicts of interest policy.

Clinical evidence

Included studies

The Cochrane review upon which this review is based (Walters et al. 2018) is an update of an earlier Cochrane review. This update included the same 8 studies from the previous version of the Cochrane review as no new evidence was found.

The systematic search was updated by the Cochrane Airways Group on behalf of the Guideline Updates Team to identify any trials that were published after the final search for the Cochrane review. This search returned 166 results. Full details of the review protocol and literature search strategy can be found in appendix A and appendix C. After title and abstract screening all studies other than the 8 original includes and 1 new study were excluded. This single new study was excluded at full text screening due to the paper being a secondary publication of an included study that did not provide any additional relevant information.

As a result, the 8 studies included in this review are the 8 studies from the original Cochrane review. Of these studies, only 5 provided sufficient data to be included in the meta-analysis.

Two abstracts were included in the meta-analysis in the Cochrane review based on additional data obtained from the authors (Wood-Baker et al.1997 and Sirichana et al. 2008). For the other 3 abstracts identified by the Cochrane review, unpublished data was sought, but was insufficient for the studies to be included in the quantitative meta-analysis (Gomaa et al. 2008, Rahman et al. 2004, Salam et al. 1998).

The process of study identification is summarised in the PRISMA diagram in appendix D.

The references of individual included studies are listed in appendix J.

Excluded studies

Details of the study excluded at full-text, with reasons for exclusion, is given in appendix I.

Summary of clinical studies included in the evidence review

Only 5 of the studies included by the Cochrane review provided data that was appropriate to be meta-analysed. These studies reported on the following outcomes of interest:

- Treatment failure (4 studies)
- Relapse (4 studies)
- Adverse events (5 studies)
- Mortality (2 studies)
- Length of hospitalisation (3 studies)
- FEV1 (5 studies)
- PaO_2 (2 studies)
- $PaCO_2$ (1 study)
- Breathlessness (4 studies)
- Quality of life (1 study)

The ≤ 7 day corticosteroid period was recorded as 3 days (2 studies), 5 days (2 studies) or 7 days (1 study) and the > 7 day corticosteroid treatment was either 10 days (2 studies) or 14 days (3 studies).

Further characteristics are presented in Table 2.

Table 2: Summary of included study characteristics

Author	Days on treatment	Number of randomised people	Steroid used	Dose	Study location
Chen (2005)	7 days or 14 days	87	Oral prednisolone	30mg / day	China
Leuppi (2013)	5 days or 14 days	314	Day 1: IV methylprednisolone Day 2 - End: Oral prednisolone	40mg / day	Switzerland
Sayiner (2001)	3 days or 10 days	36	IV methylprednisolone	Day 0-3: 0.5 mg/kg 6-hourly Day 4-6: 0.5 mg/kg 12-hourly Day 6-10: 0.5 mg/kg /day	Turkey
Sirichana (2008)	5 days or 10 days	48	Prednisolone (unspecified delivery method)	30mg / day	Thailand
Wood-Baker (1997)	3 days or 14 days	38	Oral prednisolone	Less than 7 days group: 2.5 mg/kg / day Greater than 7 days group: Day 0-7 0.6 mg/kg Day 7-14 0.3 mg/kg	Australia

Please refer to the evidence tables in the Cochrane review for more details about the included studies.

Quality assessment of clinical studies included in the evidence review

The quality assessments for the 5 studies included in the meta-analysis was based on the judgments of the Cochrane review authors who used the Cochrane Risk of Bias tool. This is the same method as used by NICE for risk of bias assessment and so the Guidelines Updates Team used this information to generate an overall study risk of bias. This is shown in appendix E. For full GRADE tables please see appendix G.

Economic evidence

A systematic search was carried out to cover this review question. The search returned 267 records, all of which were excluded on title and abstract.

Evidence statements

The format of the evidence statements is explained in appendix B. Unless stated, the results presented in the evidence statements are pooled results and are not separated by method of administration. Sub-group analysis results are only presented where there were significant differences between subgroups.

Shorter durations of ≤ 7 days of corticosteroid treatment vs longer treatments of > 7 days

Moderate to high quality evidence from up to 4 RCTs with up to 404 people found no meaningful difference in early or 15 day breathlessness symptoms, 6 day quality of life or 30 day quality of life in people with a COPD exacerbation offered corticosteroid treatment for ≤ 7 days compared to people with a COPD exacerbation offered corticosteroid treatment > 7 days.

Low to moderate quality evidence from up to 5 RCTs with up to 503 people could not differentiate treatment failure, relapse, time to re-exacerbation, adverse event levels, mortality, length of hospitalisation, FEV1, PaO_2, or $PaCO_2$ in people with a COPD exacerbation offered corticosteroid treatment for ≤ 7 days compared to people with a COPD exacerbation offered corticosteroid treatment for > 7 days.

Subgroup analyses

No subgroup differences were identified between studies with people on an oral corticosteroid course studies with people receiving corticosteroids by IV, studies with people receiving corticosteroids by IV followed by an oral corticosteroid course, or studies with people receiving corticosteroids by an unknown mechanism of delivery, apart from FEV1 at the end of treatment.

- Moderate quality evidence from 1 study with 34 people who received corticosteroids by IV showed an improvement in FEV1 at 12 months in people with a COPD exacerbation offered corticosteroid treatment for > 7 days compared to people with a COPD exacerbation offered corticosteroid treatment for ≤ 7 days
- Very low to moderate quality evidence in up to 3 studies with up to 110 patients who were on an oral corticosteroid course or who received corticosteroids by an unknown mechanism of delivery could not differentiate FEV1 in people with a COPD exacerbation offered corticosteroid treatment for

≤ 7 days compared to people with a COPD exacerbation offered corticosteroid treatment for > 7 days.

Sensitivity analyses removing studies at high risk of bias

Sensitivity analyses were carried out to remove studies at high risk of bias. These analyses did not lead to any changes in the interpretation of the evidence.

Recommendations

1. Offer oral prednisolone 30 mg daily for up to 7 days. Be aware that there is no benefit from taking corticosteroids for more than 7 days.

Rationale and impact

Why the committee made the recommendations

There are risks associated with long-term corticosteroid use, so it is important to use the shortest effective treatment duration. The evidence showed no benefit from taking corticosteroids for more than 7 days and shorter courses are routinely used in clinical practice already. Treatment is recommended for 'up to' 7 days because some people will recover from their exacerbation faster than others and may need less than 7 days of treatment. In addition, the trials looked at different durations of short courses (from 3-7 days) compared to a longer course, but due to the small sizes of the trials it was not possible to make a more specific recommendation. The dose of steroid was retained from the recommendation in the 2018 guideline.

Impact of the recommendations on practice

The recommendation may reduce the amount of corticosteroids used in clinical practice, which may result in a cost saving. However, the overall impact is likely to be small because oral corticosteroids are cheap, and because prescribing corticosteroids for 7 days or less is current practice for many clinicians.

The committee's discussion of the evidence

Interpreting the evidence

The outcomes that matter most

The committee agreed that since the corticosteroid use under review was taking place during an exacerbation, the key outcomes for a person with COPD under these circumstances were related to length of hospitalisation, breathlessness, time to re-exacerbation, mortality and quality of life. In addition, treatment failure and relapse were measures of the effectiveness of the treatment and it was important to examine the numbers of people experiencing adverse events associated with corticosteroid use to help determine the benefits of a shorter course of medication. Outcome measures such as FEV1, PaO_2 and $PaCO_2$ could be useful indicators of physiological improvement for the person taking the corticosteroid, but would not be sufficiently important in the absence of improvements in the aforementioned outcomes to make decisions regarding corticosteroid use.

The quality of the evidence

The evidence for the outcomes in this review ranged from very low to high quality, with no UK based studies. All studies were judged to be directly applicable, and one was judged to be at high risk of bias due to participants, investigators and outcome assessors not being blinded (Sirichana et al. 2008). All other studies were judged to be at low risk of bias. However, the sensitivity analysis removing Sirichana 2008 showed no difference in results for any of the outcomes.

Whilst two of the studies in the meta-analysis were published abstracts (Wood-baker 1997 and Sirichana 2008), the authors of the Cochrane review that this review is based upon obtained extra unpublished data from the authors, leading to a more reliable data source than the published abstract alone. However, neither study has been published as a full paper or been through peer review.

In regards to the study population, the committee noted that many of the studies were from lower income countries that may have different demographic characteristics that are less relevant to UK practice. The committee also noted that all of the evidence came from a hospital setting (if a setting was recorded), and expressed concerns of a lack of evidence of steroid use in outpatients, for example in community settings. However, they agreed that the findings remained sufficiently relevant for the UK population in general and decided against downgrading the evidence for indirectness. Further concerns around the quality of the evidence included the age of the data (the studies were carried out between 1997 and 2013), gender imbalance within the study population, and the doses of corticosteroids used. The committee discussed the high percentage of males in these studies, particularly in Sayiner 2001, which may be due to the difference in smoking habits between males and females in the countries the studies took place in. The committee also stated that the prednisolone dose of 2.5mg/kg per day used in Wood-baker (1997) is much higher than UK doses.

Data was only available for one of the subgroups outlined in the review protocol, regarding which mechanism of corticosteroid delivery occurred in each trial. The only subgroup difference was observed in the FEV1 end of treatment results, where the IV group showed an improvement in patients given corticosteroid for > 7 days. The committee agreed this result was not important in regards to recommendations, due to the low patient number in this study and the low relative importance of FEV1 compared to the other outcomes.

Benefits and harms

The aim of this review was to identify whether there was any detectable difference in outcomes between a ≤ 7 day course of corticosteroids and > 7 day course of corticosteroids. However, the committee noted that the use of shorter courses of corticosteroids is already widespread in clinical practice.

For most outcomes, the evidence could not differentiate between longer and shorter courses of corticosteroids. In particular, breathlessness, which was highlighted as one of the important outcomes, and quality of life, showed no meaningful difference between the longer and shorter corticosteroid courses. However, the committee noted that the absence of a meaningful difference did not necessarily mean that the treatment durations were equivalent, particularly as the small sample size of some of the trials might have prevented any differences from being detected. Despite this, based on their clinical experience, and supported by the results of the larger trials (Chen, 2005, and Leuppi, 2013), the committee agreed that it was likely that the effects of courses of ≤ 7 days of steroids of were equivalent to courses of > 7 days.

For breathlessness, the quality of this evidence was high to moderate from 4 studies and 404 patients, with the 95% confidence intervals (CIs) well within the minimal clinically important differences (MIDs), suggesting that there is an absence of clinically meaningful difference for this outcome. For quality of life, the evidence quality was high in one study with up to 290 patients in the intention to treat analysis population. The 95% CIs again were well within the MIDs. It is worth noting that these analyses had MIDs taken from the literature as opposed to taking the line of no effect as a measure of imprecision, which could explain the higher quality of evidence for these outcomes.

Based on these results, the committee agreed that there was no apparent clinical benefit to the longer corticosteroid course compared to the shorter course across all of the outcomes reviewed and wrote a recommendation to use a shorter course of up to 7 days of treatment. They noted that it was important to include the information about the lack of benefit of continued treatment in the recommendation because they wanted to emphasise the importance of stopping treatment at this point rather than starting to wean the person off prednisolone as may currently be common practice for some clinicians. The choice of 'up to 7 days' was based on the clinical trials which reported data for people taking short courses of differing lengths (3-7 days) versus longer courses. Due to the low number of included trials and limited patient numbers, there was not enough data for the committee to recommend a 3 day, 5 day or 7 day course specifically. However, the committee noted that in practice clinicians have been using less than 7 days prednisolone routinely for a few years and the current debate is about using courses of 5 or 7 days. The current recommendation allows clinicians to make a judgement on the exact duration of the course (up to 7 days) based on their experience and the requirements of the person with COPD. The dose was based on the 2018 guideline recommendation, written in 2004, as this review did not examine corticosteroid doses.

The committee stated that if there is no positive effect associated with a longer corticosteroid course the shorter course should be recommended to reduce the risk of corticosteroid side effects, including fluid retention, pneumonia, hypertension, diabetes mellitus, adrenal suppression and osteoporosis. The committee noted that while the evidence could not differentiate between the two courses for adverse events, long term corticosteroid use over time with repeated courses would likely increase the risk of adverse events. Thus, a shorter course would likely be beneficial over time as the total amount of corticosteroids prescribed and taken annually would be reduced.

Cost effectiveness and resource use

The committee discussed the cost effectiveness of prescribing ≤ 7 days versus > 7 days of corticosteroid treatment for acute exacerbations. They determined that, given the lack of evidence of any additional clinical benefit for treatment past 7 days, the more conservative choice of a shorter treatment duration is likely to be cost effective. Furthermore, the committee highlighted that outcomes included in the clinical review do not capture the potential longer-term consequences of corticosteroid use, such as osteoporosis. Therefore, it is reasonable to expect that treatment for ≤ 7 days is both less costly, and produces equivalent or better health outcomes than treatment for > 7 days.

The committee discussed the potential resource impact of their recommendation, and determined that it may produce a cost saving, due to reduced use of corticosteroids. However, the overall impact is likely to be small, given the low cost of oral corticosteroids, and given that prescribing corticosteroids for ≤ 7 days is current practice for many clinicians.

Other factors the committee took into account

The committee expressed an interest in examining the doses of corticosteroids used in addition to the duration of the courses, but this was outside of the scope of this review question and update. Instead, they retained the dose from the recommendation in the 2018 guideline, which was written in 2004 when the evidence for corticosteroid use was reviewed in detail.

The committee also discussed the importance of clearly informing people whether they are being prescribed 30mg single tablets or multiple tablets of a lower dose (commonly 6 tablets of 5mg). This may help to reduce the risk of an accidental overdose, particularly if a person moves between different care settings where the format of the dose may change. However, the committee felt that this issue was generally applicable to situations other than corticosteroid use in people with COPD and therefore did not require a specific recommendation to be made.

The committee noted that soluble and enteric coated corticosteroid tablets exist and are more expensive than other forms of tablets, but they were unable to recommend any conditions for their use because this area was not within the scope of this review question and they did not examine any evidence regarding the cost and clinical effectiveness of tablets.

Appendices

Appendix A – Review protocols

Review protocol for the duration of corticosteroid use during exacerbations

Field (based on PRISMA-P)	Content
Review question	Are shorter durations of ≤ 7 days of corticosteroid treatment effective at treating acute exacerbations in people with COPD compared to longer treatments of > 7 days?
Type of review question	Intervention
Objective of the review	To determine whether shorter durations (≤ 7 days) of corticosteroid treatment can be used to treat exacerbations in people with COPD instead of the longer treatments (>7 days) that are currently recommended by the NICE COPD guideline CG101 (2010).
Eligibility criteria – population	Inclusion criteria from Cochrane Review: Adults with an acute exacerbation of COPD. (The definition of an acute exacerbation could include any combination of an increase in breathlessness, sputum volume, sputum purulence, cough or wheeze.)
Eligibility criteria – interventions	Systemic corticosteroid (SCS) given for a period of seven or fewer days.
Eligibility criteria – comparators	Systemic corticosteroids given for longer than seven days.
Outcomes	• Treatment failure (for example, the need for additional treatment) • Relapse after treatment (e.g. treatment for new acute exacerbation, re-admission or hospitalisation for COPD) • Adverse drug effects • Mortality • Cardiac complications • Lung function (FEV1)

	• Length of hospital stay • Arterial blood gases • Breathlessness • Quality of life • Resource use and costs For continuous outcomes: early response was measured on or before day seven of treatment, and end of treatment response measurements were made at the time point equivalent to the end of the longer treatment period.
Eligibility criteria – study design	RCTs
Other exclusion criteria	• Studies that included patients with asthma and other lung diseases (e.g. interstitial lung disease, bronchiectasis), unless separate data on participants with COPD alone were available. • Studies in which participants received assisted ventilation (invasive or non-invasive).
Proposed sensitivity/sub-group analysis, or meta-regression	Subgroups: • Inpatient versus outpatient • Studies that included participants previously treated with corticosteroids (inhaled and systemic) • Oral versus IV administration
Selection process – duplicate screening/selection/analysis	10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. If meaningful disagreements were found between the different reviewers, a further 10% of the abstracts were reviewed by two reviewers, with this process continued until agreement is achieved between the two reviewers. From this point, the remaining abstracts will be screened by a single reviewer.

	This review made use of the priority screening functionality with the EPPI-reviewer systematic reviewing software. See Appendix B for more details.
Data management (software)	See Appendix B
Information sources – databases and dates	See Appendix C
Identify if an update	Partial update of 2004 COPD guideline question: Are oral steroids useful / effective in the treatment of patients with an exacerbation of COPD? (2004)
Author contacts	Guideline update
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual
Search strategy – for one database	For details please see appendix C
Data collection process – forms/duplicate	A standardised evidence table format will be used, and published as appendix E (clinical evidence tables) or I (economic evidence tables).
Data items – define all variables to be collected	For details please see evidence tables in appendix E (clinical evidence tables) or I (economic evidence tables).
Methods for assessing bias at outcome/study level	See Appendix B
Criteria for quantitative synthesis	See Appendix B
Methods for quantitative analysis – combining studies and exploring (in)consistency	See Appendix B
Meta-bias assessment – publication bias, selective reporting bias	See Appendix B

Confidence in cumulative evidence	See Appendix B
Rationale/context – what is known	For details please see the introduction to the evidence review.
Describe contributions of authors and guarantor	A multidisciplinary committee developed the evidence review. The committee was convened by the NICE Guideline Updates Team and chaired by Andrew Molyneux in line with section 3 of Developing NICE guidelines: the manual. Staff from the NICE Guideline Updates Team undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details please see Developing NICE guidelines: the manual.
Sources of funding/support	The NICE Guideline Updates Team is an internal team within NICE.
Name of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
Roles of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
PROSPERO registration number	N/A

1

2

Appendix B – Methods

Incorporating published systematic reviews

For all review questions where a literature search was undertaken looking for a particular study design, systematic reviews containing studies of that design were also included. All included studies from those systematic reviews were screened to identify any additional relevant primary studies not found as part of the initial search.

Quality assessment

Individual systematic reviews were quality assessed using the ROBIS tool, with each classified into one of the following three groups:

- High quality – It is unlikely that additional relevant and important data would be identified from primary studies compared to that reported in the review, and unlikely that any relevant and important studies have been missed by the review.
- Moderate quality – It is possible that additional relevant and important data would be identified from primary studies compared to that reported in the review, but unlikely that any relevant and important studies have been missed by the review.
- Low quality – It is possible that relevant and important studies have been missed by the review.

Each individual systematic review was also classified into one of three groups for its applicability as a source of data, based on how closely the review matches the specified review protocol in the guideline. Studies were rated as follows:

- Fully applicable – The identified review fully covers the review protocol in the guideline.
- Partially applicable – The identified review fully covers a discrete subsection of the review protocol in the guideline (for example, some of the factors in the protocol only).
- Not applicable – The identified review, despite including studies relevant to the review question, does not fully cover any discrete subsection of the review protocol in the guideline.

Using systematic reviews as a source of data

If systematic reviews were identified as being sufficiently applicable and high quality, and were identified sufficiently early in the review process (for example, from the surveillance review or early in the database search), they were used as the primary source of data, rather than extracting information from primary studies. The extent to which this was done depended on the quality and applicability of the review, as defined in Table 3. When systematic reviews were used as a source of primary data, and unpublished or additional data included in the review which is not in the primary studies was also included. Data from these systematic reviews was then quality assessed and presented in GRADE/CERQual tables as described below, in the same way as if data had been extracted from primary studies. In questions where data was extracted from both systematic reviews and primary studies, these were cross-referenced to ensure none of the data had been double counted through this process.

Table 3: Criteria for using systematic reviews as a source of data

Quality	Applicability	Use of systematic review
High	Fully applicable	Data from the published systematic review were used instead of undertaking a new literature search or data analysis. Searches were only done to cover the period of time since the search date of the review.
High	Partially applicable	Data from the published systematic review were used instead of undertaking a new literature search and data analysis for the relevant subsection of the protocol. For this section, searches were only done to cover the period of time since the search date of the review. For other sections not covered by the systematic review, searches were undertaken as normal.
Moderate	Fully applicable	Details of included studies were used instead of undertaking a new literature search. Full-text papers of included studies were still retrieved for the purposes of data analysis. Searches were only done to cover the period of time since the search date of the review.
Moderate	Partially applicable	Details of included studies were used instead of undertaking a new literature search for the relevant subsection of the protocol. For this section, searches were only done to cover the period of time since the search date of the review. For other sections not covered by the systematic review, searches were undertaken as normal.

Evidence synthesis and meta-analyses

Where possible, meta-analyses were conducted to combine the results of quantitative studies for each outcome. For continuous outcomes analysed as mean differences, where change from baseline data were reported in the trials and were accompanied by a measure of spread (for example standard deviation), these were extracted and used in the meta-analysis. Where measures of spread for change from baseline values were not reported, the corresponding values at study end were used and were combined with change from baseline values to produce summary estimates of effect. These studies were assessed to ensure that baseline values were balanced across the treatment groups; if there were significant differences at baseline these studies were not included in any meta-analysis and were reported separately. For continuous outcomes analysed as standardised mean differences, where only baseline and final time point values were available, change from baseline standard deviations were estimated, assuming a correlation coefficient of 0.5.

Evidence of effectiveness of interventions

Quality assessment

Individual RCTs and quasi-randomised controlled trials were quality assessed using the Cochrane Risk of Bias Tool. Each individual study was classified into one of the following three groups:

- Low risk of bias – The true effect size for the study is likely to be close to the estimated effect size.
- Moderate risk of bias – There is a possibility the true effect size for the study is substantially different to the estimated effect size.

- High risk of bias – It is likely the true effect size for the study is substantially different to the estimated effect size.

Each individual study was also classified into one of three groups for directness, based on if there were concerns about the population, intervention, comparator and/or outcomes in the study and how directly these variables could address the specified review question. Studies were rated as follows:

- Direct – No important deviations from the protocol in population, intervention, comparator and/or outcomes.
- Partially indirect – Important deviations from the protocol in one of the population, intervention, comparator and/or outcomes.
- Indirect – Important deviations from the protocol in at least two of the following areas: population, intervention, comparator and/or outcomes.

Methods for combining intervention evidence

Meta-analyses of interventional data were conducted with reference to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al. 2011).

Where different studies presented continuous data measuring the same outcome but using different numerical scales (e.g. a 0-10 and a 0-100 visual analogue scale), these outcomes were all converted to the same scale before meta-analysis was conducted on the mean differences. Where outcomes measured the same underlying construct but used different instruments/metrics, data were analysed using standardised mean differences (Hedges' g).

A pooled relative risk was calculated for dichotomous outcomes (using the Mantel–Haenszel method) reporting numbers of people having an event, and a pooled incidence rate ratio was calculated for dichotomous outcomes reporting total numbers of events. Both relative and absolute risks were presented, with absolute risks calculated by applying the relative risk to the pooled risk in the comparator arm of the meta-analysis (all pooled trials).

Fixed- and random-effects models (der Simonian and Laird) were used, with the choice of model based on the degree of heterogeneity for the results of each outcome. Fixed-effects models were the preferred choice, but in situations where the assumptions of a shared mean for fixed-effects model were clearly not met, random-effects results were presented. Random-effects models were selected for analysis if significant statistical heterogeneity was identified in the meta-analysis, defined as $I^2 \geq 50\%$.

In any meta-analyses where some (but not all) of the data came from studies at high risk of bias, a sensitivity analysis was conducted, excluding those studies from the analysis. Results from both the full and restricted meta-analyses are reported. Similarly, in any meta-analyses where some (but not all) of the data came from indirect studies, a sensitivity analysis was conducted, excluding those studies from the analysis.

Meta-analyses were performed in Cochrane Review Manager V5.3, with the exception of incidence rate ratio analyses which were carried out in R version 3.3.4.

Minimal clinically important differences (MIDs)

The Core Outcome Measures in Effectiveness Trials (COMET) database was searched to identify published minimal clinically important difference thresholds relevant to this guideline. Identified MIDs were assessed to ensure they had been developed and validated in a

methodologically rigorous way, and were applicable to the populations, interventions and outcomes specified in this guideline. In addition, the Guideline Committee were asked to prospectively specify any outcomes where they felt a consensus MID could be defined from their experience. In particular, any questions looking to evaluate non-inferiority (that one treatment is not meaningfully worse than another) required an MID to be defined to act as a non-inferiority margin.

MIDs found through this process and used to assess imprecision in the guideline are given in Table 4. For other continuous outcomes not specified in the table below, no MID was defined and the line of no effect was used instead.

Table 4: Identified MIDs

Outcome	MID	Source
Borg dyspnoea (breathlessness) score	2 units (-2, +2)	Ries AL. Minimally clinically important difference for the UCSD shortness of breath questionnaire, Borg Scale, and Visual Analog Scale. J COPD 2005; 2: 105–110.
Change in FEV1	0.1 L (-0.1, +0.1)	Cazzola M, MacNee W, Martinez M et al. Outcomes for COPD pharmacological trials: from lung function to biomarkers. Eur Respir J 2008; 31: 416–468.
Acute bronchitis health-related quality of life interview	0.5 units (-0.5, +0.5)	Evans AT, Husain S, Durairaj L, Sadowski LS, et al. Azithromycin for acute bronchitis: A randomised, double-blind controlled trial. The Lancet 2002, 359(9318), 1648-54.

The committee specified that any difference in mortality would be clinically meaningful, and therefore the line of no effect was used as an MID. For relative risks where no other MID was available, the GRADE default MID interval for dichotomous outcomes of 0.8 to 1.25 was used. Incidence rate ratios were treated in the same way as relative risks, with a default MID interval of 0.8 and 1.25 used for analysis.

GRADE for pairwise meta-analyses of interventional evidence

GRADE was used to assess the quality of evidence for the selected outcomes as specified in 'Developing NICE guidelines: the manual (2014)'. Data from all study designs was initially rated as high quality and the quality of the evidence for each outcome was downgraded or not from this initial point, based on the criteria given in Table 5.

Table 5: Rationale for downgrading quality of evidence for intervention studies

GRADE criteria	Reasons for downgrading quality
Risk of bias	Not serious: If less than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from studies at high risk of bias, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between studies at high and low risk of bias.

GRADE criteria	Reasons for downgrading quality
Indirectness	Not serious: If less than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between direct and indirect studies.
Inconsistency	Concerns about inconsistency of effects across studies, occurring when there is unexplained variability in the treatment effect demonstrated across studies (heterogeneity), after appropriate pre-specified subgroup analyses have been conducted. This was assessed using the I^2 statistic. N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study. Not serious: If the I^2 was less than 33.3%, the outcome was not downgraded. Serious: If the I^2 was between 33.3% and 66.7%, the outcome was downgraded one level. Very serious: If the I^2 was greater than 66.7%, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between studies with the smallest and largest effect sizes.
Imprecision	If an MID other than the line of no effect was defined for the outcome, the outcome was downgraded once if the 95% confidence interval for the effect size crossed one line of the MID, and twice if it crosses both lines of the MID. If the line of no effect was defined as an MID for the outcome, it was downgraded once if the 95% confidence interval for the effect size crossed the line of no effect (i.e. the outcome was not statistically significant), and twice if the sample size of the study was sufficiently small that it is not plausible any realistic effect size could have been detected. Outcomes meeting the criteria for downgrading above were not downgraded if the confidence interval was sufficiently narrow that the upper and lower bounds would correspond to clinically equivalent scenarios.

The quality of evidence for each outcome was upgraded if any of the following three conditions were met:

- Data from non-randomised studies showing an effect size sufficiently large that it cannot be explained by confounding alone.
- Data showing a dose-response gradient.
- Data where all plausible residual confounding is likely to increase our confidence in the effect estimate.

Publication bias

Publication bias was assessed in two ways. First, if evidence of conducted but unpublished studies was identified during the review (e.g. conference abstracts, trial protocols or trial records without accompanying published data), available information on these unpublished studies was reported as part of the review. Secondly, where 10 or more studies were

included as part of a single meta-analysis, a funnel plot was produced to graphically assess the potential for publication bias.

Evidence statements

Evidence statements for pairwise intervention data are classified in to one of four categories:

For outcomes with a defined MID, evidence statements were divided into 4 groups as follows:

- Situations where the data are only consistent, at a 95% confidence level, with an effect in one direction (i.e. one that is 'statistically significant'), and the magnitude of that effect is most likely to meet or exceed the MID (i.e. the point estimate is not in the zone of equivalence). In such cases, we state that the evidence showed that there is an effect.
- Situations where the data are only consistent, at a 95% confidence level, with an effect in one direction (i.e. one that is 'statistically significant'), but the magnitude of that effect is most likely to be less than the MID (i.e. the point estimate is in the zone of equivalence). In such cases, we state that the evidence showed there is an effect, but it is less than the defined MID.
- Situations where the confidence limits are smaller than the MIDs in both directions. In such cases, we state that the evidence demonstrates that there is no meaningful difference.
- In all other cases, we state that the evidence could not differentiate between the comparators.

For outcomes without a defined MID or where the MID is set as the line of no effect (for example, in the case of mortality), evidence statements are divided into 2 groups as follows:

- We state that the evidence showed that there is an effect if the 95% CI does not cross the line of no effect.
- The evidence could not differentiate between comparators if the 95% CI crosses the line of no effect.

Appendix C – Literature search strategies

Clinical literature search

The clinical literature search was undertaken by Cochrane, and outlined in full in the 2018 review. The approach comprises a search to populate the Cochrane Airways Trial Register, and additional searches of MEDLINE, CENTRAL and Embase. The MEDLINE search for this review is presented below.

1 COPD[MeSH Terms]

2 "adrenal cortex hormone*"

3 steroid

4 steroids

5 glucocorticoid*

6 corticoid*

7 corticosteroid*

8 beclomethasone

9 betamethasone

10 fluticasone

11 cortisone

12 dexamethasone

13 hydrocortisone

14 prednisolone

15 prednisone

16 methylprednisolone

17 methylprednisone

18 triamcinolone

#19 (#2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18)

20 randomised controlled trial [pt]

21 controlled clinical trial [pt]

22 randomised [tiab]

23 placebo [tiab]

24 clinical trials as topic [mesh: noexp]

25 randomly [tiab]

26 trial [ti]

27 (#20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26)

28 (animals [mh] NOT humans [mh])

29 (#27 NOT #28)

30 ("2012/01/01"[Date - Publication] : "3000"[Date - Publication])

31 (#1 AND #19 AND #29 AND #30)

Health economics literature search

Economics sources	Date searched	No. of results
MEDLINE (Ovid)	4th Oct 2018	96
MEDLINE in Process (Ovid)	4th Oct 2018	6
Embase (Ovid)	4th Oct 2018	212
EconLit (Ovid)	4th Oct 2018	0
NHS Economic Evaluation Database (NHS EED) (legacy database)	8th Oct 2018	37
Health Technology Assessment (HTA Database)	8th Oct 2018	12

The MEDLINE search strategy is presented below. This was translated for use in all of the other databases listed. The aim of the search was to identify evidence for the question being asked. Health Economics and Quality of Life filters were used to identify the evidence.

1 lung diseases, obstructive/
2 exp pulmonary disease, chronic obstructive/
3 (copd or coad or cobd or aecb).tw.
4 emphysema*.tw.
5 (chronic* adj4 bronch*).tw.
6 (chronic* adj3 (airflow* or airway* or bronch* or lung* or respirat* or pulmonary) adj3 obstruct*).tw.
7 (pulmonum adj4 (volumen or pneumatosis)).tw.
8 pneumonectasia.tw.
9 *Dyspnea/

10 (chronic* adj3 (breath* or respirat*) adj3 (difficult* or labor* or labour* or problem* or short*)).tw.
11 (chronic* adj3 (dyspnea* or dyspnoea* or dyspneic or breathless*)).tw.
12 or/1-11
13 exp Adrenal Cortex Hormones/
14 "adrenal cortex hormone*".tw.
15 Steroids/
16 steroid*.tw.
17 glucocorticoid*.tw.
18 cortico*.tw.
19 (beclomethasone* or beclometasone*).tw.
20 betamethasone*.tw.
21 exp Fluticasone/
22 fluticasone*.tw.
23 Cortisone/
24 cortisone*.tw.
25 deflazacort*.tw.
26 calcort*.tw.
27 dexamethasone*.tw.
28 glensoludex*.tw.
29 dexsol*.tw.
30 martapan*.tw.
31 exp Hydrocortisone/
32 hydrocortisone*.tw.
33 prednisolone*.tw.
34 pevanti*.tw.
35 prednisone*.tw.
36 deltacortril*.tw.
37 dilacort*.tw.
38 methylprednis*.tw.
39 medrone*.tw.
40 triamcinolone*.tw.
41 Pregnenediones/
42 (pregnenedi*).tw.
43 sterapred*.tw.
44 or/13-43
45 (short* adj3 (duration* or course or treatment* or therapy*)).tw.
46 (("7" or "6" or "5" or "4" or "3" or "2" or "1" or seven or six or five or four of three or two or one) adj3 day*).tw.
47 ("1 week" or "one week").tw.
48 or/45-47
49 44 and 48
50 12 and 49

The MEDLINE economic evaluations and quality of life search filters are presented below. They were translated for use in MEDLINE in Process and Embase databases.

<u>Economic evaluations</u>
1. Economics/
2. exp "Costs and Cost Analysis"/

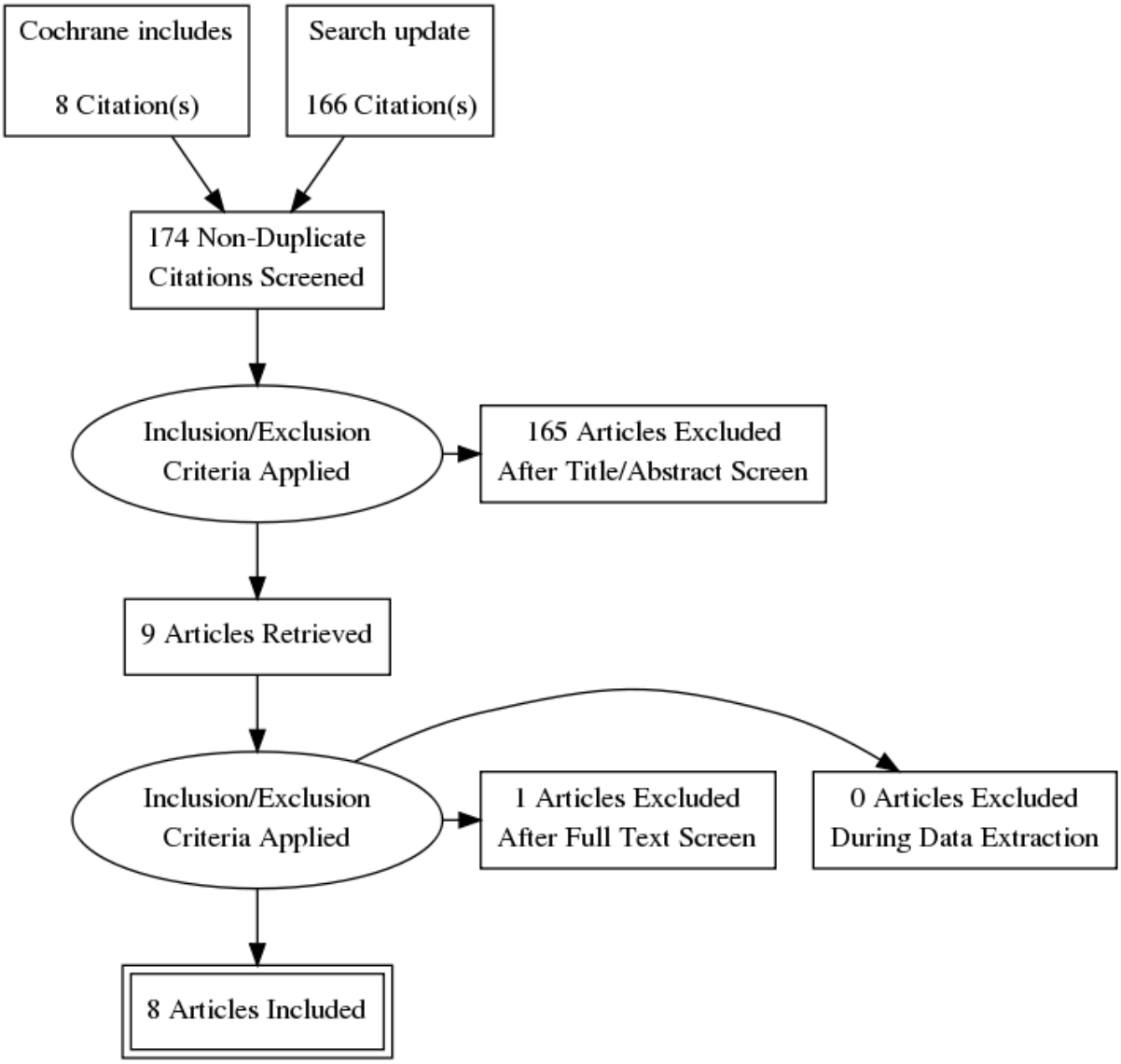
3. Economics, Dental/
4. exp Economics, Hospital/
5. exp Economics, Medical/
6. Economics, Nursing/
7. Economics, Pharmaceutical/
8. Budgets/
9. exp Models, Economic/
10. Markov Chains/
11. Monte Carlo Method/
12. Decision Trees/
13. econom$.tw.
14. cba.tw.
15. cea.tw.
16. cua.tw.
17. markov$.tw.
18. (monte adj carlo).tw.
19. (decision adj3 (tree$ or analys$)).tw.
20. (cost or costs or costing$ or costly or costed).tw.
21. (price$ or pricing$).tw.
22. budget$.tw.
23. expenditure$.tw.
24. (value adj3 (money or monetary)).tw.
25. (pharmacoeconomic$ or (pharmaco adj economic$)).tw.
26. or/1-25

Quality of Life

1. "Quality of Life"/
2. quality of life.tw.
3. "Value of Life"/
4. Quality-Adjusted Life Years/
5. quality adjusted life.tw.
6. (qaly$ or qald$ or qale$ or qtime$).tw.
7. disability adjusted life.tw.
8. daly$.tw.
9. Health Status Indicators/
10. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw.
11. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.
12. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw.
13. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.
14. (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw.
15. (euroqol or euro qol or eq5d or eq 5d).tw.
16. (qol or hql or hqol or hrqol).tw.
17. (hye or hyes).tw.
18. health$ year$ equivalent$.tw.
19. utilit$.tw.
20. (hui or hui1 or hui2 or hui3).tw.

21. disutili$.tw.
22. rosser.tw.
23. quality of wellbeing.tw.
24. quality of well-being.tw.
25. qwb.tw.
26. willingness to pay.tw.
27. standard gamble$.tw.
28. time trade off.tw.
29. time tradeoff.tw.
30. tto.tw.
31. or/1-30

Appendix D – Clinical evidence study selection



Appendix E – Clinical evidence tables

Please refer to the Cochrane review (Walters, 2018) for full evidence tables and the judgements for risk of bias. The overall risk of bias and directness in Table 6 was determined by the Guideline Updates Team based on the Cochrane review tables.

Table 6: Overall study risk of bias and reason for judgement

Author	Risk of Bias*	Reason	Directness
Chen 2005	Low	All risks low bar reporting bias, which was unclear	Directly applicable
Leuppi 2013	Low	All risks low	Directly applicable
Sayiner 2001	Low	All risks low	Directly applicable
Sirichana 2008	High	Participants, investigators and outcome assessors were not blinded. Of 6/25 participants in one group not completing study, 1 participant withdrew and no reason given for other 5.	Directly applicable
Wood-baker 1997	Low	All risks low bar allocation concealment bias, which was unclear	Directly applicable
*Risk of bias in the Cochrane review was scored for 5 types of bias (selection, detection, performance, attrition and reporting). Here all risks of bias have been combined into one final score based on the number of risks and a judgement of the importance of each risk for this review question.			

Appendix F – Forest plots

The following plots were based on data from the Cochrane review. However, the dichotomous data plots have been altered to show RR, not OR, and the choice of fixed effect or random effects model is made according to the methods in appendix B. The sensitivity analyses were carried out by NICE Guideline Updates Team using data from the Cochrane review.

Figure 1: Treatment failure

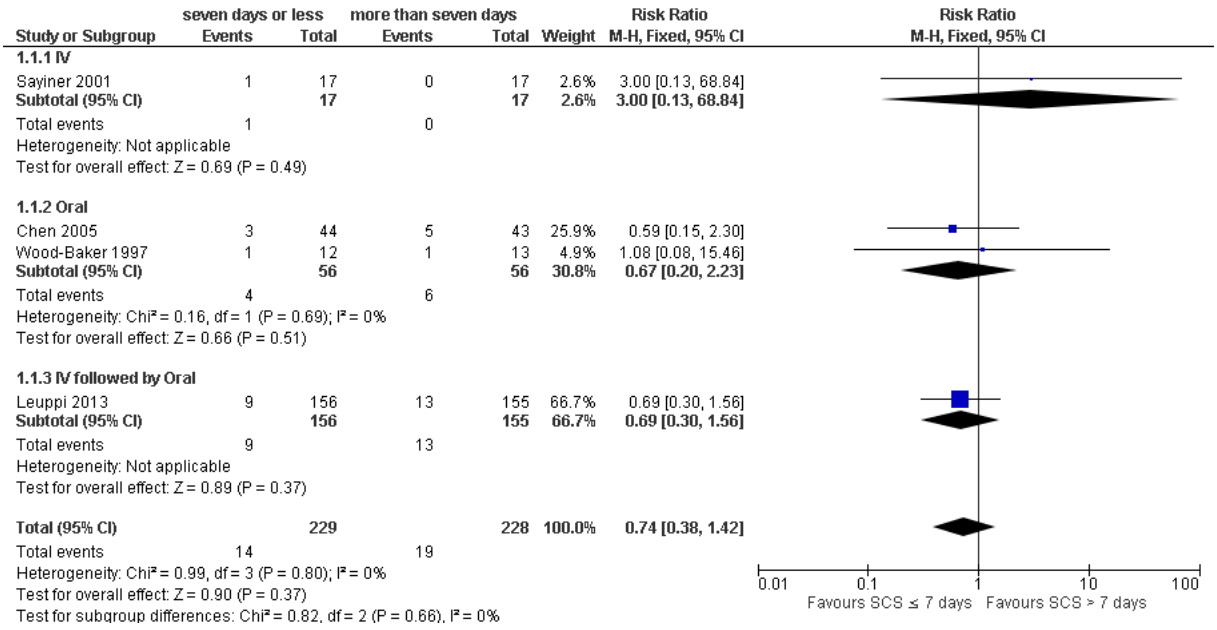


Figure 2: Relapse

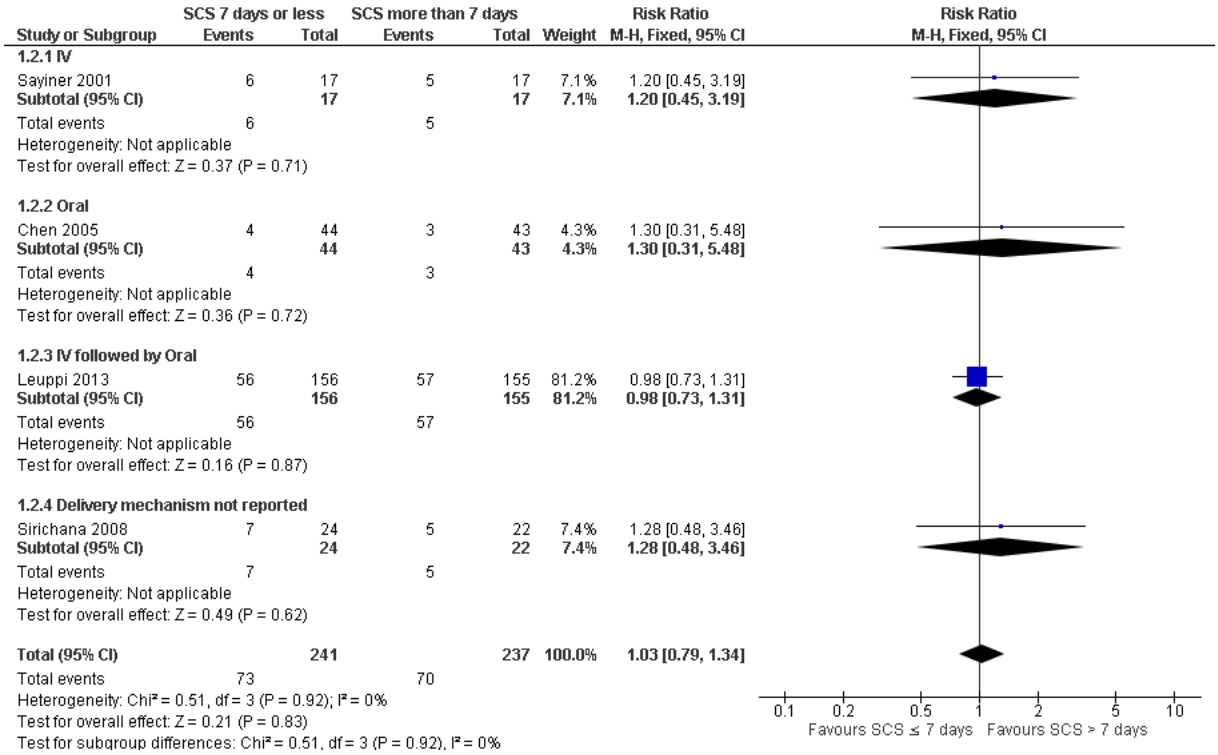


Figure 3: Sensitivity analysis: Removing studies at high risk of bias - Relapse

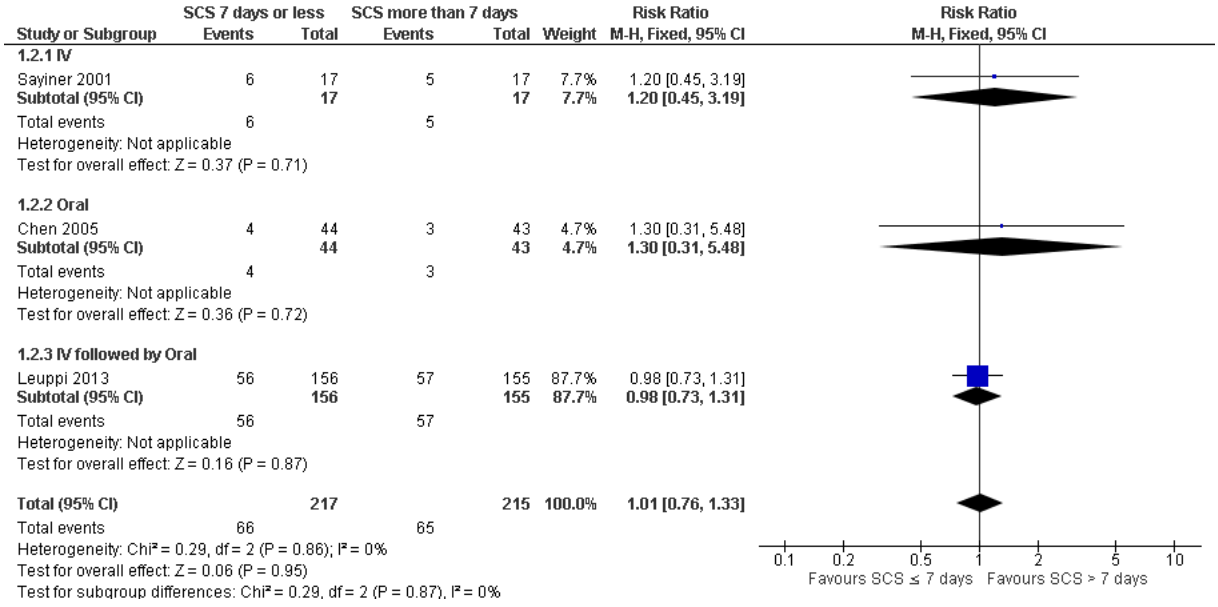


Figure 4: Time to re-exacerbation

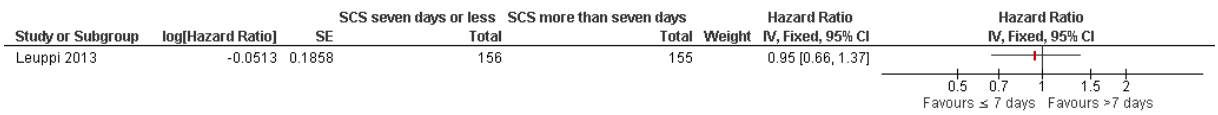


Figure 5: Adverse events – hyperglycaemia

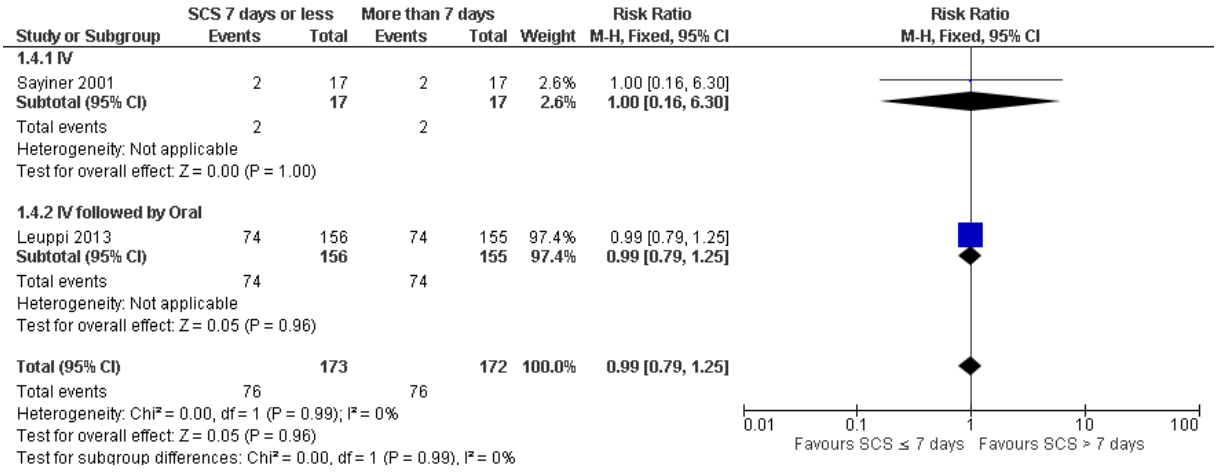


Figure 6: Adverse events – hypertension

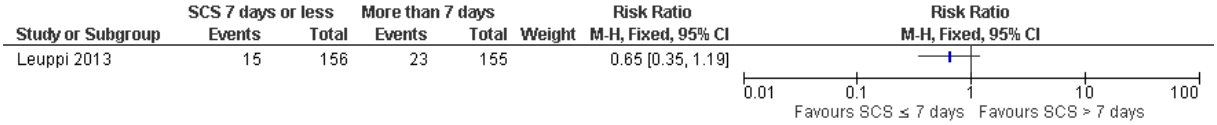


Figure 7: Other adverse events – gastrointestinal tract bleeding, symptomatic gastrointestinal reflux, symptoms of congenital heart failure or ischaemic heart disease, sleep disturbance, fractures, depression

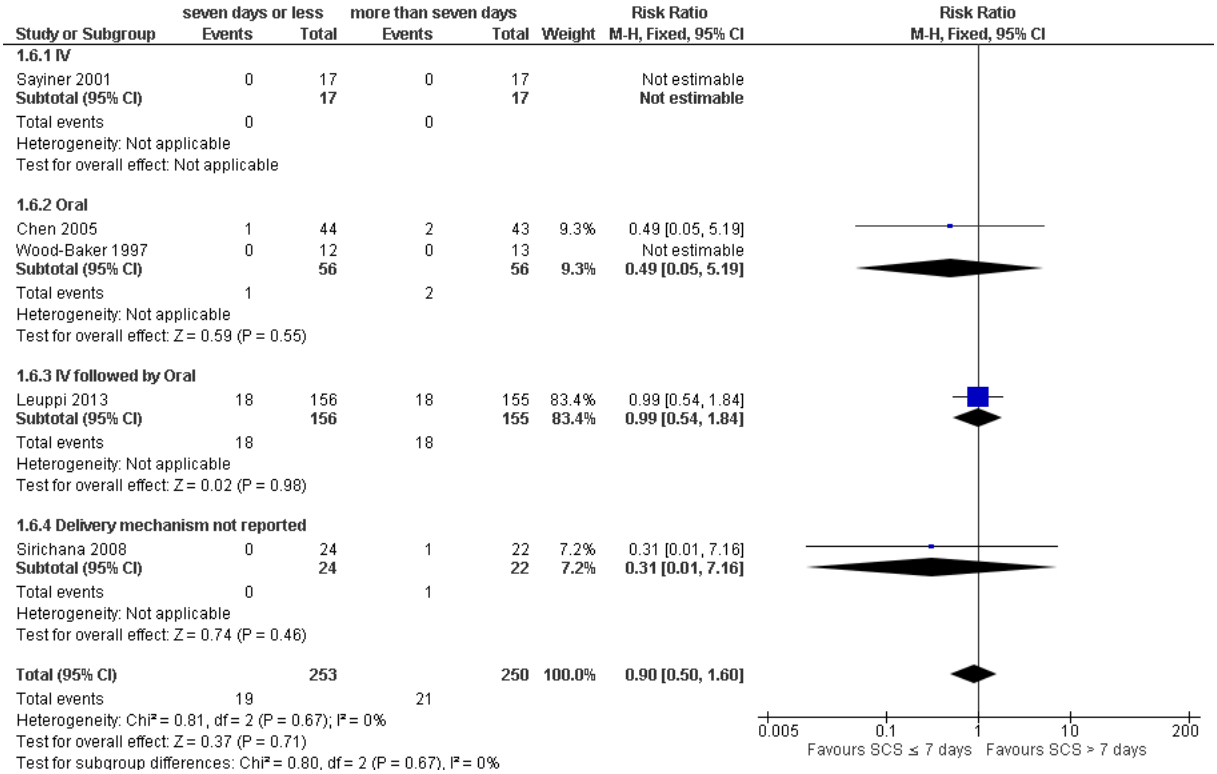


Figure 8: Sensitivity analysis: Removing studies at high risk of bias - Other adverse events – gastrointestinal tract bleeding, symptomatic gastrointestinal reflux, symptoms of congenital heart failure or ischaemic heart disease, sleep disturbance, fractures, depression

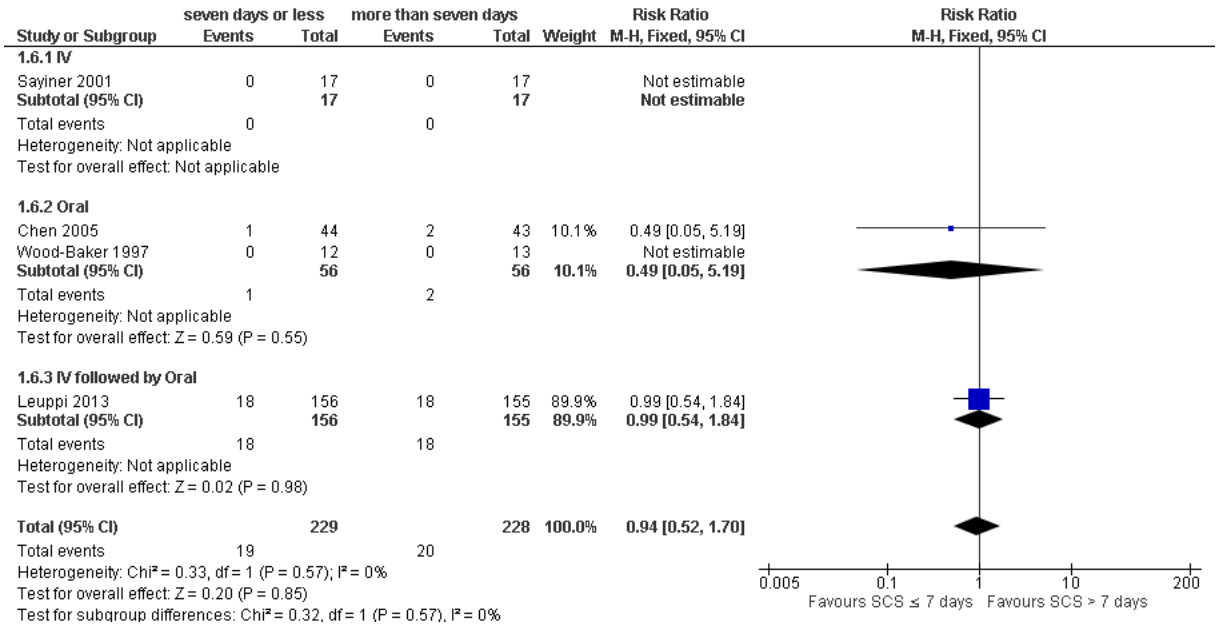


Figure 9: Mortality

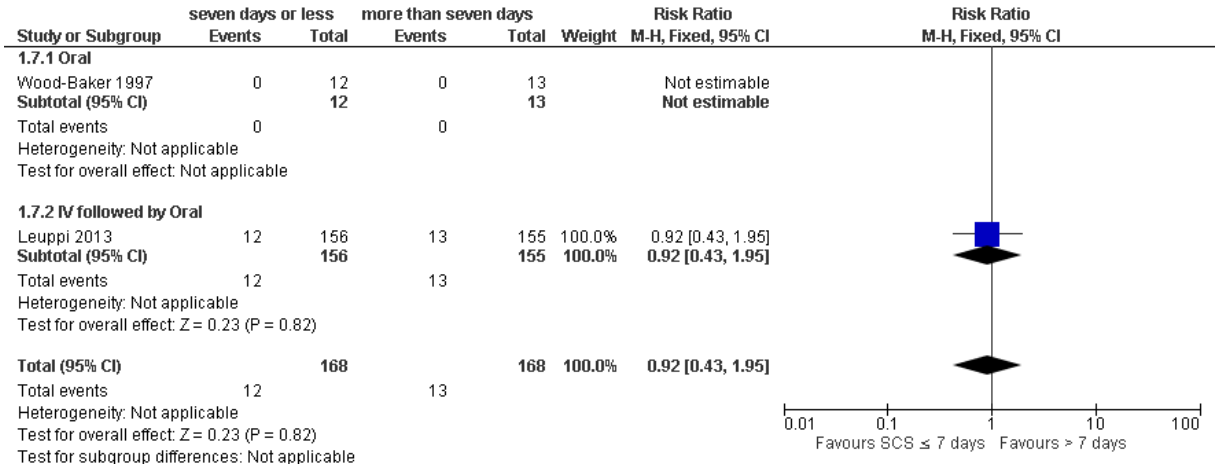


Figure 10: Length of hospitalisation

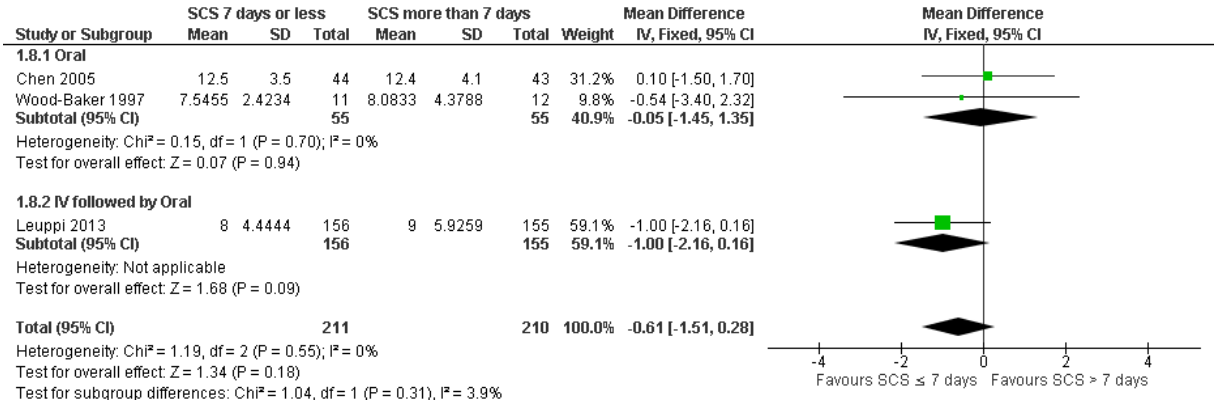


Figure 11: FEV1 (L) (Early)

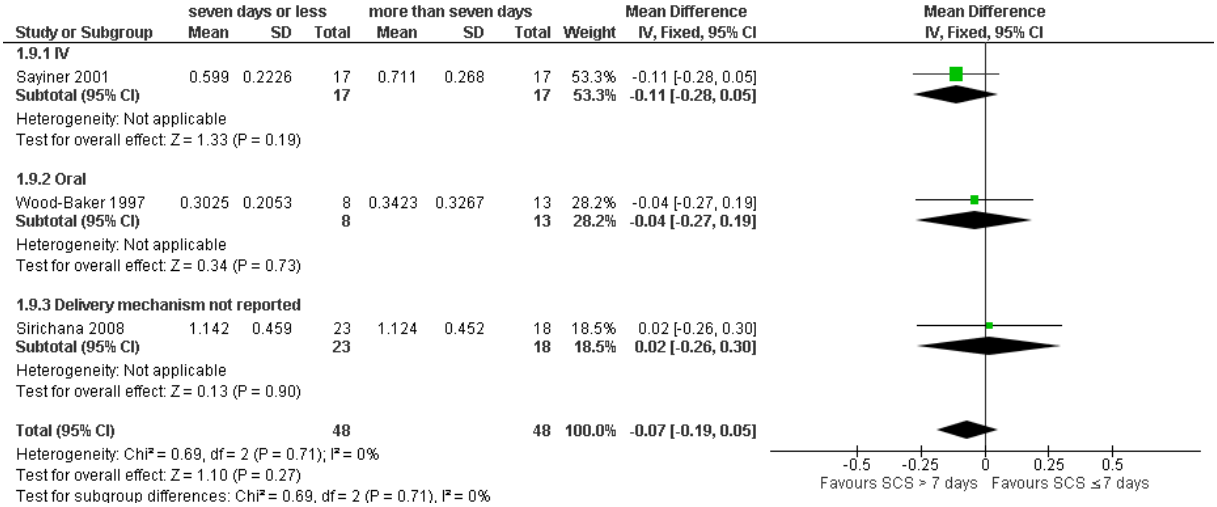


Figure 12: Sensitivity analysis: removing studies at high risk of bias - FEV1 (L) (Early)

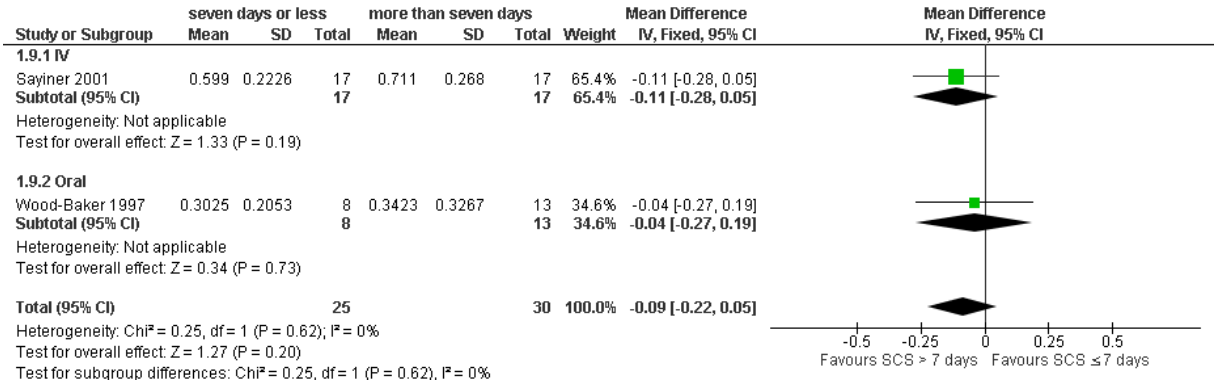


Figure 13: FEV1 % predicted (6 days)

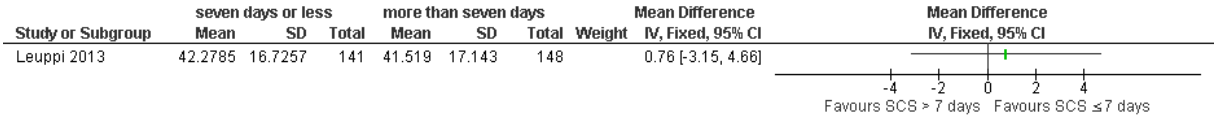


Figure 14: FEV1 (L) (End of treatment)

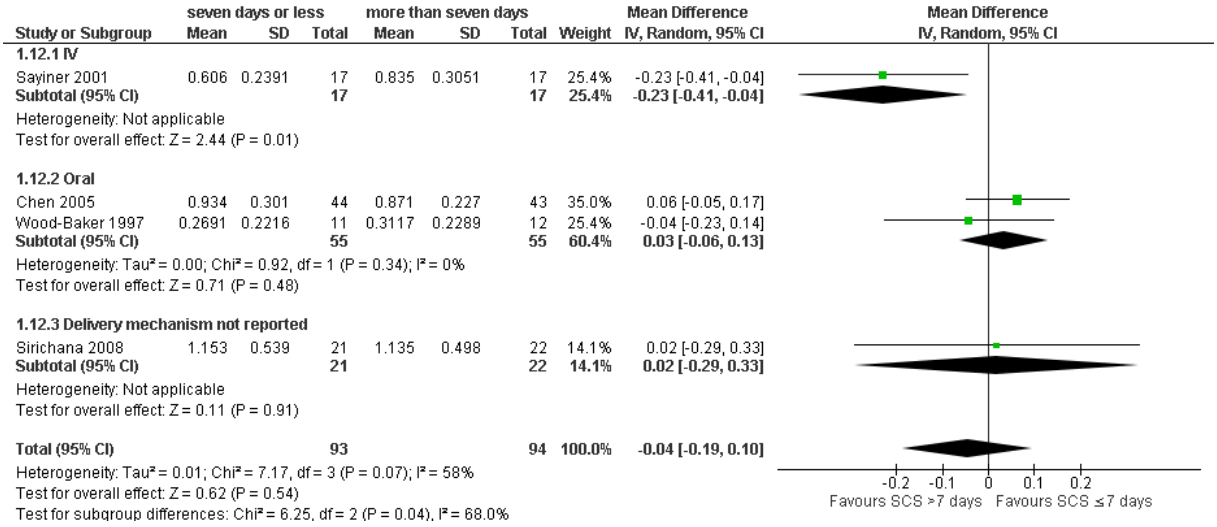


Figure 15: Sensitivity analysis: removing studies at high risk of bias - FEV_1 (L) (End of treatment)

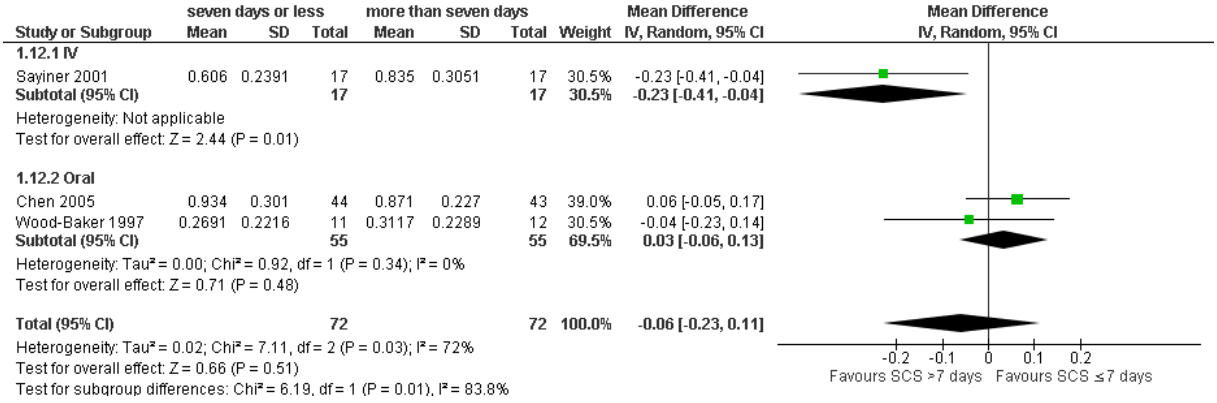


Figure 16: FEV1 % predicted (30 days)

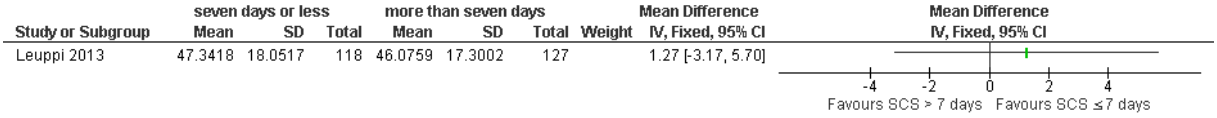


Figure 17: PaO_2 (mmHg) (Early)

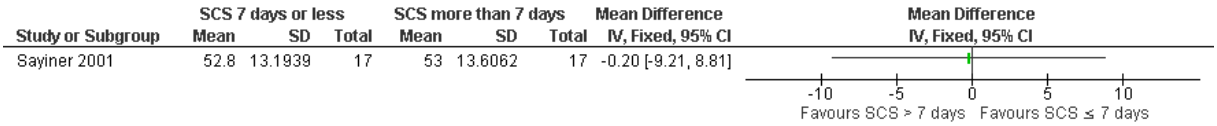


Figure 18: PaO_2 (mmHg) (End of treatment)

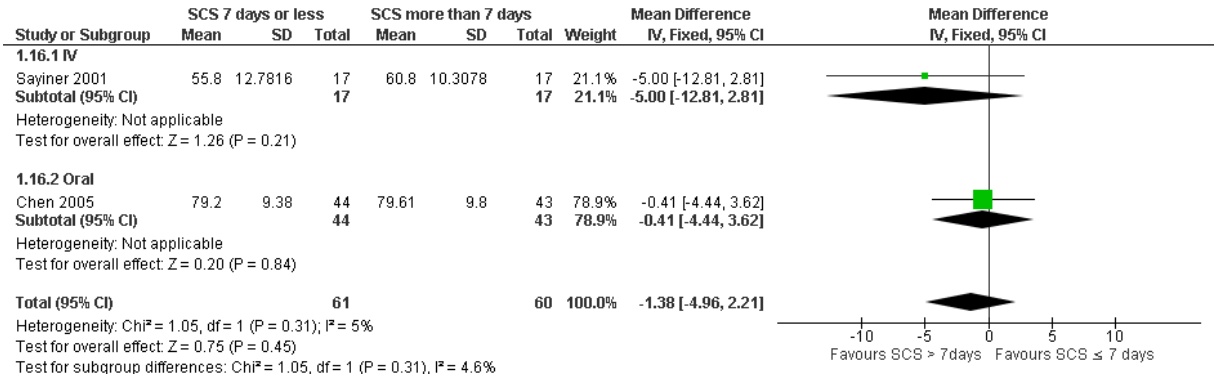


Figure 19: $PaCO_2$ (mmHg)

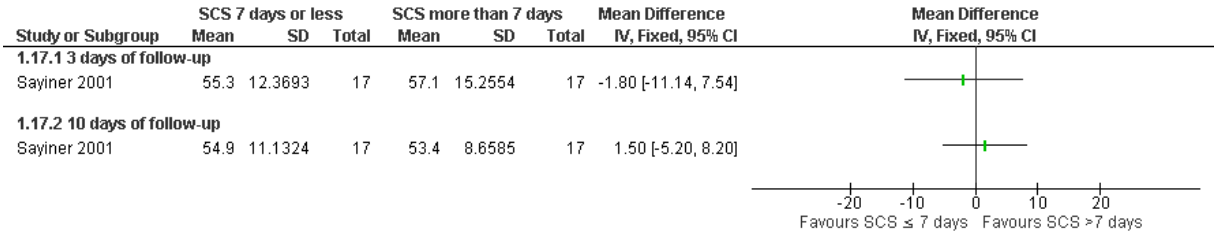


Figure 20: Symptoms – Breathlessness (Early)

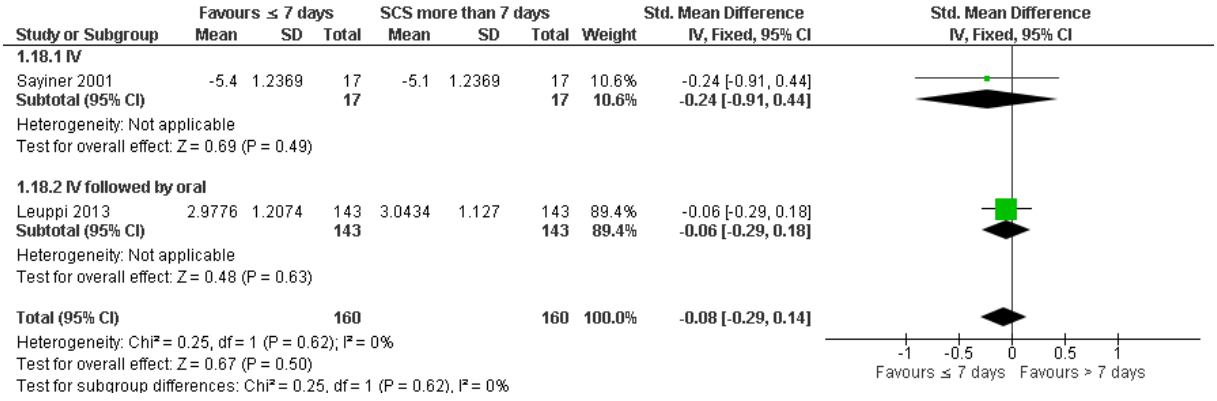


Figure 21: Symptoms – Breathlessness (15 days)

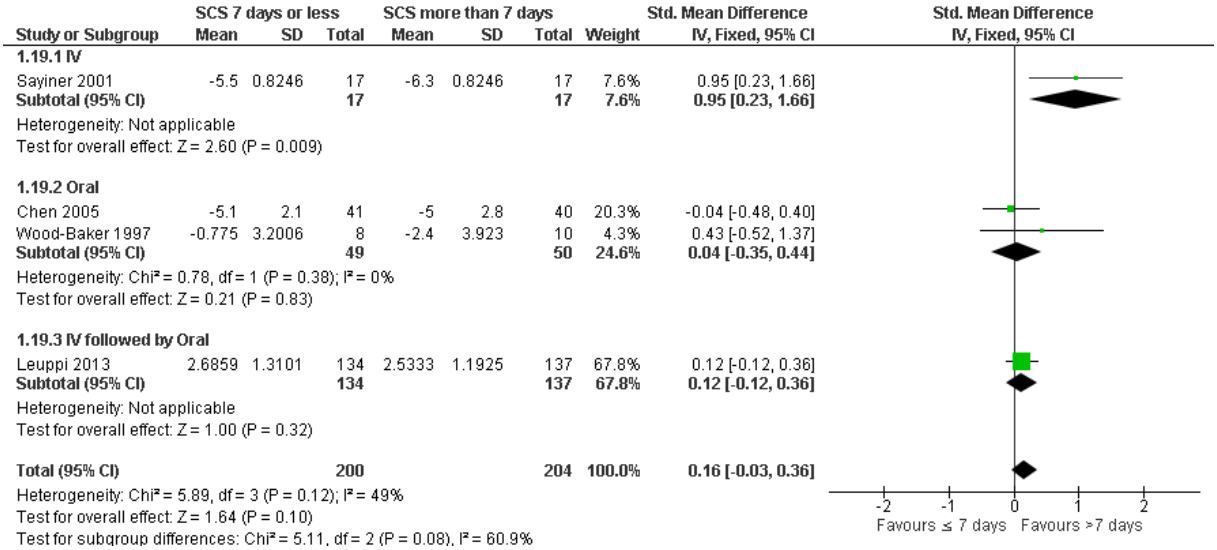


Figure 22: Quality of life - Overall (6 days)

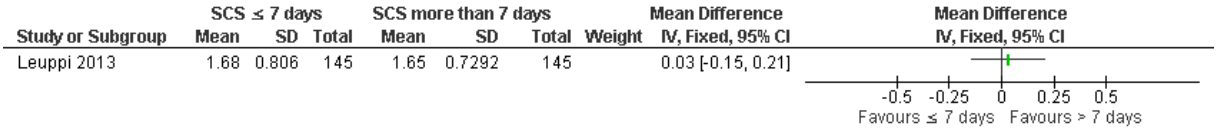
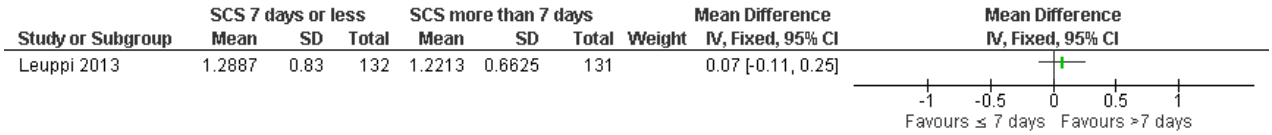


Figure 23: Quality of life - Overall (30 days)



Appendix G – GRADE tables

The following GRADE tables were completed by the NICE Guideline Updates Team tables are based on evidence on effect sizes from the Cochrane review (Walters et al. 2018). However, the dichotomous data has been altered to show RR, not OR, and the choice of fixed effect or random effects model is made according to the methods in appendix B.

Systemic corticosteroids for 7 or fewer days compared to systemic corticosteroids for longer than 7 days

No. of studies	Study design	Sample size	Effect size (95% CI)	Absolute risk: control	Absolute risk: intervention (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
Treatment failure (events) (RR <1 Favours shorter treatment)										
4 studies	RCT	457	RR 0.74 (0.38, 1.42)	8.33 per 100	6.17 per 100 (3.21, 11.87)	Not serious	Not serious	Not serious	Very serious[1]	Low
Relapse (events) (RR <1 Favours shorter treatment)										
4 studies	RCT	478	RR 1.03 (0.79, 1.34)	29.54 per 100	30.39 per 100 (23.27, 39.69)	Not serious	Not serious	Not serious	Very serious[1]	Low
Sensitivity analysis: Removing studies at high risk of bias - Relapse										
3 studies	RCT	432	RR 1.01 (0.76, 1.33)	30.23 per 100	30.49 per 100 (23.11, 40.23)	Not serious	Not serious	Not serious	Very serious[1]	Low
Time to re-exacerbation (events) (HR <1 Favours shorter treatment)										
1 Study (Leuppi 2013)	RCT	311	HR 0.95 (0.66, 1.37)	-	-	Not serious	Not serious	N/A	Serious[2]	Moderate
Adverse events – hyperglycaemia (RR <1 Favours shorter treatment)										
2 studies	RCT	345	RR 0.99 (0.79, 1.25)	44.19 per 100	43.91 per 100 (34.81, 55.39)	Not serious	Not serious	Not serious	Very serious[1]	Low
Adverse events – hypertension (RR <1 Favours shorter treatment)										

No. of studies	Study design	Sample size	Effect size (95% CI)	Absolute risk: control	Absolute risk: intervention (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
1 Study (Leuppi 2013)	RCT	311	RR 0.65 (0.35, 1.19)	14.84 per 100	9.65 per 100 (5.19, 17.66)	Not serious	Not serious	N/A	Serious[3]	Moderate
Other adverse events – gastrointestinal tract bleeding, symptomatic gastrointestinal reflux, symptoms of congenital heart failure or ischaemic heart disease, sleep disturbance, fractures, depression (RR <1 Favours shorter treatment)										
5 studies	RCT	503	RR 0.90 (0.50, 1.60)	8.40 per 100	7.53 per 100 (4.22, 13.44)	Not serious	Not serious	Not serious	Very serious[1]	Low
Sensitivity analysis: Removing studies at high risk of bias - Other adverse events – gastrointestinal tract bleeding, symptomatic gastrointestinal reflux, symptoms of congenital heart failure or ischaemic heart disease, sleep disturbance, fractures, depression										
4 studies	RCT	457	RR 0.94 (0.52, 1.70)	8.77 per 100	8.27 per 100 (4.57, 14.96)	Not serious	Not serious	Not serious	Very serious[1]	Low
Mortality (RR <1 Favours shorter treatment)										
2 studies	RCT	336	RR 0.92 (0.43, 1.95)	7.74 per 100	7.1 per 100 (3.34, 15.06)	Not serious	Not serious	N/A[4]	Serious[2]	Moderate
Length of hospitalisation (MD <0 Favours shorter treatment)										
3 studies	RCT	421	MD -0.61 (-1.51, 0.28)	-	-	Not serious	Not serious	Not serious	Serious[2]	Moderate
FEV1 (L) (Early) (MD <0 Favours shorter treatment)										
3 studies	RCT	96	MD -0.07 (-0.19, 0.05)	-	-	Not serious	Not serious	Not serious	Serious[3]	Moderate
Sensitivity analysis: removing studies at high risk of bias - FEV1 (L) (Early)										
2 studies	RCT	55	MD -0.09 (-0.22, 0.05)	-	-	Not serious	Not serious	Not serious	Serious[3]	Moderate
FEV1 % predicted (6 days) (MD <0 Favours longer treatment)										
1 study (Leuppi 2013)	RCT	289	MD 0.76 (-3.15, 4.66)	-	-	Not serious	Not serious	N/A	Serious[2]	Moderate

No. of studies	Study design	Sample size	Effect size (95% CI)	Absolute risk: control	Absolute risk: intervention (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
FEV1 (L) (End of treatment) Pooled result (MD <0 Favours longer treatment)										
4 studies	RCT	187	MD -0.04 (-0.19, 0.10)	-	-	Not serious	Not serious	Serious[5]	Serious[3]	Low
Subgroup analysis: FEV1 (L) (End of treatment) – IV **(MD <0 Favours longer treatment)**										
1 study (Sayiner 2001)	RCT	34	MD -0.23 (-0.41, -0.04)	-	-	Not serious	Not serious	N/A	Serious[3]	Moderate
Subgroup analysis: FEV1 (L) (End of treatment) – Oral **(MD <0 Favours longer treatment)**										
2 studies	RCT	110	MD 0.03 (-0.06, 0.13)	-	-	Not serious	Not serious	Not serious	Serious[3]	Moderate
Subgroup analysis: FEV1 (L) (End of treatment) - Delivery mechanism not reported **(MD <0 Favours longer treatment)**										
1 study (Sirichana 2008)	RCT	43	MD 0.02 (-0.29, 0.33)	-	-	Very serious[6]	Not serious	N/A	Very serious[1]	Very low
Sensitivity analysis: removing studies at high risk of bias - FEV1 (L) (End of treatment) **(MD <0 Favours longer treatment)**										
3 studies	RCT	144	MD -0.06 (-0.23, 0.11)	-	-	Not serious	Not serious	Very serious[7]	Very serious[1]	Very low
FEV1 % predicted (30 days) (MD <0 Favours longer treatment)										
1 Study (Leuppi 2013)	RCT	245	MD 1.27 (-3.17, 5.70)	-	-	Not serious	Not serious	N/A	Serious[2]	Moderate
PaO_2 (mmHg) (Early) (MD <0 Favours longer treatment)										
1 study (Sayiner 2001)	RCT	34	MD -0.20 (-9.21, 8.81)	-	-	Not serious	Not serious	N/A	Serious[2]	Moderate
PaO_2 (mmHg) (End of treatment) (MD <0 Favours longer treatment)										

No. of studies	Study design	Sample size	Effect size (95% CI)	Absolute risk: control	Absolute risk: intervention (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
2 studies	RCT	121	MD -1.38 (-4.96, 2.21)	-	-	Not serious	Not serious	Not serious	Serious[2]	Moderate
$PaCO_2$ (mmHg) 3 days of follow up (MD <0 Favours shorter treatment)										
1 study (Sayiner 2001)	RCT	34	MD -1.80 (-11.14, 7.54)	-	-	Not serious	Not serious	N/A	Serious[2]	Moderate
$PaCO_2$ (mmHg) 10 days of follow up (MD <0 Favours shorter treatment)										
1 study (Sayiner 2001)	RCT	34	MD 1.50 (-5.20, 8.20)	-	-	Not serious	Not serious	N/A	Serious[2]	Moderate
Symptoms – Breathlessness (Early) (SMD <0 Favours shorter treatment)										
2 studies	RCT	320	*SMD -0.08 (-0.29, 0.14)* MD -0.14 (-0.49, 0.24)*	-	-	Not serious	Not serious	Not serious	Not serious	High
Symptoms – Breathlessness (15 days) (SMD <0 Favours shorter treatment)										
4 studies	RCT	404	*SMD 0.16 (-0.03, 0.36)* MD 0.27 (-0.05, 0.61)*	-	-	Not serious	Not serious	Serious[5]	Not serious	Moderate
Quality of life - Overall (6 days) (MD <0 Favours shorter treatment)**										
1 study (Leuppi 2013)	RCT	290	MD 0.03 (-0.15, 0.21)	-	-	Not serious	Not serious	N/A	Not serious	High
Quality of life - Overall (30 days) (MD <0 Favours shorter treatment)**										
1 study (Leuppi 2013)	RCT	263	MD 0.07 (-0.11, 0.25)	-	-	Not serious	Not serious	N/A	Not serious	High

No. of studies	Study design	Sample size	Effect size (95% CI)	Absolute risk: control	Absolute risk: intervention (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality

* SMD converted to MD on the BORG scale by multiplying by the pooled SD (1.693579) from the studies included in the 15 day breathlessness meta-analysis

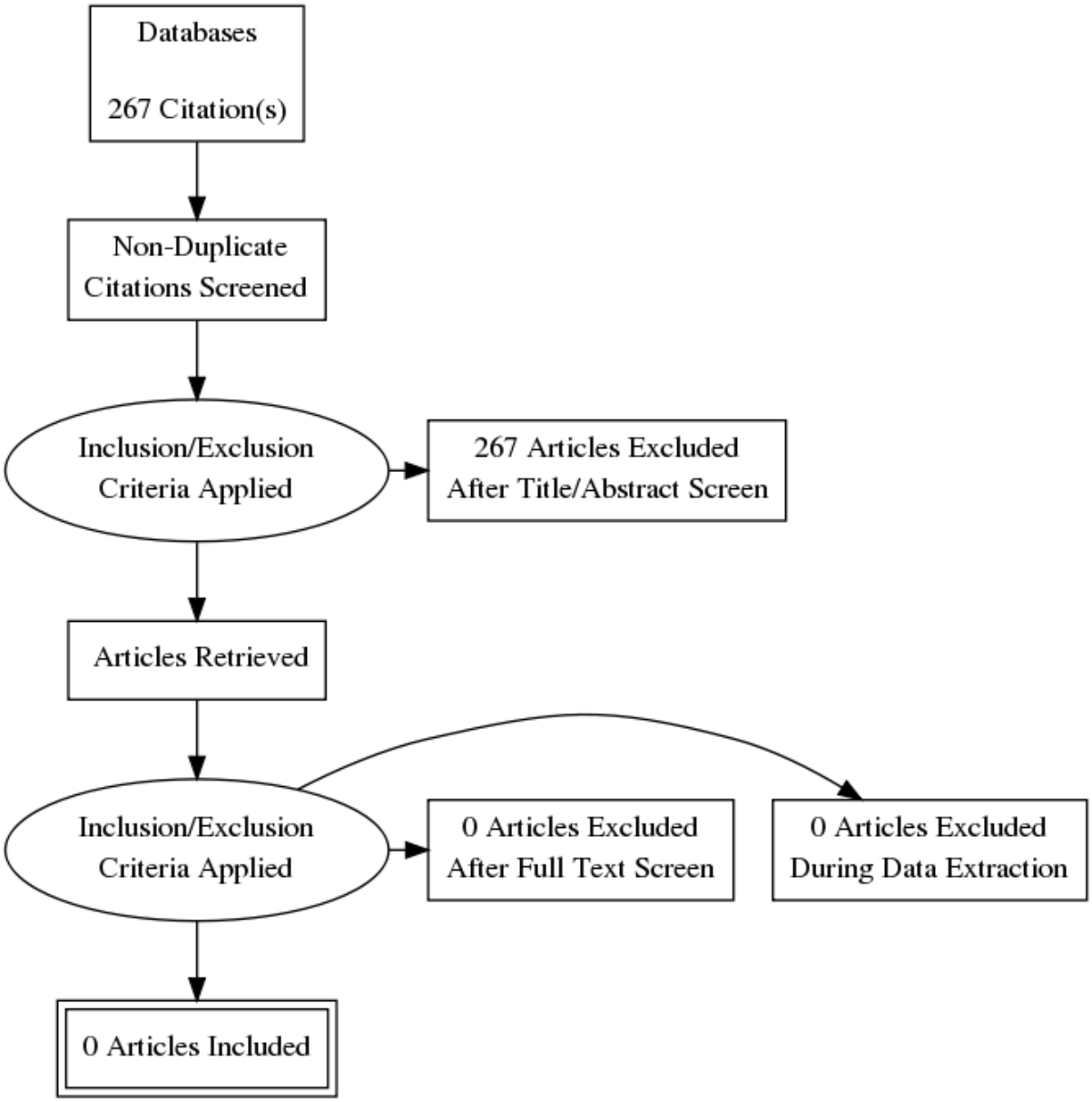
** QoL measure based on a bronchitis-associated quality-of-life score from Evans et al. 2002 [Lancet]

1. 95% confidence interval crosses both ends of the defined MID interval (0.8, 1.25)
2. 95% confidence intervals cross line of no effect
3. 95% confidence interval crosses one end of the defined MID interval (0.8, 1.25)
4. Inconsistency was non-applicable as one study reported 0 events and therefore did not contribute to the meta-analysis
5. I^2 of ≥33.3%
6. >33.3% of studies by weight in the meta-analysis were at a high risk of bias
7. I^2 of ≥66.7%

1

2

Appendix H – Economic evidence study selection



Appendix I – Excluded studies

Clinical studies

Study	Reason for exclusion
Engel B; Schindler C; Leuppi JD; Rutishauser J, Predictors of re-exacerbation after an index exacerbation of chronic obstructive pulmonary disease in the REDUCE randomised clinical trial, Swiss medical weekly, 147, w14439, 2017	Secondary publication of an included study that does not provide any additional relevant information [Post-HOC analysis of REDUCE trial looking at prognosis]

Appendix J – References

Included clinical studies

Systematic review

Walters JAE; Tan DJ; White CJ; Wood-Baker R. Different durations of corticosteroid therapy for exacerbations of chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2018, Issue 3. Art. No.: CD006897. DOI: 10.1002/14651858.CD006897.pub4.

RCTs

Included in meta-analysis

Chen G; Xie CM; Luo YF. [The effects and therapeutic duration of oral corticosteroids in patients with acute exacerbation of chronic obstructive pulmonary diseases]., Zhonghua jie he he hu xi za zhi = Zhonghua jiehe he huxi zazhi = Chinese journal of tuberculosis and respiratory diseases, 31, 8, 577-580, 2008

Leuppi JD; Schuetz P; Bingisser R; Bodmer M; Briel M; Drescher T; Duerring U; Henzen C; Leibbrandt Y; Maier S; Miedinger D; Müller B; Scherr A; Schindler C; Stoeckli R; Viatte S; von Garnier C; Tamm M; Rutishauser J, Short-term vs conventional glucocorticoid therapy in acute exacerbations of chronic obstructive pulmonary disease: the REDUCE randomized clinical trial., JAMA, 309, 21, 2223-2231, 2013

Sayiner A; Aytemur ZA; Cirit M; Unsal I, Systemic glucocorticoids in severe exacerbations of COPD., Chest, 119, 3, 726-730, 2001

Sirichana W, Sittipunt C, Kawkitinarong K WS, Comparison between 5 days and 10 days of prednisolone in treatment of acute exacerbation of chronic obstructive pulmonary disease. Respirology, 13 (Suppl 5), A120 [012-01], 2008

Wood-Baker R; Wilkinson J; Pearce M RG, A double-blind, randomised, placebo-controlled trial of corticosteroids for acute exacerbations of chronic obstructive pulmonary disease.]. Australian & New Zealand Journal of Medicine, 28, 262, 1997

Not included in meta-analysis

Gomaa M; Faramawy M IH; Duration of systemic corticosteroids treatment in COPD exacerbations., European Respiratory Society 18th Annual Congress; 2008 Oct 3-7, P3601, 2008

Rahman M; Abdullah A; Mamun SM HM, Role of 7-day and 14-day courses of oral prednisolone treatment in acute exacerbation of COPD. Chest, 839s-a, 2004

Salam T; Akers SM; Lotano R; Arnold GK; Bartter T; Pratter MR EA, Optimal duration of corticosteroid therapy in the treatment of exacerbations of chronic obstructive pulmonary disease. American Journal of Respiratory and Critical Care Medicine, 157 (3 Suppl), A801, 1998