THE MANAGEMENT OF MALIGNANT PLEURAL EFFUSIONS. THE IMPACT OF INDWELLING PLEURAL CATHETER VERSUS PLEURODESIS ON PATIENT-REPORTED OUTCOMES, NEED FOR RE-INTERVENTION, COMPLICATIONS, AND LENGTH OF STAY.

by

Richard Y. Liu

Submitted in partial fulfilment of the requirements for the degree of
Master of Science in Community Health and Epidemiology

at

Dalhousie University
Halifax, Nova Scotia
July 2021

<u>Dedication</u>

I would first like to thank my supervisor, Dr. Johnston, for his dedicated guidance and encouragement throughout this project. To my co-supervisor, Dr. Kephart, thank you sincerely for your unwavering support. To my committee members, Dr. Srinathan and Dr. Kidane, thank you for your invaluable insights and advice. To my fellow surgeons, Larry, Gord, Biniam, and Sadeesh, thank you for your mentorship and for picking up the slack while I did this.

To my parents and my sister, thank you for your love and support.

To Lesley, Margot, and William, you are my strength and everything I ever do is dedicated to you.

Table of Contents

List of Tables

List of Figures

Abstract

Background: Chemical pleurodesis and indwelling pleural catheters (IPCs) are the two most common treatments for malignant pleural effusions (MPEs). Previous systematic reviews inadequately address patient-reported outcomes (PROs).

Methods: A systematic review and meta-analysis was performed comparing IPC and pleurodesis for MPEs. Primary outcomes were patient-reported outcomes (PROs). Secondary outcomes included repeat pleural intervention, complication rates, and length of stay (LOS). PROs were analyzed using multiple methods based on established minimally important differences. Other outcomes were analyzed using standard methodology.

Results: For all PROs, there was only trivial to very small differences at specific timepoints. IPCs resulted in decreased repeat pleural intervention and decreased LOS but increased overall, infectious, and serious complications and tumour seeding.

Conclusions: Given a lack of difference in PROs, the decreased LOS and risk of repeat pleural intervention with IPC needs to be weighed against the increased risk of complications based on patient values and preferences.

List of Abbreviations

AATS: American Association for Thoracic Surgery
ARDS: Acute Respiratory Distress Syndrome
ARI: Absolute Risk increase
ARR: Absolute Risk Reduction
CDC-HRQOL-4: Center for Disease Control Health Related Quality of Life 4
CHEST: American College of Chest Physicians
CRQ: Guyatt Chronic Respiratory Questionnaire
ECOG: Eastern Cooperative Oncology Group
EORTC-QLQ-C30: European Organization for the Research and Treatment of Cancer
Quality of   Life Questionnaire Core 30
ESTS: European Society of Thoracic Surgeons
EQ5D: European Quality of Life Group 5-Dimension
EQ5D3L: European Quality of Life Group 5-Dimension 3-Level
GRADE: Grading of Recommendations, Assessment, Development and Evaluation
FACIT-PAL: Functional Assessment of Chronic Illness Therapy – Palliative
FACIT-TS: Functional Assessment of Chronic Illness Therapy Treatment Satisfaction
HRQoL: Health Related Quality of Life
IASLC: International Association for the Study of Lung Cancer
IPC: Indwelling Pleural Catheter
IR: Interventional Radiology
LCADL: London Chest Activity of Daily Living Scale
LOS: Length of Stay
MBS: Modified Borg Scale
MD: Mean Difference
MID: Minimally Important Difference
MPE: Malignant Pleural Effusion
MSAS: Memorial Symptom Assessment –Short Form
PRN: As needed
PRO: Patient-Reported Outcome
RCT: Randomized Controlled Trial
ROB: Risk of Bias
RR: Relative Risk
SD: Standard Deviation
SMD: Standardized Mean Difference
SRMA: Systematic Review and Meta-Analysis
VASD: 100mm Visual Analogue Scale for Dyspnea
WMD: Weighted Mean Difference

Acknowledgements

CHAPTER 1: Introduction

1.1 Management of Malignant Pleural Effusion

Malignant pleural effusion (MPE) is a frequent complication of cancer and can be seen in up to 15% of people who die of a solid organ malignancy (1). It is most commonly caused by lung and breast cancer but can be caused by other malignancies including lymphoma, ovarian cancer, and mesothelioma (2). An estimated 150,000 patients in the United States and 40,000 patients in the United Kingdom develop a MPE each year (3, 4). The incidence of MPE is estimated at 660 per million people, resulting in over a million people affected annually worldwide (5). While Canadian data is not available, based on global incidence, it can be estimated at approximately 20,000 patients a year.

Malignant pleural effusions represent an advanced stage of cancer with a median survival between 3 to 12 months and, in patients with poor performance status, as short as 30 days (2, 6-9). Malignant pleural effusions are symptomatic in the majority of cases, presenting with dyspnea, cough, and chest pain and significantly lowering the quality of life of patients (10). Given the poor prognosis and the significant symptom burden; interventions are targeted towards palliation of symptoms and involve drainage of the effusion.

The initial treatment for MPE is thoracentesis, whereby pleural fluid is drained through the chest wall using a needle or small bore soft tipped catheter. This procedure is generally performed by a physician and carries a small risk of complication each time (e.g. pneumothorax). Thoracentesis does not prevent re-accumulation of the pleural effusion and symptoms can recur rapidly (2). Because of this, a definitive procedure is

recommended in the setting of recurrent MPE (2, 6, 11). The two most common procedures used to provide durable palliation of symptoms are chemical pleurodesis and indwelling pleural catheters (IPCs).

Pleurodesis involves the use of a sclerosing agent to adhere the visceral pleura (lung) to the parietal pleura (chest wall) and prevent re-accumulation of fluid (12). The agent most commonly used is talc, a clay mineral composed of hydrated magnesium silicate. Pleurodesis can be performed by instillation of a liquid sclerosing agent via chest tube (e.g. talc slurry, doxycycline suspension, provodone-iodine solution) or by the application of talc powder to the surface of the pleura under direct vision during thoracoscopic surgery (i.e. talc poudrage). In both situations, a chest tube is left in the pleural space after pleurodesis to allow for apposition of the lung and parietal pleura. Some studies have examined removal of chest tubes after only 24 hours (13). Others have used outpatient pleurodesis protocols (14). However, most standard protocols require admission to hospital for several days with chest tube removal after fluid output has decreased below a threshold. Although some centers perform thoracoscopic pleurodesis in non-intubated patients with local anaesthetic and sedation, talc poudrage typically requires a general anaesthetic and single-lung ventilation (15).

The major drawbacks to chemical pleurodesis are the requirement for admission to hospital, pain related to inflammation of the pleura, and a rare but important risk of Acute Respiratory Distress Syndrome (ARDS) (16). The major benefit of chemical pleurodesis is the absence of an indwelling catheter at discharge. This alleviates the need for regular drainage and specialized care at home and may lower risk of complication related to the presence of the catheter. There may be a higher "pleurodesis rate", (defined

using several combinations of radiographic lung expansion and clinical indicators such as catheter removal) in patients treated with pleurodesis compared with indwelling pleural catheter (17-20). However, it is unclear how this marker correlates with symptoms and it is not a factor commonly used for shared decision making.

Indwelling pleural catheters are subcutaneously tunnelled catheters placed in the pleural cavity to allow for outpatient drainage of pleural effusion. They can be inserted via thoracoscopy or in an outpatient clinic using a Seldinger (wire-guided) technique (21). They have an established role in patients with trapped lung (in whom pleurodesis is ineffective because of lack of apposition) and have increasingly been used as primary management of MPE (6).

The major drawback is the presence of an indwelling catheter and the risks and inconvenience associated with this. The major benefit of IPC is the ability to insert the catheter as an outpatient procedure. Studies have consistently shown decreased length of stay in patients undergoing insertion of IPC compared with patients treated with pleurodesis and because of this, some authors suggest that IPCs should now be the procedure of choice for the management of MPE (22).

Although IPCs have been accepted as the procedure of choice by many clinicians, their use is not without risk. The presence of an indwelling catheter is a possible source of pleural or soft tissue infection and maintenance of the catheter requires diligent care. Even in the rigorously controlled setting of a Randomized Controlled Trial (RCT), there is a notable risk of cellulitis and pleural space infection associated with IPCs. Davies it al. (2012) reported more cases of infection in the IPC arm (5/52 required admission for intravenous antibiotics, and an additional 2/52 were managed as outpatients with oral

antibiotics) compared to the talc slurry arm (1/54 participants requiring hospital admission for pleural infection) (18). In retrospective series, the rate of infection associated with IPC can be as high as 25.5% (23). Additionally, catheter tract metastases have been reported in 6.7% of patients with IPC and can require treatment with external beam radiotherapy (24). Local expertise and volume may also be important, as demonstrated by the high reported rate of complications requiring additional procedural intervention (25%) at a non-teaching secondary care hospital in the United Kingdom (25).

The success of IPCs relies on access to specialized follow-up care. In a comparison of Canadian patients followed by a specialized pleural effusion clinic and those who had insertion of IPC as a "one time procedure" by Interventional Radiology (IR), there was a significantly higher rate of repeat chest procedures, infection and hospital visits in the IR group (53% vs. 32.3%, p=0.015) (26). Outpatient drain/dressing care and drainage also requires dedicated home care nurses or local hospitals and clinics comfortable with IPCs.

Because of the poor prognosis and heavy symptom burden, patient-reported outcomes (PROs) should play a central role in decision-making regarding the benefits and harms of available therapy. However, historically, the literature has focused on outcomes such as length of hospital stay and pleurodesis rate.

1.2 Patient-Reported Outcomes in Malignant Pleural Effusion

Several randomized trials have examined PROs in MPE patients treated with IPC or pleurodesis. Two RCTs suggest that there may be a difference in dyspnea scores

between patients treated with IPC and pleurodesis. Demmy et al. (2012) performed post-hoc multivariable linear regression and found that dyspnea scores favoured IPC (IPC 8.5 vs. Talc slurry 6.1, p=0.047) after adjusting for baseline dyspnea score, initial drainage, gender, inpatient status, and performance status (17). Subgroup analysis revealed that this difference was driven by poor dyspnea scores in pleurodesis patients with poor lung expansion (i.e. trapped lung) (IPC 9.0 vs. Talc 4.9, p=0.033). Davies et al. (2012) demonstrated a significant decrease in dyspnea measure on a 100mm Visual Analogue Scale for dyspnea (VASD) in patients treated with IPC at 6 months follow-up (mean difference (MD) -14.0mm [95%CI -25.2 to -2.9, p=0.01) (18). However, there was no difference between groups at any other time point. Other RCTs have not demonstrated any between group differences in patient-reported overall health related quality of life (HRQoL) and patient-reported pain (17, 18, 21, 27).

Several systematic reviews and meta-analyses (SRMAs) have been conducted since the results of two RCTs comparing IPCs to chemical pleurodesis for MPE were published in 2017 (27, 28).

A SRMA of RCTs by Yeung et al. (2020) suffered from significant flaws in methodology including a rudimentary search strategy that did not yield a well-known RCT comparing talc pleurodesis to IPC (17, 29). Inappropriately included in this analysis was a RCT that included IPC in both arms of the study (30). Mean differences were generated for available data on VASD scores from two RCTs (n=300) at four weeks (MD -2.16 [95%CI -7.59 to 3.27, I^2= 0%]) and from three RCTs (n=402) at six weeks (MD -0.42 [95%CI -5.94 to 5.10, I^2=9%]). Data from other patient-reported dyspnea instruments was not examined. Health related quality of life data was not pooled.

Another SRMA of RCTs published by Wang et al. (2020) examined patient-reported dyspnea and HRQoL (19). Post-treatment scores for VASD were pooled using data from two RCTs (n=150) at 30 and 42 days respectively (standardized mean difference (SMD) -1.50 [95%CI -3.80 to 0.80]) (18, 27). For patient-reported HRQoL, European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC-QLQ-C30) data at six weeks of follow-up from one RCT was pooled with European Quality of Life Group 5-Dimension (EQ5D) data from 12 months follow-up and presented as a SMD (-1.50 [95%CI -3.8 to 0.8, I^2=98%]).

In their SRMA of RCTs, Iyer et al. (2019) report pooling modified Borg scale (MBS) data from available RCTs (31). They do not present any data in the text or graphically and state that there was no difference between treatment groups at rest or during exercise. Health related quality of life was not examined.

The most comprehensive SRMA of RCTS on pleurodesis for MPE was performed by Dipper et al (2020) (20). The authors updated a previously published systematic review and network meta-analysis to include 80 RCTs examining multiple interventions for the management of MPEs. The majority of studies analyzed in this review compare various forms of chemical sclerosant to each other (32). The updated review examined IPC as well. The authors analyzed patient-reported dyspnea data using VASD data from two RCTs (n=160). In one trial, this was collected at 42-days follow-up. In the other, it was collected at 180-days follow-up. They reported no difference in post intervention VASD score between participants receiving an IPC compared to talc slurry (MD -6.12 [95%CI -16.21 to 4.08]) based on low certainty evidence. The vast scope of this review

did not allow for an in-depth look at PROs or other comparisons between IPC and chemical pleurodesis.

Sivakumar et al. (2020) performed a systematic review examining patient-reported HRQoL measures in observational studies and RCTs in patients undergoing IPC placement or pleurodesis for malignant pleural effusion (33). This review summarized the published studies examining patient-reported HRQoL. The authors do not directly compare IPC to pleurodesis nor do they perform a meta-analysis. No other patient-reported outcomes were examined. Several cohort studies examining patient-reported overall HRQoL were identified in this systematic review (34, 35).

1.3 Rationale

There are significant limitations to the examination of PROs in previously published SRMAs. The authors of these studies only analyzed data from studies using the same PRO instrument. Additionally, the time points from which data was extracted appeared to be arbitrary. For example, the authors of one SRMA pooled VASD scores collected from one study at 42 days with data collected from another at 180 days (20). These time points likely represent two distinct stages in the disease process with different symptom burden.

We addressed the limitations of the currently available evidence by performing a systematic review and meta-analysis comparing IPCs and pleurodesis for the treatment of MPE with a focus on PROs. Given the paucity of PRO data in this population, we included both cohort studies and RCTs. We examined PROs at multiple time points across all outcomes according to our protocol (Chapter 2). Using well-established

methodologies to determine PROs suitable for pooling, we presented pooled PRO data

using multiple presentation methods (36, 37). Available estimates of minimally important

differences (MID) were used to improve the interpretability of our results (38-41). We

examined other patient-important outcomes crucial in clinical decision making including:

need for repeat pleural intervention, complication rate, and length of stay

CHAPTER 2: Methods

      The protocol for this systematic review and meta-analysis was registered with PROSPERO. Registration ID: CRD4201443733.

2.1 Data Sources

      In consultation with two expert medical librarians (TG, JR), a comprehensive search of Medline, Embase, Scopus, Web of Science, and the Cochrane Library from inception to November 2020 was conducted. The full search strategy is available in Appendix 1.

      Grey literature search included OpenGrey.eu, clinicaltrials.gov and the last three years of abstracts from the following society meetings: the American Association for Thoracic Surgery (AATS), the European Society of Thoracic Surgeons (ESTS), American College of Chest Physicians (CHEST), and the International Association for the Study of Lung Cancer (IASLC).

      Bibliographies for all studies chosen for full-text review as well as relevant review articles were searched for additional studies not identified by the electronic searches.

2.2 Inclusion Criteria

      Published randomized trails and cohort studies comparing indwelling pleural catheter (IPC) to chemical pleurodesis in adult patients (≥18 years of age) with symptomatic pleural effusion resulting from an underlying malignant process were included. The primary outcomes of interest were patient-reported outcomes (i.e. dyspnea,

overall health related quality of life [HRQoL], pain, satisfaction with care, and overall functional status). Secondary outcomes included need for repeat pleural intervention, complication rate (overall, serious, infectious, acute respiratory distress syndrome [ARDS], catheter fracture, and tumour seeding), and length of hospital stay (initial admission, total days in hospital [including re-admission], and total effusion related days in hospital). Studies were included if they reported any outcome of interest.

Studies were excluded if they included patients with both malignant and non-malignant effusions with no clear division of groups. They were also excluded if they included patients with effusions in other body cavities (pericardial, peritoneal), as it would be difficult to distinguish the treatment effect of the pleural intervention from other interventions.

2.3 Screening

After removal of duplicate articles, two reviewers (RL and EP) independently reviewed titles and abstracts. Similarly, two reviewers (RL and EP) reviewed all full-text articles independently. Excluded articles were categorized according to reason for exclusion. Disagreements over study inclusion were resolved through consensus, or, if needed, via discussion with a third senior reviewer.

2.4 Data Extraction

Two reviewers independently extracted data using a standardized data extraction form. Data abstracted from each study included study, population, and intervention characteristics:

1. Study characteristics: Publication date, country, language, study design, trial registration, single vs. multicenter, academic vs. other hospital, and funding source.

2. Patient characteristics: Number of participants, gender (percentage of female patients), age (mean/median), comorbidity (comorbidity score used, mean/median score), functional status (e.g. Eastern Cooperative Oncology Group performance score), tumor types (number and percentage of patients with lung, breast, mesothelioma and other cancers), and inclusion of patients with trapped lung.

3. Description of intervention:

   i. Pleurodesis: type of sclerosant (percentage of talc, doxycycline, and other), mode of administration (thoracostomy tube or thoracoscopy), and technique of administration (drain size, analgesia used, duration of drainage, patient positioning, use of intrapleural fibrinolytics, timing of pleurodesis (immediately after insertion of tube or after chest tube output is below a certain volume)

   ii. Indwelling pleural catheter (IPC): type of catheter (PleurX, Tenchkoff, other), mode of insertion (bedside Seldinger or thoracoscopy), protocol for drainage (frequency, quantity, and indication for drainage), and additional methods to optimize IPC regimen

4. Primary outcome data consisted of patient-reported outcomes (PROs) subdivided into: i) dyspnea, ii) overall HRQoL, iii) pain, iv) patient satisfaction with care, and v) overall functional status (36). Patient-reported outcomes were collected at

4 different time points: i) immediately post procedure (≤2 weeks), ii) short-term follow-up (>2 weeks, ≤6 weeks), iii) medium-term follow-up (>6 weeks, <6 months), and iv) long-term follow-up (≥6 months). If there were multiple measures reported within a time point, the longest duration of follow-up was used. If both single measurements and mean scores over time were available, mean scores were used as they were felt to be less subject to momentary variability. If both endpoint data and change data were reported, change data was used as it was felt to be more representative of treatment effect and less dependent on baseline condition (42).

5. Secondary outcomes included: i) need for repeat pleural intervention, ii) complication rate (overall, serious, infectious, ARDS, catheter fracture, and tumour seeding), and iii) length of stay (index admission, total days in hospital [including readmission], total effusion related days in hospital).

2.5 Data Analysis

Due to anticipated clinical heterogeneity, random-effects meta-analysis was conducted using RevMan 5.4.1 (43). Meta-analyses were performed based on outcomes that were similarly defined and reported in at least two studies. Continuous data was analyzed using inverse-variance weighted mean difference (WMD), while dichotomous data was pooled using the Mantel-Haenszel method for relative risks (RR) with corresponding 95% confidence intervals.

When data was unavailable in published manuscripts and supplementary material, further information was sought from corresponding authors. Data that was presented only in graphical form and could not be obtained from authors was extracted using measurements from graphical representations of mean and 95% confidence intervals. When standard deviation (SD) for reported means was not available, SD was calculated using the reported interquartile range (assuming normality of the data) or imputed using the SD of the same measure and nearest time point from a similar study (44, 45).

Statistical heterogeneity was assessed via visual inspection of forest plots and using the I^2 statistic (46). To interpret heterogeneity, based on the Cochrane Handbook, I^2 values of 0-40% were considered to represent non-important heterogeneity, 30-60% to represent moderate heterogeneity, 50-90% to represent substantial heterogeneity, and 75-100% to represent considerable heterogeneity (47).

2.6 Pooling of Patient-Reported Outcomes

The decision to pool was based on the extent to which different instruments measure the same underlying construct (36). This has been addressed in previous studies by having expert reviewers independently examine each instrument in order to create a list of combinable instruments measuring the same or similar constructs. For instance, the authors of a meta-analysis of psychological interventions for premenstrual syndrome identified 25 different PRO measures. Two clinical researchers with expertise in the study area grouped 16 of these PRO measures into 6 conceptual constructs (anxiety, behavioural changes, depression, interference with daily activities, sexual relationship, and impact on perception of water retention and edema) to allow for pooling of results

(48). For our study, as needed, content experts were asked to review the instruments identified in the systematic review to determine suitability for pooling.

2.7 Data Analysis Specific to Patient-Reported Outcomes

If instruments measured the same or similar construct using the same PRO instrument, data was pooled as a weighted mean difference (WMD) in natural units of the instrument. If different instruments measured the same or similar constructs, PRO data was reported as standardized mean difference (SMD) for each time point (37).

However, SMDs have several limitations. Firstly, this method assumes that differences in SD among studies reflect a difference in measurement scale, rather than a difference in study populations (49). Secondly, the effect size may be difficult to interpret since it is reported as units of SD rather than the units of a PRO instruments (50). Because of this, where possible, effect estimates were presented in several other ways. Interpretation of these other presentation methods rely on the availability of minimally important difference (MID) estimates. The MID represents the smallest change in an outcome that an informed patient perceives as important and would lead a patient or clinician to consider a change in management (51). Several approaches for estimating an MID exist. The two most common are: i) an anchor-based approach where the relationship between a PRO instrument and an independent (anchor) measure that has clinical relevance to the patient is examined, and ii) distribution-based methods which rely solely on the distribution of scores for the PRO measure of interest (52).

1. *Conversion into units of the most familiar instrument to clinicians*. A linear transformation of trial data to the natural units of the most familiar instrument was

performed using the methodology described by Thorlund et al. (2011) (53). After

conversion into the natural units of the most familiar instrument, mean difference

estimates were generated using standard inverse-variance methods. Instruments

selected for conversion to natural units had established anchor-based MID

estimates in malignant effusion/lung cancer patient populations (39, 40).

Conversion of combinable PROs to the natural units of the most familiar

instrument can improve interpretability by the target audience (37). Interpretation

of effect size was based on the relationship between the mean difference and the

MID estimate.

    i.   A mean difference of <0.5 MID represented a trivial difference.

    ii.   A mean difference of ≥0.5 MID and <1.0 MID represented a very small

        difference.

    iii.   A mean difference of ≥1.0 MID and <1.5 MID represented a small but

        important difference.

    iv.   A mean difference ≥1.5 MID and <2.0 MID represented a moderate

        difference.

    v.   A mean difference ≥2.0 MID represented a large difference.

2. *Minimally Important Difference units*. For pooled outcomes where the MID

    could be established for all instruments, data was reported in MID units (53). The

    use of MID units helps to circumvent the assumption of similar variability in

    study populations necessary for SMD. It also aids in interpretation by decision

    makers (patients and clinicians) (50). When calculating MID units, at the study

    level, MID estimates replace the SD and estimates are pooled across studies (i.e.

unlike the SMD were the MD is divided by the SD, the MD is divided by the

MID) (50). Ideally MID units are based on anchor-based MID estimates and are

derived from our population of interest (i.e. patients with MPE) (40). Due to the

absence of appropriate anchor-based MID estimates for some PRO instruments,

distribution based methods were used (54).

2.8 Risk of Bias Assessment

All risk of bias assessments were completed independently and in duplicate

by two reviewers and differences were resolved by consensus, and, if needed, via

discussion with a third senior reviewer. Risk of bias assessments were completed both at

the study (Tables 2 and 3) and the outcome level (Appendix 2).

Using the CLARITY risk of bias instrument for randomized controlled trials,

RCTs were examined for risk of bias in 6 domains relating to: i) random sequence

generation, ii) allocation concealment, iii) blinding, iv) incomplete outcome data, v)

selective outcome reporting, and vi) other problems (55).

Similarly, included cohort studies were assessed for risk of bias using the

CLARITY risk of bias instrument for cohort studies examining risk of bias in 8 domains

relating to: i) cohort selection, ii) assessment of exposure, iii) presence of outcome prior

to intervention, iv) adjustment for prognostic variables, v) assessment of prognostic

factors, vi) assessment of outcome, vii) adequacy of follow-up and viii) similarity of co-

interventions (56).

For both instruments, the response options for each category consist of 4

categories including 'definitely yes' (indicating low risk of bias), 'probably yes',

'probably no' and 'definitely no' (indicating high risk of bias), a method previously validated (57).

Randomized trials were considered higher risk of bias if there was a response of 'probably no' or 'definitely no' in at least two categories. Otherwise, studies were considered lower risk of bias. Cohort studies were considered higher risk of bias if there is a response of 'probably no' or 'definitely no' in at least three categories. Otherwise, studies were considered lower risk of bias.

Given the high rate of anticipated attrition due to mortality at later time points, criteria were generated for assessment of risk of bias due to loss to follow-up. Studies were consider at least 'probably no' to the question "Was loss to follow-up infrequent?" if there was: i) >50% loss to follow-up for any reason including death, ii) >10% difference in loss to follow up between groups, or iii) >25% missing data amongst living patients.

2.9 Subgroup Analysis

To explore observed heterogeneity, subgroup analysis was planned separating:

1. *Inclusion of trapped lung*. Pleurodesis has been shown to be less effective in this group (58).

2. *Mode of administration of pleurodesis*. Talc poudrage in the operating room has been suggested to be more effective than talc slurry at the bedside (59-61). This subgroup analysis was planned noting that a RCT showed no impact of mode of administration (4).

3. *Type of sclerosant used*. Talc has been shown to be more effective than other sclerosants (20, 62).

4. *Method of placement of IPC*. Although it has not been demonstrated in the existing literature, it is plausible that placement of IPC in the operating room may decrease the risk of infection when compared to placement in clinic.

5. *Drainage schedule of IPC*. It has been demonstrated that a daily drainage schedule may be more effective in preventing recurrence of effusion when compared to an "as needed" drainage schedule (20).

2.10 Sensitivity Analysis

Sensitivity analysis was performed examining high and low risk of bias studies as we anticipated that high risk of bias studies would show larger treatment effects than low risk of bias studies. In a meta-epidemiological study examining 146 meta-analyses including 1346 trials across a wide range of interventions, high risk of bias studies were associated with larger estimates of intervention effect for subjectively assessed outcomes (63).

2.11 Publication Bias

A priori, visual inspection of funnel plot symmetry was planned to assess for publication bias for meta-analyses with at least 10 studies (64, 65). Egger's test was planned as a quantitative measure of funnel plot asymmetry for continuous outcomes with intervention effects measured as mean differences (65, 66).

2.12 Certainty of Evidence

The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) framework was used to report our overall certainty in estimates for each of the above stated outcomes. Using the GRADE guidelines, quality of evidence was divided into high, moderate, low, and very low for each outcome taking into account: study design, risk of bias, consistency of results, directness of evidence, precision, publication bias, effect size, dose response, and potential for residual confounding. Summary of finding tables were generated to summarize the outcomes, including the number of studies, number of participants, the relative and absolute effect size and certainty of evidence for each outcome (67).

CHAPTER 3: Results

3.1 Search Results

Primary database search identified 1894 studies and an additional 47 studies were identified through OpenGrey.eu, clinicaltrials.gov, and the last three years of abstracts from AATS, ESTS, CHEST, and IASLC conferences. After duplicates were removed, 1362 title and abstracts were independently screened. Forty-eight full text articles were collected and assessed for eligibility. After independent full-text screening, 16 studies were eligible, including six RCTs and ten cohort studies (Figure 1). Amongst the cohort studies, two of the studies presented previously unreported data from one of the randomized trials (68, 69). The data from these studies was considered as a part the Davies et al. 2012 RCT (18). One RCT and one cohort study were not suitable for meta-analysis as they included IPC in both arms of the study (30, 70). This resulted in a total of five RCTs with a total of 557 participants and seven cohort studies (two prospective and five retrospective) with a total of 988 participants that were suitable for inclusion in meta-analysis (17, 18, 21, 27, 34, 35, 71-76).

3.2 Characteristics of Included Studies

Table 1 shows the characteristics of included studies. The overall risk of bias was 'lower' in three randomized trials (Table 2) (18, 21, 27, 30). Two randomized trials were judged to have overall 'higher' risk of bias due to lack of blinding and incomplete outcome data (17, 71). Among the eligible cohort studies, the overall risk of bias was considered 'lower' in three studies (Table 2) (34, 35, 70, 72). Four cohort studies were judged to have overall 'higher' risk of bias due to problems with participant selection

bias, inappropriate or inadequate adjustment for prognostic variables, inadequate

assessment of outcome (i.e. unclear description of data collection), inadequate follow-up

(i.e. follow-up not described), and unequal exposure to co-interventions (73-76).

3.3 Patient-Reported Dyspnea

Combinable data on patient-reported dyspnea was available from four RCTs

(n=490) (18, 21, 27, 71). Data on dyspnea was reported using pooled estimates at the

immediate (three studies), short-term (four studies), medium-term (four studies), and

long-term (three studies) time points (Figure 2) (18, 21, 27, 71). Patient-reported dyspnea

was reported using the 100 mm visual analogues scale for dyspnea (VASD) in two

studies and the modified Borg scale (MBS) in two studies (18, 21, 27, 71).

Conversion to the natural units of the VASD was untaken. The VASD ranges

from 0 (maximum breathlessness) to 100 (no breathlessness) with an anchor-based MID

estimate of 19 mm among a malignant effusion population (40). A positive mean

difference represents a greater improvement in dyspnea for IPC compared to pleurodesis.

At immediate follow-up (n= 305), there was little to no effect of IPC on improvement in

dyspnea score: mean difference (MD) 1.84 [95%CI -3.82 to 7.51, I^2=0%]. At short-term

follow-up (n=309), there was a trivial benefit favouring IPC: MD 6.99 [95%CI 0.28 to

13.69, I^2=0%]. At medium-term follow-up (n=180) and at long-term follow-up (n=97),

based on the point estimate, there was little to no effect: MD 4.78 [95%CI -3.95 to 13.51,

I^2=28%] and MD 2.71 [95%CI -14.18 to 19.60, I^2=58%], respectively. While the upper

end of the 95%CI surpassed the 19 mm MID at long-term follow-up, the CI was very

wide and imprecise, suggesting the potential for minimal benefit and harm.

One randomized trial collected data on patient-reported dyspnea that could not be pooled in meta-analysis (17). In a randomized trial of 57 patients, Demmy et al. (2012) report that "multivariate regression analysis revealed that tunneled catheter drainage [i.e. IPC] had better dyspnea scores than talc pleurodesis (8.5 versus 6.1, p=0.047)". Authors did not provide any data on group means or variance for this estimate, nor did they respond to multiple requests for their data.

Two prospective cohort studies collected data on patient-reported dyspnea that could not be pooled in meta-analysis. One study utilized the London Chest Activity of Daily Living Scale (LCADL) to assess dyspnea immediately post treatment and at two and six weeks of follow-up in 104 patients. This data was only presented graphically and no data was provided on number of patients assessed at each time point (35). Authors reported no difference in LCADL score between groups. In another study of 65 patients, Fysh et al. (2012) report that a greater proportion of patients treated with IPC (93.3%) reported improved dyspnea measured on a VASD compared with pleurodesis patients (78.6%) (34). This was defined as an improvement of half a standard deviation above pre-treatment score. No mean scores or variance data was provided. Again, after multiple attempts to request the data, the authors of these two studies did not respond.

3.4 Patient-Reported Overall HRQoL

Combinable data on patient-reported overall HRQoL (n=490) was available from two RCTs and one cost analysis based on data collected from a RCT by Olfert et al. (2017) Overall HRQoL data was reported using the European Quality of Life 5 Dimension 3-Level (EQ5D3L) utility index using the UK valuation set, a modified

European Quality of Life Five Dimension (EQ5D) questionnaire, and the Chronic

Respiratory Questionnaire (CRQ) (21, 27, 69).

Data on overall HRQoL was reported using pooled estimates at the immediate

(two studies), short-term (three studies), medium-term (two studies), and long-term (two

studies) time points (Figure 3) (21, 27, 69). There were similar findings for all methods of

presentation. For all presentation methods, a positive value represents a greater

improvement in overall HRQoL for patients treated with IPC compared to pleurodesis.

There was no difference in overall HRQoL when examined as a standardized

mean difference (SMD) at the immediate, short-term or medium-term time points: SMD

0.09 [95%CI -0.19 to 0.38, I² = 0%], SMD -0.04 [95%CI -0.29 to 0.21, I² = 0%] and

SMD -0.18 [95%CI -0.55 to 0.19, I² = 0%], respectively. At the long-term time point

there was a very small difference in overall HRQoL favouring pleurodesis: SMD -0.24

[95%CI -0.68 to 0.21, I² = 0%].

Conversion to the natural units of the EQ5D3L was undertaken. The EQ5D3L

ranges from 0 (death) to 1 (perfect health) with an anchor based MID estimate of 0.1

among lung cancer patients (39). At immediate, short-term, and medium-term follow-up

there was little to no difference in overall HRQoL between groups: MD 0.03 [95%CI -

0.05 to 0.10, I²=0%], MD -0.01 [95%CI -0.07 to 0.05, I²=0%] and MD -0.04 [95%CI -

0.11 to 0.04, I²=0%], respectively. At long-term follow-up there was a very small

difference, half the MID, favouring pleurodesis: MD -0.06 [95%CI -0.17 to 0.05, I²=0%].

While the lower end of the 95%CI surpasses the 0.1 MID at both medium-term and long-

term follow-up, the CI for both is very wide and imprecise, suggesting the potential for

minimal benefit and harm.

When examined as MID units, there was little to no difference in overall HRQoL between groups at immediate, short-term, or medium-term follow-up: MD in MID units 0.2594 [95%CI -0.4818 to 1.0007, I^2=0%], MD in MID units 0.1335 [95%CI -0.5292 to 0.7961, I^2=0%], and MD in MID units -0.4370 [95%CI -1.3118 to 0.4378, I^2 = 0%] respectively. At long term follow-up there was a very small difference favouring pleurodesis, again equivalent to greater than half the MID: MD in MID units -0.598 [-1.690 to 0.491, I^2=0%].

We identified multiple other studies in our systematic review that collected overall HRQoL scores but for which reported data was insufficient for inclusion in meta-analysis. One randomized trial collected overall HRQOL data at 7 and 30 days of follow-up using the Memorial Symptom Assessment–Short Form (MSAS) but only reported a correlation coefficient between MSAS score and lung expansion on chest x-ray (17). Two cohort studies collected data on patient-reported overall HRQoL. One prospective cohort study utilized the Functional Assessment of Chronic Illness-Palliative (FACIT-PAL) to assess overall HRQoL immediately post treatment, and at two and six weeks follow-up in 104 patients but only presented this data graphically and no data was provided on number of patients assessed at each time point (35). They report no difference in FACIT-PAL score between groups. Another prospective cohort study utilized a 100 mm visual analogue scale for overall HRQoL (34). In this study of 65 patients, Fysh et al. (2012) report that a greater proportion of patients treated with IPC (93.3%) reported improved overall HRQoL compared with pleurodesis patients (50.0%) defined as an improvement of half a standard deviation above pre-treatment score. No raw scores are given that

could be utilized in meta-analysis. After multiple attempts to request data, the authors of these studies did not respond.

3.5 Patient-Reported Pain

One RCT (n=106) and no cohort studies compared IPC to chemical pleurodesis for patient-reported chest pain as measured by a 100 mm Visual Analogue Scale (18), with positive scores indicating greater improvement in pain among IPC patients. While results slightly favoured IPC, there was little to no difference in patient-reported chest pain: mean difference 3.80 [95%CI -7.32 to 14.92]. An anchor-based MID for the 100 mm VAS for chest pain was estimated at 16 mm in a cohort of patients undergoing pleural intervention (77), suggesting that IPC while slightly better than pleurodesis does not surpass the MID even at the upper bound of the 95%CI.

3.6 Patient-Reported Satisfaction with Care

One prospective cohort study (n=104) examined patient satisfaction with care as measured by the Functional Assessment of Chronic Illness Therapy-Treatment Satisfaction (FACIT-TS) score in four different groups: i) chest tube and talc slurry, ii) thoracoscopic talc poudrage, iii) bedside Tenchkoff (IPC) insertion, and iv) thoracoscopic Tenchkoff (IPC) insertion (35). Data was collected at two and six weeks of follow up. Assuming complete follow up, we graphically extracted data and combined the "chest tube and talc slurry" group with the "thoracoscopic talc poudrage" group to generate a mean score and standard deviation for patients treated with talc pleurodesis as a whole. Using the same methodology, we combined the "bedside Tenchkoff (IPC)" group with

the "thoracoscopic Tenchkoff (IPC)" group to generate a mean score and standard deviation for the indwelling pleural catheter group as a whole. There was little to no difference in treatment satisfaction between IPC and pleurodesis patients at both immediate and short-term time points: MD 2.53 [95%CI -11.26 to 16.31] and MD 3.27 [95%CI -8.06 to 14.61], respectively. Unfortunately, no MID estimates are available for the FACT-TS instruments.

3.7 Patient-Reported Overall Functional Status

Amongst RCTs, only Demmy et al. (2012) collected data on patient-reported overall functional status using the Karnofsky self-reported performance rating scale (17). This data was only presented as a predictor of lung expansion on CXR. Again, no scores or variance data was presented or accessible.

A single propensity matched cohort study by Freeman et al. (2013) reported no statistically significant difference in improvement in Eastern Cooperative Oncology Group (ECOG) performance status score (lower score means higher function) between talc pleurodesis patients (-1.2) and IPC patients (-1.6) after intervention (p=0.4) at an unspecified time point (78). No variance data was provided. Additionally, it is unclear if this measure was patient-reported or clinician assessed.

3.8 Need for Repeat Pleural Intervention

Four randomized trials (n=472) report on need for repeat pleural intervention (Figure 4). Pleural intervention was defined in all trials as an ipsilateral drainage procedure for reaccumulation of pleural effusion. Interventions for other reasons (e.g.

removal of infected IPC)   were not included in this outcome. Indwelling pleural catheters were found to result in a large reduction in need for repeat pleural intervention: relative risk (RR) 0.32 [95%CI 0.18 to 0.55, I^2=0%]. This corresponds with a15.2% absolute risk reduction (ARR) for need for repeat pleural intervention in IPC (71 per 1000) compared with pleurodesis (223 per 1000) (Table 7).

3.9 Complication Rate

Five randomized trials (n=604) report on complication rates for both IPC and pleurodesis (Figure 5). The most common complications with pleurodesis were pain and worsening dyspnea. The most common complications with IPC were cellulitis, pleural infection, catheter blockage, pain and worsening dyspnea.

Three RCTs reported on overall complications (n=384). Indwelling pleural catheters were associated with a large increased risk in overall complications compared with pleurodesis: RR 2.48 [95%CI 1.10 to 5.60, I^2=82%]. This corresponded with an absolute risk increase (ARI) of 28.1% in patients treated with IPC (471 per 1000) compared with pleurodesis (190 per 1000) (Table 7).

Four RCTs reported on serious complications (n=396). Again, IPCs were associated with a moderate increased risk in serious complications: RR 1.53 [95%CI 0.83 to 2.82, I^2=8%], corresponding to an absolute risk increase of 4.8% in IPC (138 per 1000) compared with pleurodesis (90 per 1000). Four RCTs report on infectious complications (n=475) demonstrating that IPCs are associated with a large increased risk of infectious complications: RR 4.17 [95%CI 1.61 to 10.78, I^2=0%], corresponding to an absolute risk increase of 7.4% in patients with IPC (97 per 1000) compared with pleurodesis (23 per

27

1000). Two RCTs report on tumour seeding (n=240). And once again IPCs were associated with a large increased risk of tumour seeding compared with pleurodesis: RR 3.23 [95%CI 0.37 to 28.04, I^2=0%] corresponding to an absolute risk increase of 2.8% in IPC compared to pleurodesis. No catheter fracture or ARDS was reported in any of the included trials.

3.10 Length of Stay

Based on 2 RCTs (n=243), length of stay during initial admission is nearly 3 days shorter in patients treated with IPC compared with pleurodesis: MD -2.78 [95%CI -4.41 to -1.15, I^2=92%] (Figure 6). Additionally, based on data from the same two trials, there is a reduction in effusion related days spent in hospital by more than three days in patients treated with IPC compared with pleurodesis: MD -2.55 [95%CI -4.41 to -0.68, I^2=78%].

Thomas et al. (2017) examined total days spent in hospital and found that the mean number of days in patients treated with IPC spent in hospital was 12.7 (SD 13.4) compared with a mean of 16.3 days (SD 15.2) in pleurodesis patients (27). This results in an estimated mean difference in total days spent in hospital of -3.60 [95% CI -8.29 to 1.09].

In a trial of 144 patients, Putnam et al. (1999) report a shorter length of stay during index admission in IPC patients (median 1.0) compared with pleurodesis patients (median 6.5) (21). Similarly, in a trial of 94 patients, Boshuizen et al. (2017) report lower median length of stay post procedure (0 vs. 5 days) and in total (2 vs. 7 days) in IPC

patients (71). Again, data from both of these studies could not be pooled with the studies in figure 6 as no variance data was reported.

3.11 Subgroup Analysis

Certain subgroup analysis could not be performed as planned for the following reasons:

1. *Inclusion of trapped lung.* All meta-analyzed studies included trapped lung, and data was not presented separately within studies (Table 1).
2. *Mode of administration of pleurodesis.* All meta-analyzed studies administered pleurodesis via tube thoracostomy (Table 1).
3. *Method of placement of IPC*. All meta-analyzed studies placed IPC exclusively using Seldinger technique (Table 1).

There were two subgroups where analysis was possible: type of sclerosant and drainage schedule of IPC. In both of these subgroups, the number of trials was inadequate to produce meaningful findings on heterogeneity. However, an exploratory analysis was performed.

Four RCTs used talc as the sclerosing agent and one used doxycycline. When subgroup analysis was conducted comparing talc and doxycycline pleurodesis amongst these five RCTs, there was no significant subgroup difference for immediate dyspnea, short-term dyspnea, medium-term dyspnea, short-term HRQoL, medium term HRQoL, need for repeat pleural intervention, overall complications, infectious complications, or tumour seeding (Appendix 3). No other outcomes could be examined due to the absence of studies including doxycycline pleurodesis.

One RCT did not report IPC drainage protocol and could not be included in subgroup analysis of ICP drainage schedule (71). The remaining four RCTs were divided into scheduled drainage (i.e. any regular drainage schedule) and as needed (PRN) drainage (i.e. drainage for dyspnea only) subgroups. Three RCTs performed some form of scheduled drainage, while one performed only PRN drainage (Table 1). Providing data on patient-reported dyspnea, amongst two RCTs using scheduled drainage for IPC, there is a very small increase in improvement in dyspnea at both the medium-term (two RCTs) and long-term (one RCT) time points: MD 8.94 [95%CI 0.80 to 17.08, I2=0%, n=143] and MD 14.00 [95%CI 2.80 to 25.20, n=43], respectively (Appendix 4). In contrast, the single randomized trial using a PRN drainage protocol demonstrated a very small difference in medium-term and long term-dyspnea with very wide and imprecise confidence intervals suggesting the potential for minimal benefit and harm: MD -11.88 [95%CI -30.10 to 6.35, n=61] and MD -10.40 [95%CI -30.51 to 9.71, n=42], respectively. The test for interaction suggested a statistically significant subgroup effect for both medium-term dyspnea (Chi² = 4.18, df = 1 [P = 0.04], I² = 76.1%), and long-term dyspnea (Chi² = 4.32, df = 1 [P = 0.04], I² = 76.8%). For initial admission length of stay, a single RCT using scheduled drainage showed a larger decrease in days in hospital (MD -3.67 [95%CI -4.56 to -2.78], n=99) compared with a single RCT using PRN drainage (MD -2.00 [95%CI -2.25 to -1.75], n=144) and the test for interaction suggested a statistically significant subgroup effect (Chi² = 12.51, df = 1 [P = 0.0004], I² = 92.0%). Similarly, for effusion related days in hospital, in a single trial using scheduled drainage, evidence showed a larger decrease in days in hospital (MD -3.50 [95%CI -4.75 to -2.25], n=99) compared with a single trial using PRN drainage (MD -1.60 [95%CI -2.82 to -

0.38], n=144), and the test for interaction suggested a statistically significant subgroup

effect (Chi² = 4.54, df = 1 [P = 0.03], I² = 78.0%) In terms of overall complications,

amongst the two trials using scheduled drainage, there was a large increase in risk with

IPC (RR 3.74 [95%CI 2.31 to 6.08] I2=0%, n=240), whereas, in the single trial using

PRN drainage there was a small increase in risk with IPC (RR 1.27 [95%CI 0.82 to 1.96],

n=144) with a very wide and imprecise confidence intervals suggesting the potential for

minimal benefit and harm, and suggested a statistically significant subgroup effect (Chi²

= 10.62, df = 1 [P = 0.001], I² = 90.6%). There was no subgroup difference for

immediate-dyspnea, short-term dyspnea, immediate overall HRQoL, short-term overall

HRQoL, long-term overall HRQoL, need for repeat pleural intervention, serious

complications, or infectious complications. No other outcomes could be examined due to

the absence of studies using a PRN drainage protocol.

3.12 Sensitivity Analysis

Sensitivity analysis was performed removing higher risk of bias studies

(Appendix 5). For patient-reported dyspnea at the immediate, short-term, and medium-

term time points there was no significant heterogeneity between lower and higher risk of

bias studies. For long-term dyspnea, two RCTS were higher risk of bias and only one

RCT was lower risk of bias. At this time point, removal of higher risk of bias studies

resulted in a change in the mean difference from 2.71 [95%CI -14.18 to 19.60, I^2=58%]

to 14.00 [95%CI 2.80 to 25.20], with the test of interaction demonstrating a statistically

significant subgroup effect (Chi^2 = 0.68, df=1 [p=0.03], I^2=78%).

Exclusion of higher risk of bias studies had minimal impact on estimates of overall HRQoL at the short-term and medium-term time points. Sensitivity analysis on patient-reported overall HRQoL at the immediate and long-term time points could not be performed, as there were only lower risk of bias studies at the outcome level for the immediate time point and only higher risk of bias studies at the outcome level for the long-term time point.

No studies were considered higher risk of bias for the outcomes of: need for repeat pleural intervention, complication rate (overall, serious, infectious, tumour seeding), or index admission length of stay.

Removal of one higher risk of bias trial for effusion related days in hospital left only one study one lower risk of bias trial. There was significant heterogeneity between lower and higher risk of bias studies (Chi2 = 4.54, df=1 [p=0.03], I2=78%) for this outcome. Removal of higher risk of bias studies resulted in a change of mean difference from -2.55 [95%CO -4.41 to -0.68, I^2=78%] to -1.60 [95%CI -2.82 to -0.38].

3.13 Publication Bias

None of our pooled analysis for our target outcomes met our a priori threshold of 10 studies for exploration of publication bias by funnel plot analysis.

CHAPTER 4: Discussion

4.1 Summary of Findings

In total, five RCTs (n= 557) and seven cohorts (n=988) met the eligibility criteria. Four randomized trials compared IPC to pleurodesis for patient-reported dyspnea using either the 100 mm VAS for dyspnea (VASD) (two trials) or the modified Borg scale (MBS) (two trials). The VASD is the most commonly used and familiar instrument for assessing dyspnea in patients with MPE. The instruments validity (face, construct, and criterion), responsiveness to change and intra-rater reliability as a measure of dyspnea has previously been established (79, 80). Further, there is evidence of interpretability with an available MID estimate in the MPE population (40). There was little to no difference between groups at the immediate, medium term, or long-term time points based on low to very low certainty evidence downgraded due to risk of bias and imprecision at each of these time points and inconsistency at the long-term time point (Table 4) (81). There was a trivial mean difference (MD) in favour of IPC based on low certainty of evidence downgraded due to imprecision at short-term follow up: MD 6.99 [95%CI 0.28 to 13.69, I^2=0%, n=309], with the upper end of the 95%CI not exceeding the MID estimate (19 mm) suggesting that very few patients would experience any noticeable difference in dyspnea compared with pleurodesis.

Three randomized trials compared IPC to pleurodesis for patient-reported overall HRQoL using the EQ5D3L, the modified EQ5D, and the CRQ respectively. Similarly, given the availability of an anchor-based MID (0. 1) in a lung cancer population and evidence of interpretability, reliability, validity, and responsiveness, results were converted to the natural units of the EQ5D3L (39, 82-85). There was little to no

difference at immediate, short-term, and medium-term overall HRQoL based on low and very low certainty evidence downgraded due to imprecision at all time points and further downgraded due to risk of bias at the medium-term and long-term time points (Table 5). At long-term follow-up there was a very small difference favouring pleurodesis based on very low certainty evidence: MD -0.06 [95%CI -0.17 to 0.05, I^2=0%, n=79]. Certainty in this estimate was downgraded due to risk of bias, and imprecision. While the lower end of the 95%CI surpasses the 0.1 MID, the CI is very wide and imprecise, suggesting the potential for both minimal benefit and harm. Similarly, at long-term follow-up results reported as an SMD and MID units showed a very small difference (MID units -0.60 [95%CI -1.69 to 0.49]), suggesting a consistency of results regardless of the statistical presentation method.

Only individual studies examined patient-reported pain (1 RCT, n=106) and patient satisfaction with care (1 prospective cohort, n=104). Mean difference estimates generated from individual trials demonstrated little to no difference between groups based on very low certainty of evidence (Table 6). Confidence in all estimates was lowered due to imprecision. No data was available to generate mean difference estimates for patient-reported overall functional status.

Four randomized trials assessing need for repeat pleural intervention (for reaccumulation of pleural effusion) demonstrated a large relative risk (RR) decrease in patients treated with IPC based on moderate certainty evidence: RR 0.32 [95%CI 0.18 to 0.55] (Table 7) together with a corresponding ARR of 15.2% [95%CI 10.0% to 18.3%]. The certainty of evidence was downgraded due to imprecision.

Three randomized trials showed a large relative risk increase for overall complications in patients treated with IPC based on moderate certainty evidence: RR 2.48 [95%CI 1.10 to 5.60 I^2=82%] (Table 7) and ARI of 28.1% [95%CI 1.9% to 81.0%]. The evidence was downgraded to low certainty due to inconsistency and imprecision.

Four trials demonstrated an increased relative risk of serious complications in patients treating with IPC based on low certainty evidence. The evidence was downgraded due to imprecision: RR 1.53 [95%CI 0.83to 2.82, I^2=8%, n=396] corresponding to an ARI of 4.8% [95%CI -1.5% to 16.4%].

Four randomized trials examining infectious complications demonstrated a large RR increase in patients treated with IPC based on high certainty evidence: RR 4.15 [95%CI 1.61 to 10.74 I^2=0%, n=475] and ARI 7.4% [95%CI 1.5% to 22.9%]. The certainty of evidence for infectious complications was downgraded due to imprecision.

Two trials reported an increased relative risk of tumour seeding in IPC patients based on low certainty evidence which was downgraded due to imprecision: RR 3.24[95%CI 0.37 to 28.04] and ARI of 2.8%.

Based on two trials, indwelling pleural catheter resulted in shorter index admission length of stay based on low certainty evidence (MD -2.55 [95%CI -4.41 to -0.68, I^2=92%, n=243]). The certainty in this evidence was downgraded due to inconsistency and imprecision (Table 8). Very low certainty evidence from one RCT demonstrated a greater than three-day decrease in total hospital length of stay: MD -3.6 [95%CI -8.29 to 1.09]. The certainty in this estimate was downgraded due to imprecision. Similarly, low certainty evidence suggests there may be a decrease in effusion related

days spent in hospital favoring IPC: MD -2.78 [95%CI -4.41 to -1.15]. The evidence was downgraded due to risk of bias, imprecision, and inconsistency.

As a whole, the evidence suggests a trivial benefit to IPC for patient-reported dyspnea at the short-term time point and a trivial to very small benefit to pleurodesis for patient-reported overall HRQoL at the long-term time point based on low and very-low certainty evidence respectively. There are is little to no difference at all other time points for these outcomes, and little to no difference in other PROs based on low to very low certainty evidence. There is moderate certainty evidence that there is a lower risk of repeat pleural intervention (for reaccumulation of pleural effusion) but a higher risk of infectious complications for IPC. There is low to very low certainty evidence that there is a higher risk of overall, serious, and tumour seeding complications in IPC but shorter length of stay in all categories (initial admission, total days, effusion related days).

4.2 Comparison to Existing Reviews and Ongoing RCTs

While previous single studies have generally reported no difference in patient-reported dyspnea between IPC and pleurodesis in MPE, the analysis in existing SRMAs has been overly simplistic (Chapter 1.2). There are three SRMAs that have pooled data from two trials measuring dyspnea using the VASD at one or two arbitrary time points (19, 20, 29). Another SRMA by Iyer et al. (2019) states that "pooled data for improvement in Borg score did not show any difference between groups", yet no estimates or data are reported and no time point is specified (31). This study found a trivial to very small mean difference (MD) in favour of IPC for patient-reported dyspnea, however, the upper end of the 95%CI did not exceed the MID estimate of 19 mm

suggesting that very few patients would experience any noticeable difference in dyspnea compared with pleurodesis.

Similarly, two previous SRMAs report on overall HRQoL when comparing IPC to pleurodesis but again their analysis was very limited (Chapter 1.2). Yeung et al. (2020) only describe the findings as presented in primary studies without additional analysis (29). Wang et al. (2020) pooled a mean EORTC-QLQ-C30 score from one RCT at a single time point (six weeks) with mean EQ5D score from another single RCT at a different single time point (12 months) and presented the pooled estimate as a SMD (-1.50 [95%CI -3.8 to 0.8, I^2=98%]). Our pooled estimate of two trials (n=79) reported as a MD (-0.06 [95%CI -0.17 to 0.05, I^2=0%]) in the natural units of the EQ5D3L utility index suggests that there is a very small difference (just over one half of a MID) at long-term follow-up favouring pleurodesis, however the CI is wide and imprecise, suggesting the potential for minimal benefit and harm. If there is truly a small difference in overall HRQoL at the long-term time point, it may be related to inconvenience of IPC including the out of pocket patient costs and lifestyle modifications associated with IPC. Aboudara et al. (2020) performed a cross sectional survey of American patients with IPC for MPE. Amongst 20 patients, eleven (55%) had additional costs associated with the IPC, four (20%) had significant life changes (i.e. downsizing due to cost or moving in with family due to inability to live independently), seventeen (85%) required assistance from a non-paid caregiver, six (30%) could not do activities because of the IPC; this negatively impacted overall HRQoL as measured by the Center for Disease Control Health Related Quality of Life 4 (CDC-HRQOL-4) questionnaire in three (15%) of the 20 patients surveyed (86).

No previously published SRMAs reported on other PROs and our review was only able to identify one study each reporting on patient-reported pain and patient-reported satisfaction with little to no difference between IPC and pleurodesis for these outcomes.

Previous SRMAs report a similar decrease in relative risk for repeat pleural interventions and decreased length of stay amongst IPC patients compared with pleurodesis patients (19, 20, 31). Iyer et al. (2019) also report an increased risk of infectious complications in IPC patients: pleural infections (RR 3.32 [95%CI 0.82-13.44]) and cellulitis (RR 5.83 [95%CI 1.56, 21.87]) (31). Wang et al. (2020) reported no difference in the incidence of overall adverse events but the data used in their meta-analysis is inaccurate (19). They reported adverse events in 45/48 pleurodesis patients and 43/46 IPC patients from one randomized trial (71). These numbers are actually the number of patients for which adverse event data is available. The actual estimates are 7/45 (pleurodesis) and 8/43 (IPC).

Previous SRMAs report similar decreases in length of stay in hospital in IPC patients compared with pleurodesis patients. Yeung et al. (2020) report decreased initial admission LOS (MD 2.19 [0.70-3.67, I^2=68%]) based on three trials (29). Wang et al. (2020) state that only one paper presented data on LOS, however, an additional study identified in their systematic review also reports on LOS but this data was missed in their review (19). Dipper et al. (2020) describe decreased length of stay based on two individual studies but did not conduct a meta-analysis (20). Iyer et al. (2019) also describe decreased LOS based on four individual studies but did not meta-analyze results due to multiple studies not reporting SD data (31). This study addressed this lack of

standard variance data by using validated methods for calculating SD based on reported variance data in similar studies across two RCTs (45), finding a decreased initial admission length of stay, decreased total days in hospital, and decreased effusion related length of stay in IPC patients compared with pleurodesis.

The approach to treating MPE has recently expanded to include the use of IPC *in combination w*ith talc pleurodesis. A randomized trial comparing IPC to IPC with talc demonstrated that administration of talc through the indwelling pleural catheter resulted in higher pleurodesis rate without increase in adverse events or catheter blockage (30). A randomized trial comparing talc pleurodesis alone to IPC with talc pleurodesis through the catheter recently finished recruiting (87). The primary outcome of this study is overall HRQoL as measured by the EORTC QLQ-C30. Secondary outcomes include patient-reported dyspnea and pain, complication rate, and pleurodesis failure. Another trial comparing instillation of talc through an IPC to thoracoscopic talc poudrage is currently recruiting (88). The primary outcome for this study is needed for ipsilateral pleural re-intervention. Secondary outcomes include time to symptomatic recurrence, all-cause hospital days, patient-reported dyspnea and pain as measured by respective 100mm visual analogue scales, overall HRQoL measured by EQ5D and 100mm visual analogues scale, physical activity patterns, adverse events, overall survival, and pleural-related hospital days.

4.3 Limitations

Our study has several limitations. First, the number of included studies and patients is small and there is significant attrition due, in large part, to early death. The

39

limited number of patients has an impact on the precision of results, as reflected in having to frequently rate down the evidence due to wide 95%CIs and limited number of patients (between 79 and 300 patients for PROs, between 144 and 243 patients for other continuous outcomes, and between 240 and 475 for other dichotomous outcomes).

Second, while attrition rates were high, especially for longer-term outcomes, loss to follow up due to death is a random variable unrelated to the target outcomes, so this was not considered an important risk of bias issue and criteria for assessing risk of bias due to incomplete outcome data was adjusted accordingly.

Third, due to inconsistent and limited reporting of patient-reported outcomes and variance data, while imputations or graphical extraction were used when possible, a number of primary studies could not be included in the quantitative analysis. Access to unpublished primary data would have allowed for a more complete and meaningful meta-analysis. Unfortunately, authors did not respond to multiple requests for data.

Fourth, due to the nature of the treatment options, blinding is not possible. Because of this, the Cochrane RoB 2.0 tool was felt to be too punitive as blinding would necessarily be judged as high risk of bias and therefore every study and outcome would also be judged as high risk of bias (89). Further, the Cochrane RoB 2.0 is designed with an explanatory framework in mind, adding to its punitive nature (90). The trials included in this review were, for the most part, pragmatic in nature. Therefore, the McMaster CLARITY instruments for both RCTs and cohort studies were used. These instruments include most of the same domains but are less punitive, allowing the user to make their own thresholds for lower versus higher risk of bias. It was decided that a higher risk of bias RCTs had to have at least two of six categories, and higher risk of bias cohort studies

had to have at least three of eight categories, with responses of 'probably no' or 'definitely no' (Chapter 2.8). As there is no guidance document for determining overall risk of bias, the cut offs for higher and lower risk of bias designations were arbitrarily chosen by study investigators based on what was felt to be reasonable given the outcomes and interventions examined.

Fifth, some studies reported change data for patient-reported dyspnea and patient-reported HRQoL, while others reported end-point data (18, 21, 27, 71). The Cochrane Handbook suggests that combined change and end-point data should not be meta-analyzed as a SMD (47). Contrarily, it has been demonstrated that there is little to no relevant difference between end-point and change data SMDs and combining these estimates in meta-analysis is valid as long as the decision on which type of data to use when both are available is pre-specified in the meta-analysis protocol (42). This study prioritized change data when both were available. The combination of both change and end-point was used for two outcomes, across four time points, allowing for improved precision of each estimate of effect.

Sixth, amongst two studies using the MBS for assessment of dyspnea, patients were asked to complete the MBS "at rest" and "on exertion" (21, 71). The dyspnea score "on exertion" was the only outcome selected for inclusion in meta-analysis. In this patient population "on exertion" is considered walking 100 feet on the level (21, 91). This measure was selected as the MBS is most commonly administered with a stimulus such as exercise. Additionally, this measurement was felt to likely be more responsive to small changes in dyspnea (38).

Seventh, in order to generate MD estimates in MID units for patient-reported overall HRQoL, a recently introduced meta-analytic method, MID estimates were necessary for all instruments (92). An anchor based MID estimate was available for the EQ5D3L utility index score. While anchor-based MID estimates, thought to be the most valid estimates, were available at the level of individual domain for the Chronic Respiratory Questionnaire (CRQ), they were not available for the cumulative score (37, 93, 94). Further, no anchor-based MID estimate was available for modified EQ5D used in the Thomas et al. trial (27). Therefore a distribution based method was used to determine the MID for the CRQ (2.565) and modified EQ5D (5.055) respectively (54). Subsequently, data from the RCT reporting CRQ and the RCT reporting modified EQ5D were converted to EQ5D3L utility index and the trials were pooled using recently established methods (53).

Eighth, as all included studies contained patients with trapped lung and data was not presented separately within studies, this could not be explored subgroup analysis. Patients with trapped lung would not benefit from chemical pleurodesis and would necessarily require another intervention. In fact, in several studies, patients with trapped lung randomized to the pleurodesis group did not receive talc instillation and therefore their initial treatment consisted only of chest tube placement (18, 71).

Ninth, due to the limited number of studies and the characteristics of included studies, it was not possible to perform the majority of our planned subgroup analysis (95). The subgroup analysis that was performed was underpowered and only represents early observational data, however, it does highlight some important issues. In examining the drainage schedule for IPC, there was an increased benefit in patient-reported dyspnea at

medium-term and long-term time points amongst studies that used regularly scheduled

drainage of IPC rather than PRN drainage. Additionally, the benefit for IPC in reducing

length of stay was greater in studies using regularly scheduled drainage. Conversely, the

increase in risk of overall complications associated with IPC was more pronounced

amongst studies using regularly scheduled drainage. Of note, no two studies utilized the

same drainage schedule (Table 1). These findings need to be verified in future studies,

but highlight the importance of optimizing and standardizing treatment protocols *within*

treatment modalities.

Tenth, similarly due to the limited number of studies, publication bias could not

be adequately assessed. Based on the exhaustive search of the literature (five primary

databases and three grey literature sources) and the fact that most trials were academic

investigator initiated, publication bias is not thought to be a major issue.

Eleventh, need for repeat pleural intervention is a misleading outcome. While it is

widely reported in the literature, it accounts only for pleural interventions aimed at

evacuating recurrent pleural effusion. Notably, it does not include repeat interventions for

infectious complications, which are more common amongst patients with IPC than

amongst patients who underwent pleurodesis. Additionally, the interventions for drainage

of recurrent effusion are likely to be better tolerated than interventions required for

pleural space infections or infected pleural catheters. For example, recurrent pleural

effusion might be managed simply with a thoracentesis, whereas a pleural infection from

a pleural catheter would require both removal of the catheter and possible operative

drainage of the infected space. The absence of data on interventions for indications other

than recurrent effusion is a significant limitation and systematically biases the findings of this outcome in favour of IPC.

Finally, the instruments used for assessment of patient-reported dyspnea and HRQoL have serious limitations. Patient-report dyspnea using the VASD is the measure that has the most evidence for its use measuring dyspnea in MPE patients. While there is evidence of validity, responsiveness, intra-rater reliability, and interpretability with in the MPE population its format and administration is highly variable (79) (79, 80) (40). Standardization is required for multiple facets of the VASD. Specifically, standardization of the questions asked, anchoring phrases, marking of the scale (i.e. the addition of reference points to the 100mm line), directionality of the scale (i.e. what does a lower versus higher score mean), and horizontal/vertical orientation of the scale are required (79). The modified EQ5D and EQ5D3L assess outcomes across 5 domains (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression) (96). This instrument has been shown to be acceptable, reliable, responsive, and valid across a number of other health conditions (82-85, 97). It is a valuable tool for comparing MPE patients to the general population or examining within group changes in overall HRQoL. However, it may not be a discriminatory instrument in comparing differences between treatments, particularly in MPE patients. Future work using a standardized, disease-specific VASD as well as generic quality of life instrument (such as the EQ5D3L) are required in the MPE population. In addition to these instruments, investigators should consider goal attainment scaling using instruments such as Measure Yourself Medical Outcome Profile (MYMOP) (98-100).

4.4 Strengths

Our study has strengths. First, unlike previous studies, our PRO estimates were presented at multiple pre-determined time points. Differences in patient-reported outcome data from different time points likely represent different stages of the disease process. Additionally, presentation of outcomes at one arbitrarily selected time point may be misleading if there is a change over time, or could reflect selective outcome reporting.

Second, although the number of studies and patients followed was limited, established methods were used for pooling PRO instruments and calculating missing variance data to allow for pooling of data from multiple PRO instruments (36, 44, 45). The use of these imputation methods increased the number of patients for the pooled estimates, thereby increasing the precision of the results.

Third, to improve the interpretability of our results, emerging methods were used to convert and present the PRO data using the most common, established instrument, and subsequently interpret the results based on estimates of the MD (relative to the MID) and MID units, when possible. Data was then presented using multiple statistical presentation formats (i.e. SMD, MD in natural units of the most familiar instrument, and MID units) (37, 53). Compared to SMD, reporting the MD relative the MID estimate bas been shown to improve the perceived usefulness of summary data to clinicians (50, 101). While MID units are a new and emerging statistical presentation method that clinicians and patients are unfamiliar with, with further educational initiatives, MID units could become a mainstay when studies that measure the same construct do so using different instruments (53). Importantly, for patient-reported overall HRQoL, the estimates using multiple methods were consistent with respect to the magnitude of effect, showing a very small

effects favoring pleurodesis at the long-term time point and little to no difference at all other time points.

While the presentation of PRO data in the units of the most familiar instrument is a strength of this study, natural unit presentation relative to MID can still be misleading. Mean differences for short-term dyspnea and long-term overall HRQoL were less than one MID and effect size was interpreted as trivial to very small. However, it is important to remember that a subgroup of these patients may experience a benefit that exceeds the MID. For example, Johnston et al. (2013) previously demonstrated that while there was only a mean difference of 2.5 units in the Hamilton Rating Scale for Depression (in which the MID is 7.0) between patients treated with paroxetine compared with placebo, this translated into a difference of the proportion of patients benefiting in experimental and control groups of 9.2% (37).

Fourth, the inclusion of both RCTs and cohort studies allowed for a more comprehensive assessment of available PRO data, data that represents the most patient-important outcomes to those suffering from a terminal diagnosis of cancer, for which neither IPC or pleurodesis has been shown to prolong life. The cohort studies, while lower quality and prone to confounding, provided PRO data for satisfaction with care and overall functional status. Unfortunately, there was limited reporting of data for other PROs (dyspnea, overall HRQoL, and pain) and these could not be meta-analyzed for cohort studies.

Finally, the inclusion of cohort studies in our systematic review allowed for the extraction of data that was previously unreported in one of the included RCTs. The identification of the Olfert et al. (2017) paper allowed us to examine patient-reported

overall HRQoL data at each time point for the Davies et al. (2012) RCT (18, 69). While the original Davies et al. (2012) RCT reported HRQoL data at 6 weeks only, we were able to obtain patient-reported HRQoL scores at for each time point (i.e. immediate, short-term, medium-term, and long-term) using the data subsequently reported in the Olfert et al. (2017) cost analysis.

4.5 Implications for Practice

There remains a lack of robust, comprehensive, clear assessments of patient-reported outcomes in patients with MPE. This study suggests there may be a trivial to very small improvement in dyspnea in IPC patients at short-term follow up and, conversely, there may be a very small improvement in overall HRQoL in pleurodesis patients at long-term follow-up. Neither of these differences is likely to be important in a substantial number of patients. The standardization and discriminatory capability of currently available instruments is a major limitation to interpreting the results of available PRO data.

For non-PRO outcomes, this study found that shorter hospital LOS and lower need for repeat pleural intervention (for reaccumulation of pleural effusion) with IPC. This however must be balanced with the higher risk of overall, serious, infectious, and tumour seeding complications.

In the absence of data suggesting an important difference in PROs, the decision as to which treatment modality to use should be based on patient values and preferences, however these value sensitive decisions rely on clinicians fully informing patients of the benefits, harms, inconveniences and costs of IPC versus pleurodesis based on all

47

outcomes, ideally starting with GRADE summary of findings tables. The importance of these outcomes will vary by patient.

4.6 Implications for Research

Ongoing RCTs continue to compare various treatment modalities for MPE. These include a trial comparing talc pleurodesis alone to IPC with talc pleurodesis through the catheter and another trial comparing instillation of talc through an IPC to thoracoscopic talc poudrage (87, 88).

Although the interventions have changed in these trials, the outcome measurements remain the same. As demonstrated by this study, despite the presence of multiple randomized controlled trials comparing IPC to pleurodesis, evidence is limited by issues of measurement (e.g. non-standardized measurement of PRO data), reporting, and the ability to interpret data obtained by currently used PRO instruments.

Additionally, there is a lack of standardization of the treatments themselves. Our limited subgroup analysis suggests that the drainage schedule of IPC may be important in patient-reported dyspnea, complication rate, and length of stay. However, the subgroup analysis was not powered adequately to assess this. The optimal drainage schedule for IPC requires further investigation, ideally in the form of a RCT.

In addition to work on standardization and evaluation of the measurement properties of PRO outcomes such as dyspnea and overall HRQoL specific to patients with malignant pleural effusions, investigators should also consider using goal attainment scaling as a part of a battery of tests on this population. Further, as with other fields, the development of a core outcome set by an international group of patients and clinicians

would serve to standardize the outcomes, timing and procedures for measuring outcomes in future RCTs and cohort studies (102). With this, those conducting evidence synthesis will be better equipped to summarize and interpret the totality of evidence across a set of outcomes that are important to patients. Finally, work on patient health-related values and preferences based on the current dataset should be encouraged. This could include, for example, a focus group of patients wherein the best estimates of benefit and harm are shared with patients and they are asked what treatment they would prefer. Such information could then be used to inform clinical practice guidelines that follow GRADE and Guideline International Network standards. The improved development of PRO instruments and a core outcome set highlighting patient-important factors specific to patients with malignant pleural effusions together with data on patient values and preferences would be invaluable for future research and clinical decision-making.

Table 1: Characteristics of included studies

Author, year (ref)	Study design	Country	Sample size		Sex (%Female)	Cancer types	Intervention details	Include trapped lung?	Patient-reported outcomes	Other outcomes
			IPC	P						
Boshuizen et al. 2017 (71)	RCT	Netherlands	46	48	IPC: 58.7% P: 43.8%	Lung: 33% Breast: 21% Mesothelioma: NS Other: 46%	IPC: Seldinger PleurX, drainage schedule NS P: Talc slurry, 15-20Fr, protocol NS	Yes	**Dyspnea:** Improvement in MBS (mean first 2 weeks, at 6 weeks, 3 months, and 6 months)	Need for pleural re-intervention; complications; number of hospital visits (NR); patient survival <6 weeks
Davies et al. 2012* (18)	RCT	United Kingdom	52	54	IPC: 56% P: 57%	Lung: 24% Breast: 26% Mesothelioma: 20% Other: 40%	IPC: Seldinger PleurX, drained 3x/week and as needed for dyspnea P: Talc slurry, 12Fr, protocol NS	Yes	**Dyspnea:** mean VASD over first 42 days; Proportion of patients achieving decrease in dyspnea by 10mm on VASD over first 42 days; VASD score at 6 weeks, 3 months, and 6 months; **Overall HRQoL:** EORTC-QLQ-30 at 6 weeks, 3 months, and 6 months; EQ5D3L scores at week 2, 4, 6, 10, 14, 18, 22, 26, at 9 months, and 1 year; **Pain:** mean VASC at 6 weeks, 3 months, and 6 months;	LOS randomization to discharge; all-cause mortality to 1 year; complications and serious complications
Demmy et al. 2012 (17)	RCT	United States	33	34	IPC: 39% P: 45%	Lung: 63% Breast: 12% Mesothelioma: 0% Other: 25%	IPC: Seldinger PleurX, drained once daily until volume <30cc over 3 consecutive days P: Talc slurry, ≥24Fr, removed tube when <150cc/24hrs	Yes	**Dyspnea:** "Dyspnea Index" at 7 and 30 days **Overall HRQoL:** MSAS at 7 and 30 days **Overall functional status:** Karnofsky self-reported performance rating at 7 and 30 days	Combined "success" at 30 days including: 1) alive, 2) lung re-expansion ≥90%, 3) "completion of intervention" by 2 weeks, 4) removal of chest tube for P or proper function of IPC

Author, year (ref)	Study design	Country	Sample size IPC	Sample size P	Sex (%Female)	Cancer types	Intervention details	Include trapped lung?	Patient-reported outcomes	Other outcomes
Putnam et al. 1999 (21)	RCT	United States	99	45	IPC: 58% P: 58%	Lung: 40% Breast: 27% Mesothelioma: NS Other: 33%	IPC: Seldinger, PleurX, drained 1.5L then 1L q8hours until all effusion drained, then 'completely' every 2 days. P: Doxycycline via chest tube, size NS, removed tube when <100cc/24hrs, if >4 days, re-administered doxycycline	Yes	**Dyspnea:** Improvement in MBS immediately post-procedure and at 30, 60, 90 days **Overall HRQoL:** Improvement in CRQ at 30, 60, 90 days	Spontaneous pleurodesis in IPC group; "Late failure" in P group; complications; median survival; LOS from randomization until "eligible for discharge"
Thomas et al. 2017 (27)	RCT	Australia, New Zealand, Singapore, Hong Kong	74	72	IPC: 41% P: 40%	Lung: 33% Breast: 12% Mesothelioma: 26% Other: 29%	IPC: Seldinger PleurX, drained only when symptomatic P: Talc slurry, 12-18Fr, protocol NS	Yes	**Dyspnea:** VASD at 2 weeks and 1, 3, 6, 9, 12 months **Overall HRQoL:** modified EQ5D at 8 and 14 days, and 1, 3, 6, 9, and 12 months; VAS-QoL for first 14 days then at 1, 3, 6, 9, and 12 months;	Effusion related days in hospital; need for repeat pleural intervention; survival; complications and serious complications; total days spent in hospital from intervention to death or 12 month follow up
De Abreu et al. 2019 (73)	RC	Brazil	18	34	IPC: 63.1% P: 79.4%	Lung: 46% Breast: 27% Mesothelioma: NS Other: 27%	IPC: Beside LunGO, drained every 5 days or when symptomatic P: 75% talc poudrage, 25% talc slurry, size NS, protocol NS	Yes, - all IPC patients	-	Need for repeat pleural-intervention; complications; length of stay; MPE recurrence;
Freeman et al. 2013 (72)	RC	United States	30	30	IPC: 60% P: 60%	Lung: 33% Breast: 37% Mesothelioma: NS Other: 30%	IPC: Thoracoscopy PleurX, suction on pleurevac until POD#1 then drainage when symptomatic P: Talc poudrage, size NS, 48hrs suction, removed tube when <250cc/24hrs	Yes	**Overall functional status:** ECOG performance status change (unclear if patient-reported or clinician assessed)	Need for repeat pleural intervention; hospital length of stay; interval from surgery to systemic treatment; complications; operative mortality; mean survival; rate of TPC removal;

Author, year (ref)	Study design	Country	Sample size		Sex (%Female)	Cancer types	Intervention details	Include trapped lung?	Patient-reported outcomes	Other outcomes
			IPC	P						
Fysh et al. 2012 (34)	PC	Australia	34	31	IPC: 26.5% P: 38.7%	Lung: 18% Breast: 17% Mesothelioma: 46% Other: 18%	IPC: Seldinger PleurX, drainage when symptomatic P: Mixed talc slurry/poudrage, ratio NS, 12-15Fr for slurry, 24-32Fr for poudrage, drainage protocol as per treating physician	Yes	**Dyspnea:** VASD daily for 1 week **Overall HRQoL:** VAS-QoL daily for 1 week	LOS (total days in hospital, effusion related hospital days); admissions to hospital; effusion related admissions; "control of effusion"; complications; change in albumin and protein;
Hunt et al. 2012 (74)	RC	United States	59	50	IPC: 64% P: 54%	Lung: 39% Breast: 14% Mesothelioma: 18% Other: 19%	IPC: PleurX, 20% thoracoscopy, 80% Seldinger P: Talc poudrage, size NS, 48 hours suction, removed tube when <200cc/24hrs	NS	-	LOS (total and post-procedure); complications; in-hospital mortality; repeat pleural intervention; re-admission for effusion
Liou et al. 2016 (75)	RC	United States	79	159	IPC: 69.6% P: 58.5%	Lung: 47% Breast: 18% Mesothelioma: NS Other: 35%	IPC: Type, insertion, and drainage protocol NS P: Talc poudrage, size NS, protocol NS	NS	-	LOS (overall hospital, post-procedure); repeat intervention; complications; disposition from hospital; readmissions; ICU admission; survival;
Srour et al. 2013 (76)	RC	Canada	193	167	IPC: 56.0% P: 70.7%	Lung: 43% Breast: 24% Mesothelioma: 4% Other: 29%	IPC: Seldinger PleurX, drainage 3x/week P: Talc slurry, size NS, protocol NS	NS	-	Pleural effusion control; freedom from pleural effusion and catheter; need from subsequent intervention; survival; effusion-free survival; adverse events;

Author, year (ref)	Study design	Country	Sample size		Sex (%Female)	Cancer types	Intervention details	Include trapped lung?	Patient-reported outcomes	Other outcomes
			IPC	P						
Walker et al. 2016 (35)	PC	Canada	47	57	Overall: 64%	Lung: 43% Breast: 31% Mesothelioma: NS Other: 26%	IPC: Tenchkoff, 43% thoracoscopy, 57% bedside, drainage schedule NS P: 32% talc poudrage, 68% talc slurry, size at discretion of physician, protocol NS	NS	**Dyspnea:** LCADL at 2 and 6 weeks **Overall HRQoL:** FACIT-Pal at 2 and 6 weeks **Pain:** Numerical rating scale (unclear collection time) **Satisfaction with care:** FACIT-TS at 2 and 6 weeks	LOS; Complications;

*Includes data published in Olfert et al. 2017 and Penz et al. 2013

CRQ: Guyatt Chronic Respiratory Questionnaire, ECOG: Eastern Cooperative Oncology Group, EORTC-QLQ-30: European Quality of Life Questionnaire Core-30, EQ5D: European Quality of Life Group 5-dimension questionnaire, EQ5D3L: European Quality of Life 5 Dimension 3-Level Utility Index, FACIT-Pal: Functional assessment of chronic illness therapy – palliative, FACIT-TS: Functional assessment of chronic illness therapy – treatment satisfaction, Fr: French, IPC: Indwelling pleural catheter, IQR: Interquartile range, LCADL: London chest activity of daily living scale, MBS: Modified Borg Scale, MSAS: Memorial Symptom Assessment –Short Form, MPE: Malignant pleural effusion, NA: Not assessed, NR: Not reported, NS: Not specified, P: Pleurodesis, PC: prospective cohort, RC: Retrospective cohort, RCT: Randomized controlled trial, VASC: 100mm visual analogue scale for chest pain, VASD: 100mm visual analogue scale for dyspnea, VAS-QoL: 100mm visual analogue scale for quality of life,

Table 2: Risk of bias of included randomized trials

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	PN	DY	PY	Higher
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Demmy 2012	DY	DY	DN	PN	PY	PY	Higher
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	DY	DY	PY	Lower

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Table 3: Risk of bias of included cohort studies

Study	Cohort selection	Assessment of exposure	Outcome not present at start	Matching or adjustment for prognostic variables	Assessment of prognostic factors	Assessment of outcome	Adequate follow-up	Similar co-interventions	Overall risk of bias
De Abreu 2019	DN	DY	DY	DN	PY	PN	PY	PY	Higher
Freeman 2013	PN	DY	DY	DY	PY	DY	DY	PY	Lower
Fysh 2012	PN	DY	DY	DN	PY	PY	PY	PY	Lower
Hunt 2012	DN	DY	DY	DN	PN	PY	PN	PY	Higher
Liou 2016	DN	DY	DY	DN	PY	PY	PN	PY	Higher
Srour 2013	DN	DY	DY	PN	PN	PY	PN	PN	Higher
Walker 2016	PN	DY	DY	DN	PY	PY	PY	PY	Lower

Assessed using the McMaster CLARITY instrument for assessing risk of bias in cohort studies.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Table 4: Summary of findings – Indwelling pleural catheter vs. pleurodesis for malignant pleural effusion (Improvement in patient-reported dyspnea)

Outcome	Estimated risk with pleurodesis*	Absolute change in risk with IPC (95%CI)**	Number of participants (studies)	Confidence in effect estimate[1]	Comments
Immediate dyspnea *	-	The mean improvement in dyspnea score in the IPC group was on average **1.84 higher** (3.82 lower to 7.51 higher)	305 (3 RCTs)	⊕◯◯◯ VERY LOW [a,c]	The evidence suggests that IPC results in little to no difference in improvement in immediate dyspnea compared with pleurodesis.
Short-term dyspnea *	-	The mean improvement in dyspnea score in the IPC group was on average **6.99 higher** (0.28 higher to 13.69 higher)	309 (4 RCTs)	⊕⊕◯◯ LOW [c]	The evidence suggests IPC may result in a trivial improvement in short-term dyspnea compared with pleurodesis.
Medium-term dyspnea *	-	The mean improvement in dyspnea score in the IPC group was on average **4.78 higher** (3.95 lower to 13.51 higher)	188 (4 RCTs)	⊕◯◯◯ VERY LOW [b,c]	The evidence suggests that IPC results in little to no difference in improvement in medium-term dyspnea compared with pleurodesis.
Long-term dyspnea *	-	The mean improvement in dyspnea score in the IPC group was on average **2.71 higher** (14.18 lower to 19.6 higher)	97 (3 RCTs)	⊕◯◯◯ VERY LOW [a,c,d]	The evidence suggests that IPC results in little to no difference in improvement in long-term dyspnea compared with pleurodesis.

* Based on the mean control group risk from all included trials. The risk with pleurodesis for dyspnea cannot be presented as meta-analysis was performed solely on the basis on mean difference in order to include the maximum number of trials
** Based on the assumed risk in the comparison group and the relative effect of the intervention (95% CI).
*** Dyspnea is reported on the VASD which ranges from 0 (maximum breathlessness) to 100 (no breathlessness) with an anchor-based MID estimate of 19mm among a MPE population
a. Serious concern for risk of bias due to lack of blinding and high/unequal loss to follow up in some studies
b. Very serious concern for risk of bias due to lack of blinding and high/unequal loss follow up in all studies
c. Very serious concerns for imprecision due to small sample size and wide confidence interval
c. Serious concern for inconsistency for moderate heterogeneity (I^2=58%)
CI: Confidence interval; IPC: Indwelling pleural catheter, MID: Minimally important difference, MPE: Malignant pleural effusion, RCT: Randomized controlled trial, VASD: 100mm visual analogue scale for dyspnea
1. GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Table 5: Summary of findings – Indwelling pleural catheter vs. pleurodesis for malignant pleural effusion (Patient-reported overall HRQoL)

Outcome	Presentation method	Estimated risk with pleurodesis*	Absolute change in risk with IPC (95%CI) **	Number of participants (studies)	Confidence in effect estimate[1]	Comments
Immediate overall HRQoL	**SD units (SMD)**	The overall HRQoL score in the IPC group was on average **0.09SDs higher** (0.19 lower to 0.38 higher) than in the pleurodesis group		195 (2 RCTs)	⊕⊕◯◯ LOW [a]	The evidence suggests that IPC results in little to no difference in immediate overall HRQoL compared with pleurodesis.
	Natural units ***	The mean overall HRQoL scores with pleiurodesis ranged from **0.629 to 0.680**	The mean overall HRQoL scores in the IPC group was on average **0.03 higher** (0.05 lower to 0.1 higher)		⊕⊕◯◯ LOW [a]	
	MID units	The mean overall HRQoL scores in the IPC group was on average **0.259 MID units higher** (0.482 lower to 1.001 higher) than the pleurodesis group			⊕⊕◯◯ LOW [a]	
Short-term overall HRQoL	**SD units (SMD)**	The overall HRQoL score in the IPC group was on average **0.04SDs lower** (0.29 lower to 0.21 higher) than in the pleurodesis group		266 (3 RCTs)	⊕⊕◯◯ LOW [a]	The evidence suggests that IPC results in little to no difference in short-term overall HRQoL compared with pleurodesis.
	Natural units ***	The mean overall HRQoL scores with pleiurodesis ranged from **0.681 to 0.700**	The mean overall HRQoL scores in the IPC group was on average **0.01 lower** (0.07 lower to 0.05 higher)		⊕⊕◯◯ LOW [a]	
	MID units	The mean overall HRQoL scores in the IPC group was on average **0.134 MID units higher** (0.529 lower to 0.796 higher) than the pleurodesis group			⊕⊕◯◯ LOW [a]	

Outcome	Presentation method	Estimated risk with pleurodesis*	Absolute change in risk with IPC (95%CI) **	Number of participants (studies)	Confidence in effect estimate[1]	Comments
Medium-term overall HRQoL	**SD units (SMD)**	The overall HRQoL score in the IPC group was on average **0.18 SDs lower** (0.55 lower to 0.19 higher) than in the pleurodesis group		115 (2 RCTs)	⊕◯◯◯ VERY LOW [a,b]	The evidence suggests that IPC results in little to no difference in medium-term overall HRQoL compared with pleurodesis.
	Natural units *	The mean overall HRQoL scores with pleurodesis was **0.706**	The mean overall HRQoL scores in the IPC group was on average **0.04 lower** (0.11 lower to 0.04 higher)		⊕◯◯◯ VERY LOW [a,b]	
	MID units	The mean overall HRQoL scores in the IPC group was on average **0.437 MID units lower** (1.132 lower to 0.438 higher) than the pleurodesis group			⊕◯◯◯ VERY LOW [a,b]	
Long-term overall HRQoL	**SD units (SMD)**	The overall HRQoL score in the IPC group was on average **0.24 SDs lower** (0.68 lower to 0.21 higher) than in the pleurodesis group		79 (2 RCTs)	⊕◯◯◯ VERY LOW [a,c]	The evidence suggests that IPC might result in a very small decrease in long-term overall HRQoL compared with pleurodesis.
	Natural units *	The mean overall HRQoL scores with pleiurodesis ranged from **0.650 to 0.788**	The mean overall HRQoL scores in the IPC group was on average **0.06 lower** (0.17 lower to 0.05 higher)		⊕◯◯◯ VERY LOW [a,c]	
	MID units	The mean overall HRQoL scores in the IPC group was on average **0.598 MID units lower** (1.690 lower to 0.491 higher) than the pleurodesis group			⊕◯◯◯ VERY LOW [a,c]	

* Based on the mean control group risk from all included trials

** Based on the assumed risk in the comparison group and the relative effect of the intervention (95% CI).

***Overall HRQoL measured on the EQ5D3L ranges from 0 (death) to 1 (perfect health) with an anchor based MID estimate of 0.1 among lung cancer patients

a. Very serious concern for imprecision due to small sanple size and wide confidence interval

b. Serious concern for risk of bias due to lack of bliniding and high/unequal loss to follow up in some studies

c. Very serious concern for risk of bias due to lack of blinding and high/unequal loss to follow up in all studies

CI: Confidence interval, EQ5D3L: European Quality of Life 5 Dimension 3-Level, HRQoL: Health related quality of life, IPC: Indwelling pleural catheter, MD: Mean difference, MID: Minimally important difference, RCT: Randomized controlled trial, SD: Standard deviation, SMD: Standardized mean difference

1.GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Table 6: Summary of findings – Indwelling pleural catheter vs. pleurodesis for malignant pleural effusion (Other patient-reported outcomes)

Outcome		Estimated risk with pleurodesis*	Absolute change in risk with IPC (95%CI) **	Number of participants (studies)	Confidence in effect estimate[1]	Comments
Patient-reported pain ***		The mean improvement in patient-reported pain was **4.4**	The mean improvement in patient-reported pain in the IPC group was on average **3.8 higher** (7.32 lower to 14.92 higher)	106 (1 RCT)	⊕◯◯◯ VERY LOW [a,b]	The evidence suggests that IPC results in little to no difference in patient-reported pain
Treatment satisfaction	Immediate	The mean treatment satisfaction immediate was **72.2**	The mean treatment satisfaction score in the IPC group was **2.53 higher** (11.26 lower to 16.31 higher)	104 (1 cohort study)	⊕◯◯◯ VERY LOW [b]	The evidence suggests that IPC results in little to no difference in immediate treatment satisfaction
	Short-term	The mean treatment satisfaction short-term was **79.0**	The mean treatment satisfaction in the IPC group was on average **3.27 higher** (8.06 lower to 14.61 higher)	104 (1 cohort study)	⊕◯◯◯ VERY LOW [b]	The evidence suggests that IPC results in little to no difference in short-term treatment satisfaction

* Based on the mean control group risk from all included trials
** Based on the assumed risk in the comparison group and the relative effect of the intervention (95% CI).
*** Patient-reported pain measured on 100mm visual analogue scale for chest pain ranging from 0 (no chest pain) to 100 (maximal chest pain)
a. Certainty of evidence starts at moderate due to presence of single trial only
b. Very serious concern for imprecision due to single study, small sample size, wide confidence interval spanning clinical decision threshold
CI: Confidence interval, IPC: Indwelling pleural catheter, MD: Mean difference, RCT: Randomized controlled trial
1.GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Table 7: Summary of findings – Indwelling pleural catheter vs. pleurodesis for malignant pleural effusion (Compications and need for re-intervention)

Outcome	Estimated risk with pleurodesis*	Absolute change in risk with IPC (95%CI) **	Relative effect (95%CI)	Number of participants (studies)	Confidence in effect estimate[1]	Comments
Need for repeat pleural intervention	223 per 1,000	**71 per 1,000** (40 to 123)	**RR 0.32** (0.18 to 0.55)	472 (4 RCTs)	⊕⊕⊕◯ MODERATE [a]	The evidence suggests IPC likely results in a large reduction in need for repeat pleural procedure.
Overall complications	190 per 1,000	**471 per 1,000** (209 to 1,000)	**RR 2.48** (1.10 to 5.60)	384 (3 RCTs)	⊕◯◯◯ VERY LOW [b,c]	The evidence suggests IPC likely results in a large increase in overall complications.
Serious complications	90 per 1,000	**138 per 1,000** (75 to 254)	**RR 1.53** (0.83 to 2.82)	396 (4 RCTs)	⊕⊕◯◯ LOW [b]	The evidence suggests IPC may increase serious complications.
Infectious complications	23 per 1,000	**97 per 1,000** (38 to 252)	**RR 4.15** (1.61 to 10.74)	475 (4 RCTs)	⊕⊕⊕◯ MODERATE [a]	The evidence suggests IPC results in a large increase in infectious complications.
Tumour seeding	0 per 1,000***	**0 per 1,000** (0 to 0)***	**RR 3.24 ***** (0.37 to 28.04)	240 (2 RCTs)	⊕⊕◯◯ LOW [b]	The evidence suggests IPC likely results in a large increase in tumour seeding.

* Based on the mean control group risk from all included trials
** Based on the assumed risk in the comparison group and the relative effect of the intervention (95% CI).
*** Relative risk based on Revman assumption imputation into zero count cells. True rate of tumour seeding was 0/97 in pleurodesis group and 4/143 in IPC group.
a. Serious concern for imprecision due to small sample size and wide confidence interval
b. Very serious concern for imprecision due to very small sample size and wide confidence interval
c. Serious concern for inconsistency due to considerable heterogeneity (I^2=82%)
CI: Confidence interval, IPC: Indwelling pleural catheter, RCT: Randomized controlled trial, RR: Relative risk
1.GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Table 8: Summary of findings – Indwelling pleural catheter vs. pleurodesis for malignant pleural effusion (Length of stay)

Outcome	Estimated risk with pleurodesis*	Absolute change in risk with IPC (95%CI) **	Number of participants (studies)	Confidence in effect estimate[1]	Comments
Initial admission	The mean initial admission length of stay ranged from **3.3 to 4 days**	The mean initial admission length of stay in the IPC group was on average **2.78 days lower** (4.41 lower to 1.15 lower)	243 (2 RCTs)	⊕⊕◯◯ LOW [a,c]	IPC likely results in a reduction in initial admission.
Total days in hospital	The mean total days in hospital was **16.3 days**	The mean total days in hospital in the IPC group was on average **3.6 days lower** (8.29 lower to 1.09 higher)	144 (1 RCT)	⊕◯◯◯ VERY LOW [a,b]	IPC likely results in a reduction in total days in hospital.
Effusion related days in hospital	The mean effusion related days in hospital ranged from **4.7 to 4.8 days**	The mean effusion related days in hospital in the IPC group was on average **2.55 lower** (4.41 lower to 0.68 lower)	243 (2 RCTs)	⊕⊕◯◯ LOW [a,c]	IPC may reduce effusion related days in hospital.

* Based on the mean control group risk from all included trials
** Based on the assumed risk in the comparison group and the relative effect of the intervention (95% CI).
a. Very serious concern for imprecision due to very small sample size and wide confidence interval
b. Certainty of evidence starts at moderate due to presence of single trial only
c. Serious concern for inconsistency due to considerable heterogeneity
CI: Confidence interval, IPC: Indwelling pleural catheter, RCT: Randomized controlled trial
1.GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Figure 1: Summary of evidence search and selection

Records identified through primary databases (n = 1894) MEDLINE: 445 EMBASE: 728 SCOPUS: 220 Cochrane: 264 Web of Science: 237	Additional records identified through other sources (n = 47)

Records after duplicates removed
(n = 1362)

Records screened
(n = 1362)

Records excluded
(n = 1314)

Full-text articles assessed for eligibility
(n= 48)

Full-text articles excluded
(n = 32)

n = 25 study design not eligible
n = 7 intervention not eligible
n = 1 did not report an eligible outcome

16 studies identified in systematic review

(n= 6 RCTs)
(n=10 cohort studies)

Excluded
(n = 5)

n = 2 IPC in both arms
n= 2 studies presenting data collected during already included RCT

12 studies for possible inclusion in quantitative synthesis.

(n= 5 RCTs)
(n= 7 cohort studies)

Figure 2: Effect of indwelling pleural catheter vs. pleurodesis on patient-reported dyspnea presented as mean differences in natural units of the100mm visual analogue scale for dyspnea*

a. Immediate (≤2 weeks)
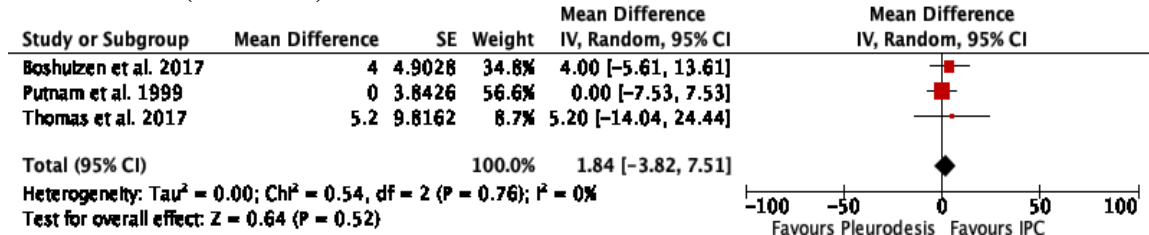
b. Short-term (>2 weeks, ≤6 weeks)
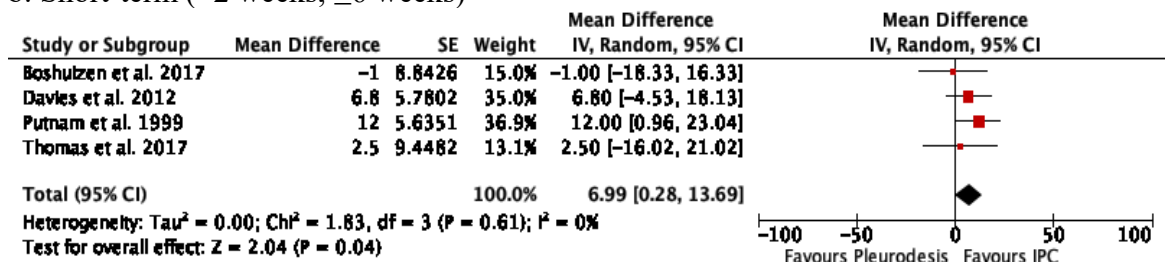
c. Medium-term (>6 weeks, <6 months)

d. Long-term (≥6 months)

*A positive mean difference indicates a greater improvement in dyspnea score in IPC patients
IPC: Indwelling pleural catheter, IV: Inverse variance

Figure 3: Effect of indwelling pleural catheter vs. pleurodesis on patient-reported overall health related quality of life presented as mean differences in the natural units of the European Quality of Life 5 Dimension 3-Level utility index*

a. Immediate (≤2 weeks),

Study or Subgroup	IPC Mean	SD	Total	Pleurodesis Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Olfert et al. 2017	0.656	0.2	45	0.629	0.25	51	68.3%	0.03 [-0.06, 0.12]
Thomas et al. 2017	0.71	0.33	46	0.68	0.34	53	31.7%	0.03 [-0.10, 0.16]
Total (95% CI)			**91**			**104**	**100.0%**	**0.03 [-0.05, 0.10]**

Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.97); I² = 0%
Test for overall effect: Z = 0.74 (P = 0.46)

b. Short-term (>2 weeks, ≤6 weeks)

Study or Subgroup	IPC Mean	SD	Total	Pleurodesis Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Olfert et al. 2017	0.658	0.22	40	0.681	0.21	43	44.9%	-0.02 [-0.12, 0.07]
Putnam et al. 1999	0.01	0.25	27	0.01	0.21	59	32.8%	0.00 [-0.11, 0.11]
Thomas et al. 2017	0.69	0.32	44	0.7	0.34	53	22.3%	-0.01 [-0.14, 0.12]
Total (95% CI)			**111**			**155**	**100.0%**	**-0.01 [-0.07, 0.05]**

Heterogeneity: Tau² = 0.00; Chi² = 0.10, df = 2 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.40 (P = 0.69)

c. Medium-term (>6 weeks, <6 months)

Study or Subgroup	IPC Mean	SD	Total	Pleurodesis Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Olfert et al. 2017	0.659	0.21	31	0.706	0.2	30	57.5%	-0.05 [-0.15, 0.06]
Putnam et al. 1999	0.04	0.23	21	0.06	0.2	33	42.5%	-0.02 [-0.14, 0.10]
Total (95% CI)			**52**			**63**	**100.0%**	**-0.04 [-0.11, 0.04]**

Heterogeneity: Tau² = 0.00; Chi² = 0.11, df = 1 (P = 0.74); I² = 0%
Test for overall effect: Z = 0.89 (P = 0.37)

d. Long-term (≥6 months)

Study or Subgroup	IPC Mean	SD	Total	Pleurodesis Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Olfert et al. 2017	0.674	0.32	18	0.788	0.16	18	44.0%	-0.11 [-0.28, 0.05]
Thomas et al. 2017	0.63	0.24	20	0.65	0.25	23	56.0%	-0.02 [-0.17, 0.13]
Total (95% CI)			**38**			**41**	**100.0%**	**-0.06 [-0.17, 0.05]**

Heterogeneity: Tau² = 0.00; Chi² = 0.70, df = 1 (P = 0.40); I² = 0%
Test for overall effect: Z = 1.10 (P = 0.27)

*A positive mean difference indicates higher overall HRQoL in IPC patients
IPC: Indwelling pleural catheter, IV: Inverse Variance

Note: Olfert 2017 is a study presenting unpublished overall HRQoL data from the Davies et al. 2012 RCT.

Figure 4: Effect of indwelling pleural catheter vs. pleurodesis on risk of repeat intervention

Study or Subgroup	IPC Events	Total	Pleurodesis Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Boshuizen et al. 2017	7	43	15	45	47.4%	0.49 [0.22, 1.08]
Davies et al. 2012	3	51	12	52	20.6%	0.25 [0.08, 0.85]
Putnam et al. 1999	2	94	4	43	10.9%	0.23 [0.04, 1.20]
Thomas et al. 2017	3	73	16	71	21.1%	0.18 [0.06, 0.60]
Total (95% CI)		261		211	100.0%	0.32 [0.18, 0.55]
Total events	15		47			

Heterogeneity: Tau² = 0.00; Chi² = 2.30, df = 3 (P = 0.51); I² = 0%
Test for overall effect: Z = 4.09 (P < 0.0001)

IPC: Indwelling pleural catheter, M-H: Mantel-Haenszel

Figure 5: Effect of indwelling pleural catheter vs. pleurodesis on risk of complications

a. Overall

Study or Subgroup	IPC Events	Total	Pleurodesis Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Davies et al. 2012	21	52	7	54	30.1%	3.12 [1.45, 6.70]
Putnam et al. 1999	22	43	11	91	33.1%	4.23 [2.26, 7.92]
Thomas et al. 2017	30	73	23	71	36.8%	1.27 [0.82, 1.96]
Total (95% CI)		168		216	100.0%	2.48 [1.10, 5.60]
Total events	73		41			

Heterogeneity: Tau² = 0.42; Chi² = 11.11, df = 2 (P = 0.004); I² = 82%
Test for overall effect: Z = 2.18 (P = 0.03)

Favours IPC   Favours Pleurodesis

b. Serious

Study or Subgroup	IPC Events	Total	Pleurodesis Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Boshuizen et al. 2017	8	43	7	45	38.0%	1.20 [0.47, 3.01]
Davies et al. 2012	9	52	5	54	31.7%	1.87 [0.67, 5.21]
Demmy et al. 2012	8	28	3	30	23.0%	2.86 [0.84, 9.71]
Thomas et al. 2017	1	73	3	71	7.3%	0.32 [0.03, 3.04]
Total (95% CI)		196		200	100.0%	1.53 [0.83, 2.82]
Total events	26		18			

Heterogeneity: Tau² = 0.03; Chi² = 3.27, df = 3 (P = 0.35); I² = 8%
Test for overall effect: Z = 1.36 (P = 0.17)

Favours IPC   Favours Pleurodesis

c. Infectious

Study or Subgroup	IPC Events	Total	Pleurodesis Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Boshuizen et al. 2017	2	43	2	45	24.6%	1.05 [0.15, 7.10]
Davies et al. 2012	13	52	2	54	43.6%	6.75 [1.60, 28.47]
Putnam et al. 1999	7	94	0	43	11.2%	6.95 [0.41, 118.95]
Thomas et al. 2017	6	73	1	71	20.6%	5.84 [0.72, 47.26]
Total (95% CI)		262		213	100.0%	4.15 [1.61, 10.74]
Total events	28		5			

Heterogeneity: Tau² = 0.00; Chi² = 2.70, df = 3 (P = 0.44); I² = 0%
Test for overall effect: Z = 2.94 (P = 0.003)

Favours IPC   Favours Pleurodesis

d. Tumour Seeding

Study or Subgroup	IPC Events	Total	Pleurodesis Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Davies et al. 2012	1	52	0	54	46.1%	3.11 [0.13, 74.74]
Putnam et al. 1999	3	91	0	43	53.9%	3.35 [0.18, 63.41]
Total (95% CI)		143		97	100.0%	3.24 [0.37, 28.04]
Total events	4		0			

Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.97); I² = 0%
Test for overall effect: Z = 1.07 (P = 0.29)

Favours IPC   Favours Pleurodesis

IPC: Indwelling pleural catheter, M-H: Mantel-Haenszel

Figure 6: Effect of indwelling pleural catheter vs. pleurodesis on length of stay in hospital.

a. Initial admission

	Experimental			Control				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Davies et al. 2012	0.33333333	0.76286553	51	4	3.05695439	48	46.6%	-3.67 [-4.56, -2.78]	
Thomas et al. 2017	1.33333333	0.75628711	73	3.33333333	0.75671456	71	53.4%	-2.00 [-2.25, -1.75]	
Total (95% CI)			124			119	100.0%	-2.78 [-4.41, -1.15]	

Heterogeneity: Tau² = 1.28; Chi² = 12.51, df = 1 (P = 0.0004); I² = 92%
Test for overall effect: Z = 3.34 (P = 0.0008)

Favours IPC   Favours Pleurodesis

b. Effusion related total days in hospital

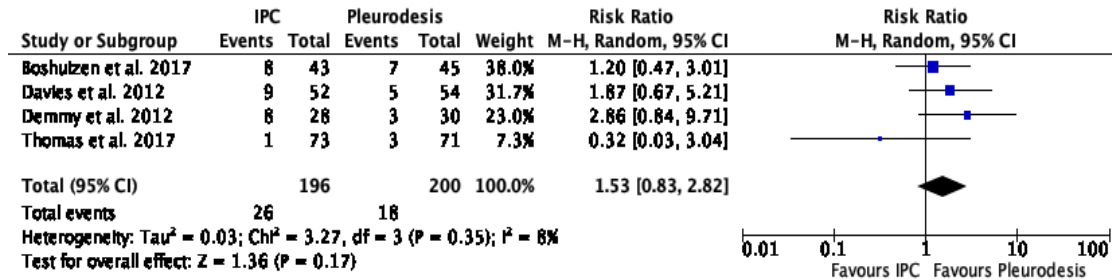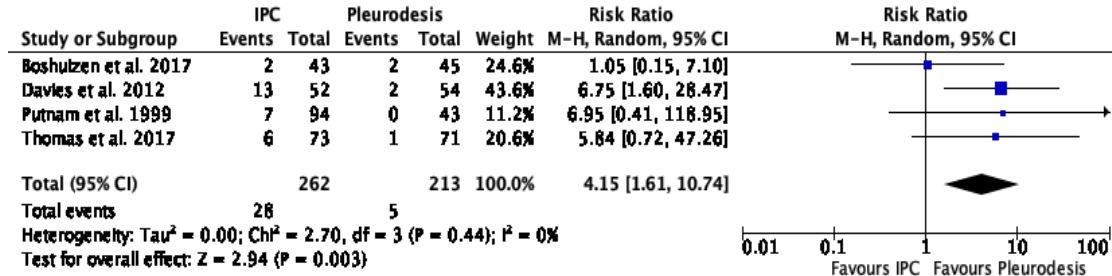|	Experimental			Control				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Davies et al. 2012	1.33333333	2.28859659	51	4.83333333	3.821193	48	49.7%	-3.50 [-4.75, -2.25]	
Thomas et al. 2017	3.1	4.3	73	4.7	3.1	71	50.3%	-1.60 [-2.82, -0.38]	
Total (95% CI)			124			119	100.0%	-2.55 [-4.41, -0.68]	

Heterogeneity: Tau² = 1.41; Chi² = 4.54, df = 1 (P = 0.03); I² = 78%
Test for overall effect: Z = 2.68 (P = 0.007)

Favours IPC   Favours Pleurodesis

IPC: Indwelling pleural catheter, M-H: Mantel-Haenszel

References

1. Rodriguez-Panadero F, Borderas Naranjo F, Lopez Mejias J. Pleural metastatic tumours and effusions. Frequency and pathogenic mechanisms in a post-mortem series. Eur Respir J. 1989;2(4):366-9.
2. Roberts ME, Neville E, Berrisford RG, Antunes G, Ali NJ, Group BTSPDG. Management of a malignant pleural effusion: British Thoracic Society Pleural Disease Guideline 2010. Thorax. 2010;65 Suppl 2:ii32-40.
3. American Thoracic S. Management of malignant pleural effusions. Am J Respir Crit Care Med. 2000;162(5):1987-2001.
4. Bhatnagar R, Luengo-Fernandez R, Kahan BC, Rahman NM, Miller RF, Maskell NA. Thoracoscopy and talc poudrage compared with intercostal drainage and talc slurry infusion to manage malignant pleural effusion: The tapps rct. Health Technology Assessment. 2020;24(26):1-89.
5. Marel M, Zrustova M, Stasny B, Light RW. The incidence of pleural effusion in a well-defined region. Epidemiologic study in central Bohemia. Chest. 1993;104(5):1486-9.
6. Feller-Kopman DJ, Reddy CB, DeCamp MM, Diekemper RL, Gould MK, Henry T, et al. Management of Malignant Pleural Effusions. An Official ATS/STS/STR Clinical Practice Guideline. Am J Respir Crit Care Med. 2018;198(7):839-49.
7. Laisaar T, Palmiste V, Vooder T, Umbleja T. Life expectancy of patients with malignant pleural effusion treated with video-assisted thoracoscopic talc pleurodesis. Interact Cardiovasc Thorac Surg. 2006;5(3):307-10.
8. Zamboni MM, da Silva CT, Jr., Baretta R, Cunha ET, Cardoso GP. Important prognostic factors for survival in patients with malignant pleural effusion. BMC Pulm Med. 2015;15:29.
9. Clive AO, Kahan BC, Hooper CE, Bhatnagar R, Morley AJ, Zahan-Evans N, et al. Predicting survival in malignant pleural effusion: development and validation of the LENT prognostic score. Thorax. 2014;69(12):1098-104.
10. Gadewad N, Deokar K, Ghorpade S. Clinical Profile of Patients Presenting with Malignant Pleural Effusion to a Tertiary Health Care Centre. J Assoc Physicians India. 2017;65(8):28-31.
11. Kvale PA, Selecky PA, Prakash UB, American College of Chest P. Palliative care in lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):368S-403S.
12. Antony VB. Pathogenesis of malignant pleural effusions and talc pleurodesis. Pneumologie. 1999;53(10):493-8.
13. Goodman A, Davies CW. Efficacy of short-term versus long-term chest tube drainage following talc slurry pleurodesis in patients with malignant pleural effusions: a randomised trial. Lung cancer (Amsterdam, Netherlands). 2006;54(1):51-5.
14. Ozkul S, Turna A, Demirkaya A, Aksoy B, Kaynak K. Rapid pleurodesis is an outpatient alternative in patients with malignant pleural effusions: a prospective randomized controlled trial. Journal of thoracic disease. 2014;6(12):1731-5.
15. Mineo TC, Sellitri F, Tacconi F, Ambrogi V. Quality of life and outcomes after nonintubated versus intubated video-thoracoscopic pleurodesis for malignant pleural effusion: comparison by a case-matched study. Journal of palliative medicine. 2014;17(7):761-8.
16. Gonzalez AV, Bezwada V, Beamis JF, Jr., Villanueva AG. Lung injury following thoracoscopic talc insufflation: experience of a single North American center. Chest. 2010;137(6):1375-81.
17. Demmy TL, Gu L, Burkhalter JE, Toloza EM, D'Amico TA, Sutherland S, et al. Optimal management of malignant pleural effusions (results of CALGB 30102). JNCCN Journal of the National Comprehensive Cancer Network. 2012;10(8):975-82.

18.	Davies HE, Mishra EK, Kahan BC, Wrightson JM, Stanton AE, Guhan A, et al. Effect of an indwelling pleural catheter vs chest tube and talc pleurodesis for relieving dyspnea in patients with malignant pleural effusion: the TIME2 randomized controlled trial. JAMA. 2012;307(22):2383-9.

19.	Wang L, Deng H, Chen X, Li C, Yi F, Wei Y, et al. Talc pleurodesis versus indwelling pleural catheter among patients with malignant pleural effusion: A meta-analysis of randomized controlled trials. World Journal of Surgical Oncology. 2020;18(1):184.

20.	Dipper A, Jones HE, Bhatnagar R, Preston NJ, Maskell N, Clive AO. Interventions for the management of malignant pleural effusions: a network meta-analysis. The Cochrane database of systematic reviews. 2020;4:CD010529.

21.	Putnam JB, Jr., Light RW, Rodriguez RM, Ponn R, Olak J, Pollak JS, et al. A randomized comparison of indwelling pleural catheter and doxycycline pleurodesis in the management of malignant pleural effusions. Cancer. 1999;86(10):1992-9.

22.	Sidhu C, Louw A, Brims F, Creaney J, Lee YCG. Malignant Pleural Mesothelioma: an Update for Pulmonologists. Current Pulmonology Reports. 2019;8(2):40-9.

23.	Akram MJ, Khalid U, Ashraf M, Butt F, Khan F. INDICATIONS AND CLINICAL OUTCOMES OF INDWELLING PLEURAL CATHETER PLACEMENT IN PATIENTS WITH MALIGNANT PLEURAL EFFUSION IN A CANCER SETTING HOSPITAL. Chest. 2020;157(6 Supplement):A358.

24.	Janes SM, Rahman NM, Davies RJO, Lee YCG. Catheter-tract metastases associated with chronic indwelling pleural catheters. Chest. 2007;131(4):1232-4.

25.	Khalid T, Khan OU, Blaxill P. A review of the indwelling pleural catheter (IPC) service in a busy District General Hospital (DGH). Lung Cancer. 2020;139(Supplement 1):S85.

26.	Yan A, Li P, Williams D, Gillson A. Outcomes of indwelling pleural catheters in malignant pleural effusions: Do patients who are followed in a specialized pleural effusion clinic have better outcomes? Canadian Journal of Respiratory, Critical Care, and Sleep Medicine. 2019;3(Supplement 1):51-2.

27.	Thomas R, Fysh ETH, Smith NA, Lee P, Kwan BCH, Yap E, et al. Effect of an Indwelling Pleural Catheter vs Talc Pleurodesis on Hospitalization Days in Patients With Malignant Pleural Effusion: the AMPLE Randomized Clinical Trial. JAMA. 2017;318(19):1903-12.

28.	Boshuizen RC, vd Noort V, Burgers JA, Herder GJM, Hashemi SMS, Hiltermann TJN, et al. A randomized controlled trial comparing indwelling pleural catheters with talc pleurodesis (NVALT-14). Lung Cancer. 2017;108:9-14.

29.	Yeung M, Loh E-W, Tiong T-Y, Tam K-W. Indwelling pleural catheter versus talc pleurodesis for malignant pleural effusion: a meta-analysis. Clinical & experimental metastasis. 2020;37(4):541-9.

30.	Bhatnagar R, Keenan EK, Morley AJ, Kahan BC, Stanton AE, Haris M, et al. Outpatient Talc Administration by Indwelling Pleural Catheter for Malignant Effusion. New England journal of medicine. 2018;378(14):1313-22.

31.	Iyer NP, Reddy CB, Wahidi MM, Lewis SZ, Diekemper RL, Feller-Kopman D, et al. Indwelling Pleural Catheter versus Pleurodesis for Malignant Pleural Effusions. A Systematic Review and Meta-Analysis. Annals of the American Thoracic Society. 2019;16(1):124-31.

32.	Clive AO, Jones HE, Bhatnagar R, Preston NJ, Maskell N. Interventions for the management of malignant pleural effusions: a network meta-analysis. The Cochrane database of systematic reviews. 2016(5):CD010529.

33.	Sivakumar P, Saigal A, Ahmed L. A systematic review of interventions to improve health related quality of life in malignant pleural effusion. Thorax. 2017;72(Supplement 3):A211-A2.

34.   Fysh ETH, Waterer GW, Kendall PA, Bremner PR, Dina S, Geelhoed E, et al. Indwelling pleural catheters reduce inpatient days over pleurodesis for malignant pleural effusion. Chest. 2012;142(2):394-400.

35.   Walker S, Zubrinic M, Massey C, Shargall Y, Bedard E, Darling G. A prospective study of patient-centred outcomes in the management of malignant pleural effusions. International journal of palliative nursing. 2016;22(7):351-8.

36.   Johnston BC, Patrick DL, Busse JW, Schunemann HJ, Agarwal A, Guyatt GH. Patient-reported outcomes in meta-analyses--Part 1: assessing risk of bias and combining outcomes. Health Qual Life Outcomes. 2013;11:109.

37.   Johnston BC, Patrick DL, Thorlund K, Busse JW, da Costa BR, Schunemann HJ, et al. Patient-reported outcomes in meta-analyses-part 2: methods for improving interpretability for decision-makers. Health Qual Life Outcomes. 2013;11:211.

38.   Ries AL. Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. COPD. 2005;2(1):105-10.

39.   Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007;5:70.

40.   Mishra EK, Corcoran JP, Hallifax RJ, Stradling J, Maskell NA, Rahman NM. Defining the minimal important difference for the visual analogue scale assessing dyspnea in patients with malignant pleural effusions. PLoS ONE. 2015;10(4):e0123798.

41.   Giesinger JM, Kuijpers W, Young T, Tomaszewski KA, Friend E, Zabernigg A, et al. Thresholds for clinical importance for four key domains of the EORTC QLQ-C30: physical functioning, emotional functioning, fatigue and pain. Health Qual Life Outcomes. 2016;14:87.

42.   da Costa BR, Nuesch E, Rutjes AW, Johnston BC, Reichenbach S, Trelle S, et al. Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. J Clin Epidemiol. 2013;66(8):847-55.

43.   Review Manager Web (RevMan Web). The Cochrane Collaboration 2020.

44.   Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. J Clin Epidemiol. 2006;59(1):7-10.

45.   Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014;14:135.

46.   Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.

47.   Deeks JJ HJ, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

48.   Busse JW, Montori VM, Krasnik C, Patelis-Siotis I, Guyatt GH. Psychological intervention for premenstrual syndrome: a meta-analysis of randomized controlled trials. Psychother Psychosom. 2009;78(1):6-15.

49.   McKenzie JE BS, Ryan RE, Thomson HJ, Johnston RV. Chapter 9: Summarizing study characteristics and preparing for synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

50.   Johnston BC, Thorlund K, Schunemann HJ, Xie F, Murad MH, Montori VM, et al. Improving the interpretation of quality of life evidence in meta-analyses: the application of minimal important difference units. Health Qual Life Outcomes. 2010;8:116.

51.	Schunemann HJ, Guyatt GH. Commentary--goodbye M(C)ID! Hello MID, where do you come from? Health Serv Res. 2005;40(2):593-7.
52.	Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61(2):102-9.
53.	Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH. Pooling health-related quality of life outcomes in meta-analysis-a tutorial and review of methods for enhancing interpretability. Res Synth Methods. 2011;2(3):188-203.
54.	Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003;41(5):582-92.
55.	CLARITY group at McMaster university. Tools to assess risk of bias in randomized controlled trials. Available at:. https://wwwevidencepartnerscom/wp-content/uploads/2017/09/Tool-to-Assess-Risk-of-Bias-in-Randomized-Controlled-Trialspdf [accessed July 10, 2020].
56.	CLARITY group at McMaster university. Tools to assess risk of bias in cohort studeis. Available at:. https://wwwevidencepartnerscom/wp-content/uploads/2017/09/Tool-to-Assess-Risk-of-Bias-in-Cohort-Studiespdf [Accessed July 10,2020].
57.	Akl EA, Sun X, Busse JW, Johnston BC, Briel M, Mulla S, et al. Specific instructions for estimating unclearly reported blinding status in randomized trials were reliable and valid. J Clin Epidemiol. 2012;65(3):262-7.
58.	Lan RS, Lo SK, Chuang ML, Yang CT, Tsao TC, Lee CH. Elastance of the pleural space: a predictor for the outcome of pleurodesis in patients with malignant pleural effusion. Ann Intern Med. 1997;126(10):768-74.
59.	Stefani A, Natali P, Casali C, Morandi U. Talc poudrage versus talc slurry in the treatment of malignant pleural effusion. A prospective comparative study. Eur J Cardiothorac Surg. 2006;30(6):827-32.
60.	Shaw P, Agarwal R. Pleurodesis for malignant pleural effusions. Cochrane Database Syst Rev. 2004(1):CD002916.
61.	Dresler CM, Olak J, Herndon JE, 2nd, Richards WG, Scalzetti E, Fleishman SB, et al. Phase III intergroup study of talc poudrage vs talc slurry sclerosis for malignant pleural effusion. Chest. 2005;127(3):909-15.
62.	Shaw PHS, Agarwal R. Pleurodesis for malignant pleural effusions. Cochrane Database of Systematic Reviews. 2013;2013(11):CD002916.
63.	Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ. 2008;336(7644):601-5.
64.	Page MJ HJ, Sterne JAC. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.
65.	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;343:d4002.
66.	Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.
67.	Schünemann H BJ, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook. 2013.

68.   Penz E, Mishra E, Davies H, Manns B, Miller R, Rahman N. Comparing cost of indwelling pleural catheter vs talc pleurodesis for malignant pleural effusion. Chest [Internet]. 2014; 146(4):[991-1000 pp.]. Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/291/CN-01015291/frame.html.

69.   Olfert JA, Penz ED, Manns BJ, Mishra EK, Davies HE, Miller RF, et al. Cost-effectiveness of indwelling pleural catheter compared with talc in malignant pleural effusion. Respirology (carlton, vic). 2017;22(4):764-70.

70.   Frost N, Brunger M, Ruwwe-Glosenkamp C, Raspe M, Tessmer A, Temmesfeld-Wollbruck B, et al. Indwelling pleural catheters for malignancy-associated pleural effusion: report on a single centre's ten years of experience. BMC pulmonary medicine. 2019;19(1):232.

71.   Boshuizen RC, Vd Noort V, Burgers JA, Herder GJM, Hashemi SMS, Hiltermann TJN, et al. A randomized controlled trial comparing indwelling pleural catheters with talc pleurodesis (NVALT-14). Lung cancer (Amsterdam, Netherlands). 2017;108:9-14.

72.   Freeman RK, Ascioti AJ, Mahidhara RS. A propensity-matched comparison of pleurodesis or tunneled pleural catheter in patients undergoing diagnostic thoracoscopy for malignancy. The Annals of thoracic surgery. 2013;96(1):259-4.

73.   De Abreu IRLB, Biazzotto AFB, Santos MA, De Souza AAC, Abrao FC, Cavalcante MGC. The Low-Cost, Long-Term Indwelling Pleural Catheter (LunGO): An Inexpensive and Effective Alternative for the Management of Dyspnea Associated with Malignant Pleural Effusions. Respiration. 2019;98(2):151-6.

74.   Hunt BM, Farivar AS, Vallieres E, Louie BE, Aye RW, Flores EE, et al. Thoracoscopic talc versus tunneled pleural catheters for palliation of malignant pleural effusions. The Annals of thoracic surgery. 2012;94(4):1053-9.

75.   Liou DZ, Serna-Gallegos D, Chan JL, Borgella J, Akhmerov S, Soukiasian HJ. Survival Difference in Patients with Malignant Pleural Effusions Treated with Pleural Catheter or Talc Pleurodesis. The American surgeon. 2016;82(10):995-9.

76.   Srour N, Amjadi K, Forster A, Aaron S. Management of malignant pleural effusions with indwelling pleural catheters or talc pleurodesis. Canadian respiratory journal. 2013;20(2):106-10.

77.   Dahlberg GJ, Maldonado F, Chen H, Rickman O, Roller L, Walston C, et al. Minimal clinically important difference for chest discomfort in patients undergoing pleural interventions. BMJ Open Respir Res. 2020;7(1).

78.   Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5(6):649-55.

79.   Mishra EK. Measurement of breathlessness in patients with malignant pleural effusions. Current Pulmonology Reports. 2016;5(1):1-6.

80.   Mador MJ, Kufel TJ. Reproducibility of visual analog scale measurements of dyspnea in patients with chronic obstructive pulmonary disease. Am Rev Respir Dis. 1992;146(1):82-7.

81.   Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol. 2011;64(4):407-15.

82.   EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.

83.   Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ. 1998;316(7133):736-41.

84.   Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.

85.   van Agt HM, Essink-Bot ML, Krabbe PF, Bonsel GJ. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. Soc Sci Med. 1994;39(11):1537-44.

86.   Aboudara M, Roller L, Lentz R, Rickman OB, Gillaspie EA, Maldonado F. A Survey-Based Study of Patient-Centered Costs Associated With Indwelling Pleural Catheters. The American journal of hospice & palliative care. 2020:1049909120954810.

87.   Sivakumar P, Douiri A, West A, Rao D, Warwick G, Chen T, et al. OPTIMUM: a protocol for a multicentre randomised controlled trial comparing Out Patient Talc slurry via Indwelling pleural catheter for Malignant pleural effusion vs Usual inpatient Management. BMJ Open. 2016;6(10):e012795.

88.   NCT04322136. AMPLE-3: IPC plus talc vs VATS in management of malifnant pleural effusion. https://clinicaltrialsgov/show/NCT04322136. 2020.

89.   Higgins JPT SJ, Page MK, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized tiral. In: Higgins JPT, Thomas J, Chadler K, Cumpston M, Li T, Page MJ, Welch VA (editores). Cochrane Handbook for Systematic Reviews of Interventons version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook

90.   Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol. 2009;62(5):464-75.

91.   Reddy C, Ernst A, Lamb C, Feller-Kopman D. Rapid pleurodesis for malignant pleural effusions: a pilot study. Chest. 2011;139(6):1419-23.

92.   Johnston BC, Thorlund K, da Costa BR, Furukawa TA, Guyatt GH. New methods can extend the use of minimal important difference units in meta-analyses of continuous outcome measures. J Clin Epidemiol. 2012;65(8):817-26.

93.   Schunemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic respiratory disease questionnaire (CRQ). COPD. 2005;2(1):81-9.

94.   Jones PW, Beeh KM, Chapman KR, Decramer M, Mahler DA, Wedzicha JA. Minimal clinically important differences in pharmacological trials. Am J Respir Crit Care Med. 2014;189(3):250-5.

95.   Fu R GG, Grany M, Shamliyan T, Sedrakyan A, Wilt T, Griffith L, Oremus M, Raina P, Ismaila A, Santaguida P, Lau J, Trikalinos T. Conducting Quantitative Synthesis When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. 2010 Oct 25. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK49407/.

96.   Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013;22(7):1717-27.

97.   Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol. 1997;36(5):551-9.

98.   Paterson C, Langan CE, McKaig GA, Anderson PM, Maclaine GD, Rose LB, et al. Assessing patient outcomes in acute exacerbations of chronic bronchitis: the measure your medical outcome profile (MYMOP), medical outcomes study 6-item general health survey (MOS-6A) and EuroQol (EQ-5D). Qual Life Res. 2000;9(5):521-7.

99.   Paterson C. Measuring outcomes in primary care: a patient generated measure, MYMOP, compared with the SF-36 health survey. BMJ. 1996;312(7037):1016-20.

100.   Paterson C, Britten N. In pursuit of patient-centred outcomes: a qualitative evaluation of the 'Measure Yourself Medical Outcome Profile'. J Health Serv Res Policy. 2000;5(1):27-36.
101.   Johnston BC, Alonso-Coello P, Friedrich JO, Mustafa RA, Tikkinen KAO, Neumann I, et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. CMAJ. 2016;188(1):25-32.
102.   Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials. 2016;17(1):449.

Appendix 1: Search strategy

In November 2020, a comprehensive search of Medline, Embase, Scopus, Web of Science, and the Cochrane Library from inception to 2020 was performed irrespective of publication status or language.

Hand searches
Bibliographies for all studies chosen for full-text review as well as relevant review articles were searched for additional studies not identified by the electronic searches. The last three years of abstracts from the following society meetings were searched: the American Association for Thoracic Surgery (AATS), the European Society of Thoracic Surgeons (ESTS), American College of Chest Physicians (CHEST), and the International Association for the Study of Lung Cancer (IASLC).

Additional searches
We searched for ongoing trials and grey literature using OpenGrey.eu and clinicaltrials.gov.

Electronic Search strategies
The full electronic search strategies are listed below.

Medline
1. exp Pleural Effusion, Malignant/ **2.** (malig* adj10 pleural).tw. **3.** 1 or 2 **4.** (indwell* adj10 catheter*).tw. **5.** exp Catheters, Indwelling/ **6.** 4 or 5 **7.** 3 and 6 **8.** exp Pleurodesis/ **9.** (chemical adj10 pleurodesis).tw. **10.** exp Bleomycin/ **11.** exp Tetracycline/ **12.** exp Povidone-Iodine/ **13.** exp Talc/ **14.** exp Minocycline/ **15.** (bleomycin or tetracycline or "povidone-iodine" or talc or minocycline).tw. **16.** 8 or 10 or 11 or 12 or 13 or 14 **17.** 9 or 16 **18.** 16 or 17 **19.** 3 and 18 **20.** 7 or 19 [subject search] **21.** Randomized Controlled Trials as Topic/ **22.** randomized controlled trial/ **23.** Random Allocation/ **24.** Double Blind Method/ **25.** Single Blind Method/ **26.** clinical trial/ **27.** clinical trial, phase i.pt. **28.** clinical trial, phase ii.pt. **29.** clinical trial, phase iii.pt. **30.** clinical trial, phase iv.pt. **31.** controlled clinical trial.pt. **32.** randomized controlled trial.pt. **33.** multicenter study.pt. **34.** clinical trial.pt. **35.** exp Clinical Trials as topic/ **36.** or/21-35 **37.** (clinical adj trial$).tw. **38.** ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw. **39.** PLACEBOS/ **40.** placebo$.tw. **41.** randomly allocated.tw. **42.** (allocated adj2 random$).tw. **43.** or/37-42 **44.** 36 or 43 **45.** case report.tw. **46.** letter/ **47.** historical article/ **48.** or/45-47 **49.** 44 not 48 [SIGN rct hedge] **50.** 20 and 49 [pe strings and rct hedge] **51.** Epidemiologic studies/ **52.** exp case control studies/ **53.** exp cohort studies/ **54.** Case control.tw. **55.** (cohort adj (study or studies)).tw. **56.** Cohort analy$.tw. **57.** (Follow up adj (study or studies)).tw. **58.** (observational adj (study or studies)).tw. **59.** Longitudinal.tw. **60.** Retrospective.tw. **61.** Cross sectional.tw. **62.** Crossectional studies/ **63.** or/51-62 [observational study SIGN hedge] **64.** 20 and 63 **65.** Meta-Analysis as Topic/ **66.** meta analy$.tw. **67.** metaanaly$.tw. **68.** Meta-Analysis/ **69.** (systematic adj (review$1 or overview$1)).tw. **70.** exp Review Literature as Topic/ **71.** or/65-70 **72.** cochrane.ab. **73.** embase.ab. **74.** science citation index.ab. **75.** web of science.ab. **76.** scopus.ab. **77.** 72 or 73 or 74 or 75 or 76 **78.** reference list$.ab. **79.** bibliograph$.ab. **80.** hand-search$.ab. **81.** relevant

journals.ab. **82.** manual search$.ab. **83.** 78 or 79 or 80 or 81 or 82 **84.** selection criteria.ab. **85.** data extraction.ab. **86.** 84 or 85 **87.** Review/ **88.** 86 and 87 **89.** Comment/ **90.** Letter/ **91.** Editorial/ **92.** animal/ **93.** human/ **94.** 92 not (92 and 93) **95.** 89 or 90 or 91 or 94 **96.** 71 or 77 or 83 or 88 **97.** 96 not 95 [meta analysis SIGN hedge] **98.** 20 and 97 **99.** 50 or 64 **100.** 50 or 64 or 98

Embase
1. exp malignant pleura effusion/ **2.** (malignant adj10 pleural).tw. **3.** 1 or 2 **4.** (indwell* adj10 catheter*).tw. **5.** exp indwelling catheter/ **6.** 4 or 5 **7.** 3 and 6 **8.** exp pleurodesis/ **9.** (chemical adj10 pleurodesis).tw. **10.** exp Bleomycin/ **11.** exp Tetracycline/ **12.** exp Povidone-Iodine/ **13.** exp Talc/ **14.** exp Minocycline/ **15.** 8 or 10 or 11 or 12 or 13 or 14 **16.** (bleomycin or tetracycline or "povidone-iodine" or talc or minocycline).tw. **17.** 9 or 13 or 16 **18.** 15 or 17 **19.** 3 and 18 **20.** 7 or 19 [subject search] **21.** Clinical Trial/ **22.** Randomized Controlled Trial/ **23.** controlled clinical trial/ **24.** multicenter study/ **25.** Phase 3 clinical trial/ **26.** Phase 4 clinical trial/ **27.** exp RANDOMIZATION/ **28.** Single Blind Procedure/ **29.** Double Blind Procedure/ **30.** Crossover Procedure/ **31.** PLACEBO/ **32.** randomi?ed controlled trial$.tw. **33.** rct.tw. **34.** (random$ adj2 allocat$).tw. **35.** single blind$.tw. **36.** double blind$.tw. **37.** ((treble or triple) adj blind$).tw. **38.** placebo$.tw. **39.** Prospective Study/ **40.** 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 **41.** Case Study/ **42.** case report.tw. **43.** abstract report/ or letter/ **44.** Conference abstract.pt. **45.** Editorial.pt. **46.** Letter.pt. **47.** Note.pt. **48.** 41 or 42 or 43 or 44 or 45 or 46 or 47 **49.** 40 not 48 [rct SIGN hedge] **50.** 20 and 49 **51.** Clinical study/ **52.** Case control study/ **53.** Family study/ **54.** Longitudinal study/ **55.** Retrospective study/ **56.** Prospective study/ **57.** Randomized controlled trials/ **58.** 56 not 57 **59.** Cohort analysis/ **60.** (Cohort adj (study or studies)).tw. **61.** (Case control adj (study or studies)).tw. **62.** (follow up adj (study or studies)).tw. **63.** (observational adj (study or studies)).tw. **64.** (epidemiologic$ adj (study or studies)).tw. **65.** (cross sectional adj (study or studies)).tw. **66.** 51 or 52 or 53 or 54 or 55 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 [obs study SIGN hedge] **67.** 20 and 66 **68.** exp Meta Analysis/ **69.** ((meta adj analy$) or metaanalys$).tw. **70.** (systematic adj (review$1 or overview$1)).tw. **71.** 68 or 69 or 70 **72.** cochrane.ab. **73.** embase.ab. **74.** science citation index.ab. **75.** web of science.ab. **76.** scopus.ab. **77.** 72 or 73 or 74 or 75 or 76 **78.** reference lists.ab. **79.** bibliograph$.ab. **80.** hand-search$.ab. **81.** manual search$.ab. **82.** relevant journals.ab. **83.** 78 or 79 or 80 or 81 or 82 **84.** data extraction.ab. **85.** selection criteria.ab. **86.** 84 or 85 **87.** review.pt. **88.** 86 and 87 **89.** letter.pt. **90.** editorial.pt. **91.** animal/ **92.** human/ **93.** 91 not (91 and 92) **94.** 89 or 90 or 91 or 93 **95.** 71 or 77 or 83 or 88 **96.** 95 not 94 [sr SIGN hedge] **97.** 20 and 96 **98.** 50 or 67 **99.** 50 or 67 or 97

Scopus
1 TITLE-ABS (malign* W/10 pleural)
2 TITLE-ABS (indwell* W/10 catheter*)
3 (TITLE-ABS (malign* W/10 pleural)) AND (TITLE ABS (indwell* W/10 catheter*))
4 TITLE-ABS (chemical W/10 pleurodesis)
5 TITLE-ABS (bleomycin OR tetracycline OR "povidone iodine" OR talc OR minocycline)

6 (TITLE-ABS (bleomycin OR tetracycline OR "povidone

iodine" OR talc OR minocycline)) OR (TITLE-ABS (chemical W/10 pleurodesis))

7 (TITLE-ABS (malign* W/10 pleural)) AND ((TITLE-

ABS (bleomycin OR tetracycline OR "povidone-

iodine" OR talc OR minocycline)) OR (TITLE-ABS (chemical W/10 pleurodesis)))

8 ((TITLE-ABS (malign* W/10 pleural)) AND ((TITLE-

ABS (bleomycin OR tetracycline OR "povidone-

iodine" OR talc OR minocycline)) OR (TITLE-

ABS (chemical W/10 pleurodesis)))) OR ((TITLE-

ABS (malign* W/10 pleural)) AND (TITLE-ABS (indwell* W/10 catheter*)))

9PMID (1*) OR PMID (2*) OR PMID (3*) OR PMID (4*) OR PMID (5*) OR PMID (6*)

OR PMID (7*) OR PMID (8*) OR PMID (9*)

10 (((TITLE-ABS (malign* W/10 pleural)) AND ((TITLE-

ABS (bleomycin OR tetracycline OR "povidone-

iodine" OR talc OR minocycline)) OR (TITLE-

ABS (chemical W/10 pleurodesis)))) OR ((TITLE-

ABS (malign* W/10 pleural)) AND (TITLE-ABS (indwell* W/10 catheter*)))) AND

NOT (PMID (1*) OR PMID (2*) OR PMID (3*) OR PMID (4*) OR PMID (5*) OR PMID (

6*) OR PMID (7*) OR PMID (8*) OR PMID (9*)) View Less

11 (((TITLE-ABS (malign* W/10 pleural)) AND ((TITLE-

ABS (bleomycin OR tetracycline OR "povidone-

iodine" OR talc OR minocycline)) OR (TITLE-

ABS (chemical W/10 pleurodesis)))) OR ((TITLE-

ABS (malign* W/10 pleural)) AND (TITLE-ABS (indwell* W/10 catheter*)))) AND

NOT (PMID (1*) OR PMID (2*) OR PMID (3*) OR PMID (4*) OR PMID (5*) OR PMID (

6*) OR PMID (7*) OR PMID (8*) OR PMID (9*)) AND (LIMIT-

TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "cp")) View

Less

Web of Science

1. TOPIC: ((malign* NEAR/10 pleural))
    Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-
    EXPANDED, IC Timespan=All years
2. TOPIC: ((indwel* NEAR/10 catheter*))
    Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-
    EXPANDED, IC Timespan=All years
3. #2 AND #1
    Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-
    EXPANDED, IC Timespan=All years
4. TOPIC: ((chemical NEAR/10 pleurodesis))
    Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-
    EXPANDED, IC Timespan=All years
5. TOPIC: ((bleomycin OR tetracycline OR "povidone-iodine" OR talc OR minocycline))

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

6. #5 OR #4
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

7. #6 AND #1
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

8. #7 OR #3
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

9. PMID=1* OR PMID=2* OR PMID=3* OR PMID=4* OR PMID=5* OR PMID=6* OR PMID=7#9 * OR PMID=8* OR PMID=9*
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR EXPANDED, IC Timespan=All years

10. #8 NOT #9
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

11. #8 NOT #9
Refined by: DOCUMENT TYPES: (MEETING ABSTRACT OR ARTICLE OR REVIEW OR PROCEEDINGS PAPER)
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

12. #8 NOT #9
Refined by: DOCUMENT TYPES: (REVIEW)
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

13. #8 NOT #9
Refined by: DOCUMENT TYPES: (MEETING ABSTRACT OR ARTICLE OR PROCEEDINGS PAPER)
Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

Cochrane Library
1. malignant near/10 pleural
2. pleurodesis
3. indwell* near/10 catheter*
4. #2 or #3
5. #1 and #4

Appendix 2: Risk of bias assessments by outcome included in meta-analysis

Immediate dyspnea

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	PN	DY	PY	Higher
Putnam 1999	DY	DY	DN	PN	PY	PY	Higher
Thomas 2017	DY	DY	DN	DY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Short-term dyspnea

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	PN	DY	PY	Higher
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Medium-term dyspnea

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	DN	DY	PY	Higher
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	DN	PY	PY	Higher
Thomas 2017	DY	DY	DN	PN	DY	PY	Higher
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Very serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Long-term dyspnea

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	DN	DY	PY	Higher
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Thomas 2017	DY	DY	DN	PN	DY	PY	Higher
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Immediate overall HRQoL

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012 Olfert 2017	DY	DY	DN	DY	DY	PY	Lower
Thomas 2017	DY	DY	DN	DY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Short-term overall HRQoL

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012 Olfert 2017	DY	DY	DN	DY	DY	PY	Lower
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	PN	DY	PY	Higher
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Medium-term overall HRQoL

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012 Olfert 2017	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	DN	PY	PY	Higher
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Long-term overall HRQoL

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012 Olfert 2017	DY	DY	DN	DN	DY	PY	Higher
Thomas 2017	DY	DY	DN	DN	DY	PY	Higher
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Very serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Patient-reported pain

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Immediate treatment satisfaction

Study	Cohort selection	Assessment of exposure	Outcome not present at start	Matching or adjustment for prognostic variables	Assessment of prognostic factors	Assessment of outcome	Adequate follow-up	Similar co-interventions	Overall risk of bias
Walker 2016	PN	DY	DY	DN	PY	PY	PY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)									Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in cohort studies.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely N

Short-term treatment satisfaction

Study	Cohort selection	Assessment of exposure	Outcome not present at start	Matching or adjustment for prognostic variables	Assessment of prognostic factors	Assessment of outcome	Adequate follow-up	Similar co-interventions	Overall risk of bias
Walker 2016	PN	DY	DY	DN	PY	PY	PY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)									Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in cohort studies.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely N

Need for repeat pleural interventions

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	PY	DY	PY	Lower
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Overall complications

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Infectious complications

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	PY	DY	PY	Lower
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Serious complications

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Boshuizen 2017	DY	DY	DN	PY	DY	PY	Lower
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Demmy 2012	DY	DY	DN	PY	PY	PY	Lower
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Tumour Seeding

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Putnam 1999	DY	DY	DN	PY	PY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Length of stay - initial admission

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012	DY	DY	DN	PY	DY	PY	Lower
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Length of stay – total days in hospital (including readmission)

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Not serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Length of stay - effusion related days in hospital

Study	Random sequence	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other	Overall risk of bias
Davies 2012	DY	DY	DN	PY	PN	PY	Higher
Thomas 2017	DY	DY	DN	PY	DY	PY	Lower
Overall concern for risk of bias for outcome (Not serious, serious, very serious)							Serious

Assessed using the McMaster CLARITY instrument for assessing risk of bias in randomized controlled trials.
DY = 'Definitely Yes', PY = 'Probably Yes', PN = 'Probably No', DN = 'Definitely No'

Appendix 3: Subgroup analysis (talc vs. doxycycline)

Outcome	Talc	Doxycycline	Mean difference units	Mean difference [95%CI] Talc	Mean difference [95%CI] doxycycline	Test for interaction
Immediate dyspnea	Boshuizen 2017 Thomas 2017	Putnam 1999	VASD	4.24 [-4.36, 12.84] I^2=0%	0.00 [-7.53, 7.53]	Chi² = 0.53, df = 1 (P = 0.47) I^2 = 0%
Short-term dyspnea	Boshuizen 2017 Davies 2012 Thomas 2017	Putnam 1999	VASD	4.06 [-4.38, 12.50] I^2=0%	12.00 [0.96, 23.04]	Chi² = 1.25, df = 1 (P = 0.26) I^2 = 20.3%
Medium-term dyspnea	Boshuizen 2017 Davies 2012 Thomas 2017	Putnam 1999	VASD	2.10 [-10.95, 15.15] I^2=47%	9.00 [-3.70, 21.70]	Chi² = 0.55, df = 1 (P = 0.46) I^2 = 0%
Short-term HRQoL	*Olfert 2017 Thomas 2017	Putnam 1999	EQ5D3L	-0.02 [-0.09, 0.06] I^2=0%	0.00 [-0.11, 0.11]	Chi² = 0.08, df = 1 (P = 0.78) I^2 = 0%
Medium-term HRQoL	*Olfert 2017	Putnam 1999	EQ5D3L	-0.05 [-0.15, 0.06]	-0.02 [-0.14, 0.10]	Chi² = 0.11, df = 1 (P = 0.74) I^2 = 0%
				Relative risk [95%CI] Talc	Relative risk [95%CI] doxycycline	
Need for repeat pleural intervention	Boshuizen 2017 Davies 2012 Thomas 2017	Putnam 1999	-	0.33 [0.18, 0.60] I^2=7%	0.23 [0.04, 1.20]	Chi² = 0.16, df = 1 (P = 0.69) I^2 = 0%
Overall complications	Davies 2012 Thomas 2017	Putnam 1999	-	1.88 [0.77, 4.56] I^2=76%	4.23 [2.26, 7.92]	Chi² = 2.15, df = 1 (P = 0.14) I^2 = 53.4%
Infectious complications	Boshuizen 2017 Davies 2012 Thomas 2017	Putnam 1999	-	3.76 [1.18, 11.97] I^2=22%	6.95 [0.41, 118.95]	Chi² = 0.15, df = 1 (P = 0.70) I^2 = 0%
Tumour seeding	Davies 2012	Putnam 1999	-	0.02 [-0.03, 0.07]	0.03 [-0.02, 0.08]	Chi² = 0.14, df = 1 (P = 0.71) I^2 = 0%

* Olfert 2017 is a study presenting unpublished overall HRQoL data from the Davies et al. 2012 RCT.

df: Degrees of freedom, EQ5D3L: European Quality of Life Group 5-Dimension 3-Level, HRQoL: health related quality of life, LOS: Length of stay, VASD: 100mm visual analogue scale for dyspnea

Appendix 4: Subgroup analysis (regularly scheduled IPC drainage vs. as needed IPC drainage)

Outcome	Scheduled drainage	Symptom based drainage only	Excluded (i.e. no described protocol)	Mean difference units	Mean difference [95%CI] regularly scheduled drainage	Mean difference [95%CI] as needed drainage	Test for interaction
Immediate Dyspnea	Putnam 1999	Thomas 2017	Boshuizen 2017	VASD	0.00 [-7.53, 7.53]	5.20 [-14.04, 24.44]	Chi² = 0.24, df = 1 (P = 0.62) I² = 0%
Short-term Dyspnea	Davies 2012 Putnam 1999	Thomas 2017	Boshuizen 2017	VASD	9.47 [1.56, 17.37] I²=0%	2.50 [-16.02, 21.02]	Chi² = 0.46, df = 1 (P = 0.50) I² = 0%
Medium-term Dyspnea	Davies 2012 Putnam 1999	Thomas 2017	Boshuizen 2017	VASD	8.94 [0.80, 17.08] I²=0%	-11.88 [-30.10, 6.35]	Chi² = 4.18, df = 1 (P = 0.04) I² = 76.1%
Long-term Dyspnea	Davies 2012	Thomas 2017	Boshuizen 2017	VASD	14.00 [2.80, 25.20]	-10.40 [-30.51, 9.71]	Chi² = 4.32, df = 1 (P = 0.04) I² = 76.8%
Immediate overall HRQoL	*Olfert 2017	Thomas 2017	Boshuizen 2017	EQ5D3L	0.03 [-0.06, 0.12]	0.03 [-0.10, 0.16]	Chi² = 0.00, df = 1 (P = 0.97) I² = 0%
Short-term overall HRQoL	*Olfert 2017 Putnam 1999	Thomas 2017	-	EQ5D3L	-0.01 [-0.08, 0.06] I²=0%	-0.01 [-0.14, 0.12]	Chi² = 0.00, df = 1 (P = 0.97) I² = 0%
Long-term overall HRQoL	*Olfert 2017	Thomas 2017	-	EQ5D3L	-0.11 [-0.28, 0.05]	-0.02 [-0.17, 0.13]	Chi² = 0.70, df = 1 (P = 0.40) I² = 0%
LOS – initial admission	Davies 2012	Thomas 2017	-	Days	-3.67 [-4.56, -2.78]	-2.00 [-2.25, -1.75]	Chi² = 12.51, df = 1 (P = 0.0004) I² = 92.0%
LOS – effusion related days	Davies 2012	Thomas 2017	-	Days	-3.50 [-4.75, -2.25]	-1.60 [-2.82, -0.38]	Chi² = 4.54, df = 1 (P = 0.03), I² = 78.0%
					Relative risk [95%CI] regular drainage	Relative risk [95%CI] as needed drainage	
Need for repeat pleural intervention	Davies 2012 Putnam 1999	Thomas 2017	Boshuizen 2017	-	0.25 [0.09, 0.65] I²=0%	0.18 [0.06, 0.60]	Chi² = 0.14, df = 1 (P = 0.70) I² = 0%
Overall complications	Davies 2012 Putnam 1999	Thomas 2017	-	-	3.74 [2.31, 6.08] I²=0%	1.27 [0.82, 1.96]	Chi² = 10.62, df = 1 (P = 0.001) I² = 90.6%
Serious complications	Davies 2012 Demmy 2012	Thomas 2017	Boshuizen 2017	-	2.23 [1.02, 4.88] I²=0%	0.32 [0.03, 3.04]	Chi² = 2.53, df = 1 (P = 0.11) I² = 60.5%
Infectious complications	Davies 2012 Putnam 1999	Thomas 2017	Boshuizen 2017	-	1.72 [0.94, 3.12] I²=0%	0.32 [0.03, 3.04]	Chi² = 1.98, df = 1 (P = 0.16) I² = 49.6%

*Olfert 2017 is a study presenting unpublished overall HRQoL data from the Davies et al. 2012 RCT.

df: Degrees of freedom, EQ5D3L: European Quality of Life Group 5-Dimension 3-Level, HRQoL: health related quality of life, LOS: Length of stay, VASD: 100mm visual analogue scale for dyspnea

Appendix 5: Sensitivity analysis (higher/lower risk of bias studies)

Outcome	Lower ROB studies **	Higher ROB studies	Mean difference units	Mean difference [95%CI]	Mean difference excluding higher ROB studies [95%CI]	Test for interaction
Immediate dyspnea	Thomas 2017	Boshuizen 2017 Putnam 1999	VASD	1.84 [-3.82 to 7.51] I^2=0%	5.20 [-14.04, 24.44]	Chi² = 0.13, df = 1 (P = 0.72) I^2 = 0%
Short-term dyspnea	Davies 2012 Putnam 1999 Thomas 2017	Boshuizen 2017	VASD	6.99 [0.28, 13.69] I^2=0%	8.39 [1.12, 15.66] I^2=0%]	Chi² = 0.96, df = 1 (P = 0.33) I^2 = 0%
Medium-term dyspnea	Davies 2012	Boshuizen 2017 Putnam 1999 Thomas 2017	VASD	4.78 [-3.95, 13.51] I^2=28%]	8.90 [-1.70, 19.50]	Chi² = 0.68, df = 1 (P = 0.41) I^2 = 0%
Long-term dyspnea	Davies 2012	Boshuizen 2017 Thomas 2017	VASD	2.71 [-14.18, 19.60] I^2=58%]	14.00 [2.80, 25.20]	Chi² = 4.54, df = 1 (P = 0.03) I^2 = 78.0%
Short-term overall HRQoL	*Olfert 2017 Putnam 1999	Thomas 2017	EQ5D3L	-0.01 [-0.07, 0.05] I^2=0%	-0.01 [-0.08, 0.06] I^2=0%	Chi² = 0.00, df = 1 (P = 0.97) I^2 = 0%
Medium-term overall HRQoL	*Olfert 2017	Putnam 1999	EQ5D3L	-0.04 [-0.11, 0.04] I^2=0%	-0.05 [-0.15, 0.06]	Chi² = 0.11, df = 1 (P = 0.74) I^2 = 0%
Effusion related days in hospital	Thomas 2017	Davies 2012	Days	-2.55 [-4.41, -0.68] I^2=78%	-1.60 [-2.82, -0.38]	Chi² = 4.54, df = 1 (P = 0.03) I^2 = 78.0%

*Olfert 2017 is a study presenting unpublished overall HRQoL data from the Davies et al. 2012 RCT.

** Risk of bias assessed at the outcome level

df: Degrees of freedom, EQ5D3L: European Quality of Life Group 5-Dimension 3-Level, HRQoL: health related quality of life, ROB: risk of bias, VASD: 100mm visual analogue scale for dyspnea