European Respiratory Society guidelines on transbronchial lung cryobiopsy in the diagnosis of interstitial lung diseases

Daniël A. Korevaar[1], Sara Colella[2], Markus Fally[3], Juliette Camuset[4], Thomas V. Colby[5], Lars Hagmeyer[6,7], Juergen Hetzel[8,9], Fabien Maldonado[10], Antonio Morais[11,12,13], Claudia Ravaglia[14,15], René Spijker[16,17], Sara Tomassetti[15,18], Lauren K. Troy[19,20], Johny A. Verschakelen[21], Athol U. Wells[22,23], Thomy Tonia[24], Jouke T. Annema[1,25] and Venerino Poletti[14,15,25]

[1]Dept of Respiratory Medicine, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands. [2]Pulmonary Unit, "G. Mazzini" Hospital, Teramo, Italy. [3]Dept of Respiratory Medicine, Copenhagen University Hospital – Bispebjerg and Frederiksberg, Copenhagen, Denmark. [4]Dept of Thoracic and Vascular Surgery, Tenon Hospital, AP-HP, Paris, France. [5]Dept of Laboratory Medicine and Pathology (Emeritus), Mayo Clinic, Scottsdale, AZ, USA. [6]Bethanien Hospital Solingen, Clinic for Pneumology and Allergology, Center for Sleep Medicine and Respiratory Care, Solingen, Germany. [7]Institute of Pneumology, University of Cologne, Solingen, Germany. [8]Dept of Medical Oncology and Pulmonary Medicine, University Hospital Tuebingen, Tuebingen, Germany. [9]Dept of Internal Medicine, Division of Pneumology, Kantonssiptal Winterthur, Winterthur, Switzerland. [10]Vanderbilt University Medical Center, Nashville, TN, USA. [11]Dept of Pulmonology, Centro Hospitalar e Universitário São João, Porto, Portugal. [12]Faculty of Medicine, University of Porto, Porto, Portugal. [13]Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal. [14]Dept of Thoracic Diseases, G.B. Morgagni Hospital and University of Bologna, Forlì, Italy. [15]ERN-LUNG member. [16]Medical Library, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands. [17]Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands. [18]Dept of Experimental and Clinical Medicine, Careggi University Hospital and University of Florence, Florence, Italy. [19]Royal Prince Alfred Hospital, Camperdown, Australia. [20]Sydney Medical School, University of Sydney, Sydney, Australia. [21]Dept of Radiology, Pulmonary Radiology, University Hospitals Leuven, Leuven, Belgium. [22]Interstitial Lung Disease Unit, Pulmonary Medicine, Royal Brompton Hospital, London, UK. [23]Margaret Turner Warwick Centre for Fibrosing Lung Disease, National Heart and Lung Institute, Imperial College London, London, UK. [24]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland. [25]Shared senior authorship.

Corresponding author: Venerino Poletti (venerino.poletti@gmail.com)

Shareable abstract (@ERSpublications)
Transbronchial lung cryobiopsy provides important diagnostic information in patients with undiagnosed interstitial lung diseases https://bit.ly/3MAR4wc

Cite this article as: Korevaar DA, Colella S, Fally M, *et al*. European Respiratory Society guidelines on transbronchial lung cryobiopsy in the diagnosis of interstitial lung diseases. *Eur Respir J* 2022; 60: 2200425 [DOI: 10.1183/13993003.00425-2022].

This article has an editorial commentary: https://doi.org/10.1183/13993003.01648-2022

Received: 25 Feb 2022
Accepted: 7 May 2022

Abstract

Background In patients with interstitial lung diseases (ILD), histopathological input is often required to obtain a diagnosis. Surgical lung biopsy (SLB) is considered the reference standard, but many patients are clinically unfit to undergo this invasive procedure, and adverse events, length of hospitalisation and costs are considerable. This European Respiratory Society (ERS) guideline provides evidence-based clinical practice recommendations for the role of transbronchial lung cryobiopsy (TBLC) in obtaining tissue-based diagnosis in patients with undiagnosed ILD.

Methods The ERS Task Force consisted of clinical experts in the field of ILD and/or TBLC and methodological experts. Four PICO (Patient, Intervention, Comparator, Outcomes) questions and two narrative questions were formulated. Systematic literature searches were performed in MEDLINE and Embase (up to June 2021). GRADE (Grading, Recommendation, Assessment, Development and Evaluation) methodology was applied.

Results In patients with undiagnosed ILD and an indication to obtain histopathological data: 1) TBLC is suggested as a replacement test in patients considered eligible to undergo SLB, 2) TBLC is suggested in patients not considered eligible to undergo SLB, 3) SLB is suggested as an add-on test in patients with a non-informative TBLC, 4) no recommendation is made for or against a second TBLC in patients with a non-informative TBLC and 5) TBLC operators should undergo training, but no recommendation is made for the type of training required.

Conclusions TBLC provides important diagnostic information in patients with undiagnosed ILD. Diagnostic yield is lower compared to SLB, at reduced serious adverse events and length of hospitalisation. Certainty of the evidence is mostly "very low".

Introduction

Accurate diagnosis of interstitial lung diseases (ILD) is important for guiding treatment decisions and prognosis. In the majority of patients with ILD, integration of clinical, laboratory and radiological data within a multidisciplinary discussion (MDD) results in a diagnosis [1, 2]. For a subset of patients, however, a diagnosis cannot be made with sufficient confidence based on these data and histopathological evaluation of lung tissue may be indicated [3].

Multiple tests can be used to obtain cyto- or histopathological information in the diagnostic work-up of ILD. Bronchoalveolar lavage (BAL) is associated with a very low rate of adverse events, but its diagnostic value is mostly limited to disorders that are typically intra-alveolar (*e.g.* infection, alveolar proteinosis, eosinophilic pneumonia, organising pneumonia, alveolar haemorrhage and diffuse alveolar damage) [4]. Transbronchial lung biopsy (TBLB) with regular forceps is mainly indicated in disorders that involve the centrilobular zones and are characterised by "easy-to-identify" morphological alterations (*e.g.* carcinomatous lymphangitis, sarcoidosis, organising pneumonia and diffuse alveolar damage) [5]. Complications are rare, but diagnostic yield is limited by small specimens, sampling errors and crush artefacts. In particular, TBLB is poorly sensitive for the diagnosis of complex histopathological patterns such as usual interstitial pneumonia (UIP) [6].

Surgical lung biopsy (SLB) is generally obtained thoracoscopically and is currently considered the reference standard when less invasive approaches fail or are not feasible. Samples are large and contain peripheral structures of the secondary pulmonary lobule, with a diagnostic yield of ∼90% [6, 7]. However, SLB is associated with significant morbidity and mortality. In-hospital mortality in elective procedures is estimated to be ∼2% and significantly higher in non-elective procedures [8]. Many patients are not clinically fit to undergo this invasive procedure. Risk is particularly increased in those who may have UIP, are at older age, have significant lung function impairment or are experiencing an acute exacerbation of ILD. In addition, length of hospital admission and associated costs can be considerable [8, 9].

In recent years, transbronchial lung cryobiopsy (TBLC) has been explored as a less invasive alternative to SLB [10]. With this approach, larger samples without crush artefacts can be obtained compared to standard TBLB. Although consensus statements and guidelines dealing with the standardisation of TBLC are available [11–14], to date there have been no guidelines for its clinical application. The European Respiratory Society (ERS) established a Task Force to develop guidelines aimed at providing evidence-based clinical practice recommendations on the role of TBLC in patients with undiagnosed ILD.

Materials and methods

Scope and purpose

The purpose of this project was to evaluate the role of TBLC in obtaining tissue-based diagnosis in patients with undiagnosed ILD, aiming to provide evidence-based clinical practice recommendations for its application. Advantages and disadvantages of TBLC, with respect to diagnostic confidence, diagnostic yield, diagnostic accuracy, adverse events and patient-important outcomes, were assessed and compared with those of SLB. This was done across various subgroups, including patients eligible to undergo SLB, patients not considered eligible to undergo SLB (*e.g.* due to lung function impairment, rapidly progressive disease or comorbid disease), patients at high risk of undergoing TBLC, patients with an initial non-informative TBLC and patients with specific high-resolution computed tomography (HRCT) findings.

Task Force composition and conflict of interest declaration

The Task Force consisted of 16 members: 11 clinical experts in the field of TBLC and/or ILD (J.C., L.H., J.H., F.M., A.M., C.R., S.T., L.K.T., A.U.W., J.T.A. and V.P.), one pathologist (T.V.C.), one thoracic radiologist (J.A.V.) and three junior pulmonologists (in training) with experience in literature syntheses (D.A.K., S.C. and M.F.). An ERS methodologist had the overview of all the methodological steps (T.T.). Task Force members disclosed all potential financial conflicts of interest, which are reported at the end of this guideline.

Formulation of questions

A list of potential guideline questions (both PICO (Patients, Intervention, Comparator, Outcomes) and narrative questions) was developed by two Task Force members (S.C. and V.P.). These were then discussed in detail, prioritised and refined in a live Task Force meeting (November 2019, Florence, Italy),

in a subsequent Task Force telephone meeting (November 2019) and through Task Force e-mail discussions. Guideline questions were finalised in a Task Force telephone meeting in January 2020. Six questions were selected for the guideline. Of these, four were PICO questions that formed the basis of this guideline. In addition, two were narrative questions to be addressed in a descriptive manner, with no intention of making clinical practice recommendations. An overview of guideline questions is provided in table 1, with detailed questions in supplementary appendix S1.

Literature searches and study selection

A single search strategy was developed by a medical information specialist (R.S., with the help of D.A.K.), which covered all guideline questions. MEDLINE and Embase were searched from inception in May 2020 and searches were updated in June 2021. Search terms focused on a combination of the tests of interest (TBLC or SLB) and the condition of interest (ILD). The full search strategy is provided in supplementary appendix S2.

Study selection was done in a two-step approach. First, two Task Force members (D.A.K. and S.C.) independently assessed titles and abstracts of all search results, and those that were considered potentially relevant for at least one of the guideline questions by at least one of them were selected. After this, two Task Force members (D.A.K., S.C., M.F., J.C., C.R. and S.T.) per guideline question independently assessed all full-texts of selected studies, to determine final inclusion. Disagreements were resolved through discussion.

Detailed selection criteria per guideline question are reported in supplementary appendix S3. The study selection process was summarised in PRISMA-DTA (Preferred Reporting Items for Systematic Review and Meta-Analysis of Diagnostic Test Accuracy) flowcharts [15, 16]. Not all studies fulfilling the inclusion criteria were (directly) considered in the evidence syntheses. Instead, for each outcome, we primarily focused on included studies that directly compared TBLC and SLB in patients with undiagnosed ILD, either by performing both tests in each patient (paired direct comparison) or by randomising patients to undergo either procedure (unpaired direct comparison). If direct comparisons were not available for a specific outcome, we focused on studies that indirectly compared TBLC and SLB (*i.e.* a group of patients undergoing TBLC was compared with a group of patients undergoing SLB, without randomisation). Finally, in the absence of direct or indirect comparisons for a specific outcome, we focused on

TABLE 1 Guideline questions and recommendations

PICO Question 1	Question: In patients with undiagnosed ILD considered eligible to undergo SLB, is TBLC a valid replacement test? Recommendation: *For patients with undiagnosed ILD considered eligible to undergo SLB, the Task Force suggests performing TBLC if obtaining histopathological data is indicated* (conditional recommendation for the intervention, "very low" certainty of evidence). Remark: This recommendation applies to centres experienced in performing TBLC.
PICO Question 2	Question: In patients with undiagnosed ILD not considered eligible to undergo SLB, does TBLC increase the diagnostic confidence of the multidisciplinary team discussion? Recommendation: *For patients with undiagnosed ILD not considered eligible to undergo SLB, the Task Force suggests TBLC if obtaining histopathological data is indicated* (conditional recommendation, "very low" certainty of evidence). Remark: This recommendation applies to centres experienced in performing TBLC; the advantages of potentially increasing diagnostic certainty by performing TBLC against the disadvantages of potential serious adverse events should be weighed in each individual patient.
PICO Question 3	Question: In patients with undiagnosed ILD and a non-informative TBLC, is step-up SLB or a second TBLC a valid add-on test? Recommendation: *For patients with undiagnosed ILD and a non-informative TBLC, the Task Force suggests performing step-up SLB if obtaining histopathological data is indicated* (conditional recommendation, "very low" certainty of evidence). *For patients with undiagnosed ILD and a non-informative TBLC, the Task Force makes no recommendation about performing a second TBLC if obtaining histopathological data is indicated, as there is no evidence.*
PICO Question 4	Question: Is formal training in TBLC recommended to optimise yield and minimise adverse events in patients with undiagnosed ILD? Recommendation: *The Task Force suggests that TBLC operators should undergo training* (conditional recommendation, "very low" certainty of evidence)*, but a recommendation on the optimal type of training cannot be made due to lack of evidence.*
Narrative Question 1	Question: Are there specific HRCT findings which would lead to TBLC as the first choice for biopsy?
Narrative Question 2	Question: What are the procedural risks of TBLC in patients with undiagnosed ILD?

Detailed questions are provided in supplementary appendix S1. ILD: interstitial lung diseases; SLB: surgical lung biopsy; TBLC: transbronchial lung biopsy; HRCT: high-resolution computed tomography.

non-comparative studies that only evaluated TBLC or only evaluated SLB in patients with undiagnosed ILD. If available for a specific outcome, we selected a previously published systematic review summarising non-comparative studies, rather than focusing on individual studies, to avoid duplication of review efforts. In such cases, we used the most recently published systematic review that was not part of a clinical guideline or position statement and in which an adequate study quality assessment had been performed.

Assessment of evidence quality and recommendation strength

In line with the GRADE (Grading, Recommendation, Assessment, Development and Evaluation) approach, Task Force members participated in an online survey to rate the importance of each outcome per PICO question on a scale from 1 to 9 for its perceived importance for clinical decision making. Mean scores of 7–9 were considered a "critical" outcome, of 4–6 an "important but not critical" outcome and of 1–3 a "not important" outcome. Survey results are provided in supplementary appendix S4. For each PICO question, one Task Force member (D.A.K., S.C. or M.F.) developed an evidence profile, with input from the ERS methodologist (T.T.) and the two chairs (V.P. and J.T.A.). Data were summarised in evidence tables. The QUADAS-2 tool was used to assess risk of bias and applicability concerns of studies [17]. The evidence was graded according to GRADE [18]. Certainty of the evidence of each outcome was initially rated as "high" if it originated from randomised trials or from well-developed diagnostic accuracy studies and as "low" if it originated from observational data [19]. Certainty was subsequently downgraded if there was high risk of bias, serious inconsistency in results across studies, indirectness of the evidence, imprecision in effect sizes or point estimates, or evidence of publication bias. Data extraction, study quality assessment and performing GRADE was done by one of the three Task Force members (D.A.K., S.C. or M.F.) and checked by another, with disagreements being resolved through discussion.

One Task Force member (D.K. or M.F.) then drafted GRADE evidence-to-decision tables for each PICO question, taking into account the quality of evidence, balance of desirable and undesirable effects, patient values and preferences, resources required, health equity (*i.e.* potential differences in effectiveness in disadvantaged subgroups), acceptability of the tests by key stakeholders, and feasibility of implementation of the tests [20]. This resulted in a recommendation that could either be "strong" (phrased as "the Task Force recommends") or "conditional" (phrased as "the Task Force suggests"). The evidence-to-decision process was discussed in detail in a Task Force video meeting in December 2021, in which 15 of the 16 members plus the ERS methodologist participated, where recommendations were finalised and agreed upon. Members who did not participate in this meeting confirmed their agreement *via* e-mail. A draft manuscript was prepared by one Task Force member (D.K., with input from S.C., M.F., T.T., J.T.A. and V. P.), and shared with the complete Task Force for input and approval.

Patient input

Three patient representatives (one with experience in TBLC, one in SLB, and one in TBLC and SLB) from the European Lung Foundation's Pulmonary Fibrosis Patient Advisory Group provided input on the evidence-to-decision tables, recommendations and manuscript.

Results

An overview of recommendations per PICO question is provided in table 1, with a proposed diagnostic algorithm in figure 1.

Search results

Overall, 4325 records were retrieved in our literature searches (n=3969 in the initial search and n=356 in the update), of which 250 remained after screening of titles and abstracts. Of these, 119 were included: all of them fulfilled the inclusion criteria for PICO Question 1, whereas a subset also fulfilled the inclusion criteria for any of the other guideline questions (PICO Question 2: n=2; PICO Question 3: n=26; PICO Question 4: n=3; Narrative Question 1: n=0; Narrative Question 2: n=10). Flowcharts and lists of included studies per guideline question are provided in supplementary appendix S5.

PICO Question 1: In patients with undiagnosed ILD considered eligible to undergo SLB, is TBLC a valid replacement test?

Recommendation

For patients with undiagnosed ILD considered eligible to undergo SLB, the Task Force suggests performing TBLC if obtaining histopathological data is indicated (conditional recommendation for the intervention, "very low" certainty of evidence).

Remark: This recommendation applies to centres experienced in performing TBLC.

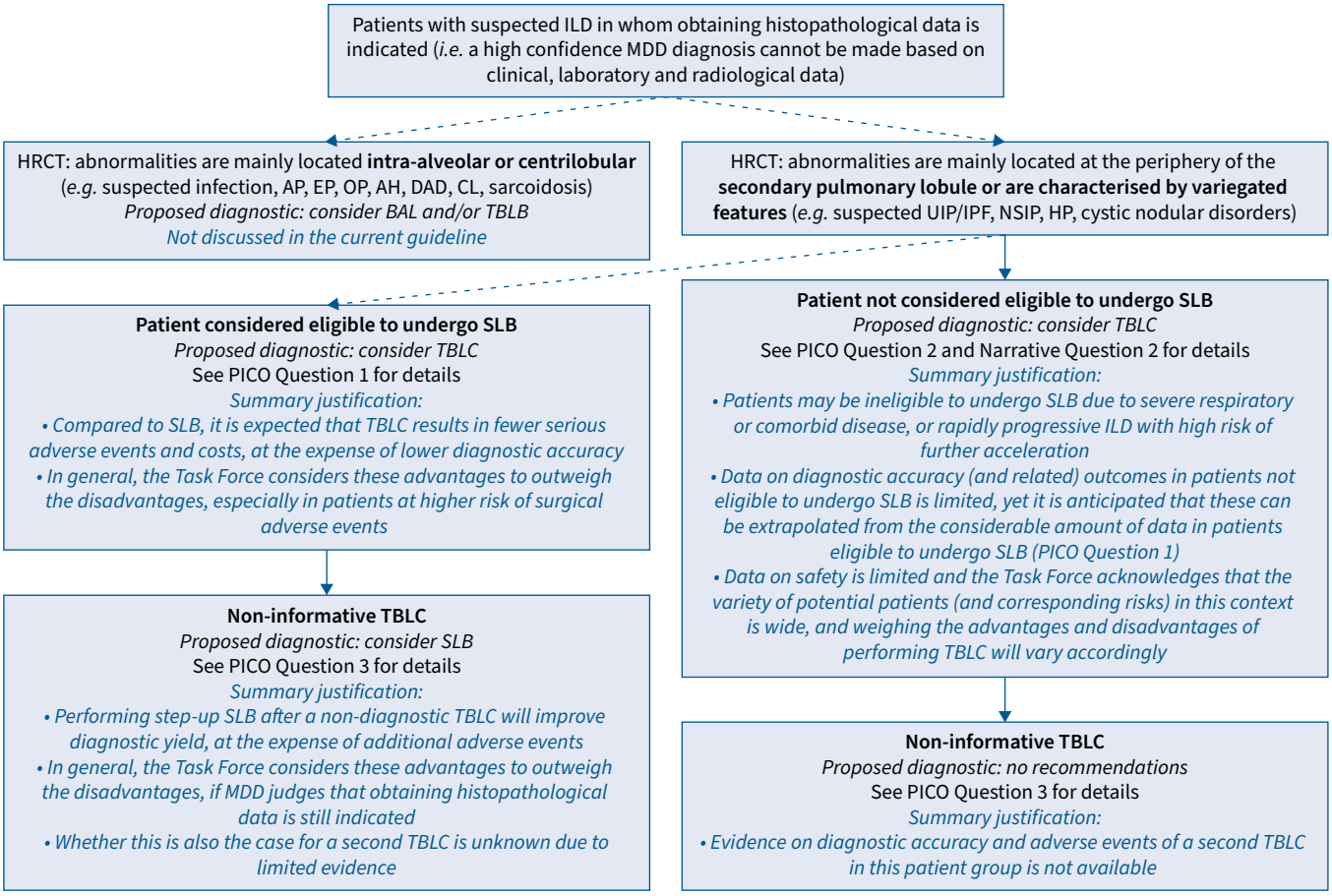

FIGURE 1 Proposed diagnostic algorithm in patients with undiagnosed interstitial lung diseases (ILD). AH: alveolar haemorrhage; AP: alveolar proteinosis; BAL: bronchoalveolar lavage; CL: carcinomatous lymphangitis; DAD: diffuse alveolar damage; EP: eosinophilic pneumonia; HP: hypersensitivity pneumonitis; HRCT: high-resolution computed tomography; IPF: idiopathic pulmonary fibrosis; MDD: multidisciplinary discussion; NSIP: non-specific interstitial pneumonia; OP: organising pneumonia; SLB: surgical lung biopsy; TBLB: transbronchial lung biopsy; TBLC: transbronchial lung cryobiopsy; UIP: usual interstitial pneumonia.

Background

In the majority of patients with ILD, amalgamation of clinical, laboratory and radiological data will result in a diagnosis at MDD. However, in a considerable proportion, lung biopsy is recommended by MDD to establish a confident diagnosis [21]. Historically, SLB has been considered the reference standard for lung tissue acquisition in these patients, but costs, adverse events and length of hospital admission can be considerable [7]. TBLC could serve as a replacement test in these patients, especially if sufficiently accurate (*i.e.* high diagnostic agreement between TBLC and SLB) and resulting in fewer serious adverse events [22].

Evidence summary

An overview of studies included in PICO Question 1, ordered by type of study, is provided in supplementary appendix S6. Evidence summary tables, results from study quality assessment, GRADE tables and evidence-to-decision tables for PICO Question 1 are provided in supplementary appendix S7. Overall, 119 studies were included for this PICO question, but the majority of these were not (directly) considered in the evidence syntheses.

Regarding direct comparisons between TBLC and SLB, no randomised trials were found (although the Task Force is aware of one in progress: Netherlands Trial Registry/International Clinical Trials Registry Platform number NL7634), but two studies were identified that performed both tests in a group of patients with undiagnosed ILD (paired direct comparison; tables S1–S3 in supplementary appendix S7) [23, 24]. Romagnoli *et al*. [23] performed both tests in 21 patients with a non-definite UIP pattern on HRCT, with

blinded pathologists and one final MDD per patient that was informed by results from both tests. Troy *et al*. [24] performed both tests in 65 patients with undiagnosed ILD, with blinded pathologists and two separate MDDs per patient: one informed by TBLC results and one informed by SLB results. Both studies were considered at high risk of bias.

In addition, three studies were found that indirectly compared TBLC and SLB, by comparing a group of patients that underwent TBLC with a group of patients that underwent SLB (table S4 in supplementary appendix S7) [25–27]. Risk of selection bias was considered high in these studies, because no randomisation was performed. Ravaglia *et al*. [25] included 297 patients undergoing TBLC and 150 patients undergoing SLB, all with ILD in whom a diagnosis could not be achieved non-invasively. Tomassetti *et al*. [26] included patients with fibrotic ILD without a typical UIP pattern on HRCT, of whom 58 had TBLC and 59 had SLB. A second study by Tomassetti *et al*. [27] included patients with suspected ILD without a definite UIP pattern on HRCT, of whom 266 had TBLC and 160 had SLB.

Finally, a large number of non-comparative studies were found that evaluated TBLC only or SLB only (n=54 and n=50, respectively) or systematic reviews thereof (n=11). The Task Force focused on two recent systematic reviews (table S5 in supplementary appendix S7) [7, 22]. Sethi *et al*. [22] included 31 studies (18 full-texts and 13 abstracts) on patients with suspected ILD undergoing TBLC, of which 27 could be included in the meta-analysis (n=1443 patients). Risk of bias according to QUADAS-2 was considered high or unclear in 80.6% (n=25) of studies. Sharp *et al*. [7] included 24 studies (n=2665 patients) on patients with ILD undergoing video-assisted thoracoscopic biopsy. Risk of selection bias according to the Cochrane Collaboration Risk of Bias tool was high in all studies.

Overall, outcomes that could be taken into account in PICO Question 1 were diagnostic agreement, diagnostic confidence, diagnostic yield, diagnostic accuracy, survival after idiopathic pulmonary fibrosis (IPF) diagnosis and adverse events. Limited evidence was available on costs and this was only discussed narratively. No comparative evidence of TBLC *versus* SLB was identified on (long-term) patient-important outcomes (*i.e.* quality of life, lung function, mortality, exercise tolerance or survival).

Diagnostic agreement

This is moderate, based on two direct comparisons (tables S2 and S7 in supplementary appendix S7) [23, 24]. Romagnoli *et al*. [23] reported a diagnostic agreement between the TBLC result and final MDD (which was informed by both TBLC and SLB results) of 47.6% (95% CI 26–70%) and a fair κ agreement of 0.31 (95% CI 0.06–0.56). Troy *et al*. [24] reported a diagnostic agreement between MDD informed by TBLC results and MDD informed by SLB results of 76.9%, and a substantial κ agreement of 0.62 (95% CI 0.47–0.78). The evidence was judged as "very low" (downgraded for risk of bias, indirectness and imprecision).

High or definite confidence final diagnosis

This can be obtained in the majority of patients, with both TBLC and SLB, based on one direct comparison (tables S2 and S7 in supplementary appendix S7) [24]. Troy *et al*. [24] reported that a high or definite confidence diagnosis could be obtained in MDD informed by TBLC results in 60.0% (n=39) and in MDD informed by SLB results in 73.8% (n=48; p=0.090). In 94.8% (n=37) of patients with a high or definite confidence diagnosis in MDD informed by TBLC results, the same diagnosis was reached in MDD informed by SLB results. In 23.1% (n=6) of patients with a low confidence or unclassifiable diagnosis in MDD informed by TBLC results, a high or definite confidence diagnosis was reached in MDD informed by SLB results. The evidence was judged as "very low" (downgraded for risk of bias, indirectness and imprecision).

Increase in diagnostic confidence

This is significant for TBLC, based on an indirect comparison and a non-comparative study (tables S6 and S7 in supplementary appendix S7) [26, 28]. Tomassetti *et al*. [26] reported that the percentage increase in IPF diagnosis made with a high level of confidence in MDD changed from 29% to 63% before and after adding TBLC results (p=0.0003), and from 30% to 65% before and after adding SLB results (p=0.0016). Hetzel *et al*. [28] reported among 128 patients a percentage increase in confidence (*i.e.* confident diagnosis or provisional diagnosis with high confidence) from 60.2% (n=77) after clinico-radiological discussion and BAL to 81.2% (n=104) when adding TBLC results (p<0.0001); this implies that in 51 patients with no consensus diagnosis or with a provisional diagnosis with low confidence after BAL, TBLC led to a definite or confident provisional diagnosis in 62.7% (n=32). The evidence was judged as "very low" (downgraded for risk of bias and imprecision).

Diagnostic yield for a histopathological diagnosis

This is high for TBLC, yet somewhat higher for SLB, based on both comparative and non-comparative studies (tables S2, S4, S5 and S7 in supplementary appendix S7) [7, 22–25]. In the direct comparison by Romagnoli *et al*. [23], a diagnostic pattern was obtained in 81.0% for TBLC and 100% for SLB, with histopathological agreement between the two in 38.1% (95% CI 18–62%) and a κ agreement of 0.22 (95% CI 0.01–0.44). In the direct comparison by Troy *et al*. [24], a diagnostic pattern was obtained in 90.8% and 96.9%, respectively, with a histopathological agreement (for guideline-refined pattern) of 70.8% and a weighted κ agreement of 0.70 (95% CI 0.55–0.86). The indirect comparison by Ravaglia *et al*. [25] reported a diagnostic yield of 82.8% for TBLC and 98.7% for SLB (p=0.013). The evidence of the comparative studies was judged as "very low" (downgraded for risk of bias and imprecision). Similar results were found in the non-comparative studies. In the meta-analyses of studies only reporting on TBLC or only reporting on SLB, summary diagnostic yield was 72.9% (95% CI 67.9–77.7%) and 91.1% (95% CI 86.9–93.2%), respectively [7, 22]. The evidence of the non-comparative studies was judged as "very low" (downgraded for risk of bias and indirectness).

Diagnostic accuracy for diagnosing IPF

This is moderate for (MDD informed by) TBLC, based on two direct comparisons (tables S3 and S7 in supplementary appendix S7) [23, 24]. In Romagnoli *et al*. [23], sensitivity and specificity for diagnosing IPF against a reference standard of final MDD (informed by TBLC and SLB results) were 66.7% (95% CI 31–91%) and 75.0% (95% CI 43–93%), respectively (recalculated based on reported data). In Troy *et al.* [24], sensitivity and specificity for diagnosing IPF against a reference standard of MDD (informed by SLB results) were 91.4% (95% CI 76–98%) and 80.0% (95% CI 61–92%), respectively (recalculated based on reported data). The evidence was judged as "very low" (downgraded for risk of bias, indirectness and imprecision).

Survival after IPF diagnosis

In the indirect comparison by Tomassetti *et al*. [27], MDD diagnoses of IPF (*versus* another ILD) based on TBLC or SLB were both significantly associated with 5-year transplant-free survival (TBLC adjusted HR 2.98 (95% CI 1.19–1.47; p=0.02) and SLB adjusted HR 4.07 (95% CI 2.01–8.24; p<0.0001)) (tables S4 and S7 in supplementary appendix S7). The evidence was judged as "very low" (downgraded for risk of bias and indirectness).

Adverse events: mortality

This is lower in TBLC compared to SLB, based on two indirect comparisons and on non-comparative studies (tables S4, S5 and S7 in supplementary appendix S7) [7, 22, 25, 26]. Ravaglia *et al*. [25] reported that mortality due to an adverse event occurred in 0.3% (n=1) in the TBLC group and in 2.7% (n=4) in the SLB group (p=0.045). In Tomassetti *et al*. [26], mortality was 1.7% (n=1) in the TBLC group and 3.4% (n=2) in the SLB group. The evidence for the indirect comparison was judged as "very low" (downgraded for risk of bias and indirectness). Similar results were found in the non-comparative studies. In the systematic review on studies evaluating TBLC only, summary incidence of mortality within 30 days was 0.3% (95% CI not reported; based on 33 studies) [22]. In the systematic review on studies evaluating SLB only, summary incidence of mortality within 30 days was 2.3% (95% CI 1.3–3.6%; based on 21 studies) [7]. The evidence of the non-comparative studies was judged as "very low" (downgraded for risk of bias).

Adverse events: time of hospitalisation

This is shorter for TBLC compared to SLB, based on two indirect comparisons (tables S4 and S7 in supplementary appendix S7) [25, 26]. Ravaglia *et al.* [25] reported a mean (range) time of hospitalisation of 2.6 (0–17) days for TBLC and 6.1 (3–48) days for SLB (p<0.0001). Tomassetti *et al*. [26] reported a mean (range) time of hospitalisation of 3 (0–9) days for TBLC and 6 (3–17) days for SLB (p-value not reported). The evidence was judged as "very low" (downgraded for risk of bias and indirectness).

Adverse events: other

These are more frequent for TBLC, based on comparative and non-comparative studies (tables S2, S4, S5 and S7 in supplementary appendix S7) [7, 22–26]. However, this is mainly because pneumothorax as a complication is only relevant for TBLC, as it occurs by definition in 100% of cases for SLB. In the direct comparison by Romagnoli *et al*. [23], serious adverse events occurred in 9.5% (n=2: pneumothorax) for TBLC and in 0% for SLB. In the direct comparison by Troy *et al*. [24], no serious adverse events occurred for TBLC (one pneumothorax was not considered as such by the study authors) and in 3.1% for SLB (n=1: re-hospitalisation due to chest pain; n=1: bleeding requiring intervention). In the indirect comparison by Ravaglia *et al*. [25], pneumothorax occurred in 20.2% (n=60) of patients undergoing TBLC; no severe bleeding was reported for either TBLC or SLB. In the indirect comparison by

Tomassetti *et al*. [26], pneumothorax occurred in 32.8% (n=19) of patients undergoing TBLC; no severe bleeding was reported for either TBLC or SLB. The evidence for the comparative studies was judged as "very low" (downgraded for risk of bias, indirectness and imprecision). In the systematic review on TBLC, overall complication rate was 23.1% (95% CI not reported; based on 31 studies), with summary incidence of pneumothorax of 9.4% (95% CI 6.7–12.5%) and summary incidence of moderate–severe bleeding of 14.2% (95% CI 7.9–21.9%) [22]. In the systematic review on SLB, summary incidence of surgical morbidity was 12.9% (95% CI 9.3–16.9%; based on 18 studies) [7]. The evidence for the studies on TBLC or SLB only was judged as "very low" (downgraded for risk of bias).

Costs

These appear to be lower for TBLC compared to SLB, based on two studies reporting on a cost analysis [7, 29]. Hernández-González *et al*. [29] estimated that the systematic use of TBLC in their clinic (involving 33 patients over a 3-year period) had reduced overall costs up to EUR 59 846, compared to systematically performing SLB. Sharp *et al*. [7] (theoretical cost analysis) estimated that the systematic use of TBLC (followed by SLB if inconclusive) reduced costs up to GBP 647 per patient per year. No evidence grading was performed for this outcome due to limited data.

Justification of the recommendation

Overall certainty of the evidence was considered "very low" (tables S7 and S8 in supplementary appendix S7). Taking the aforementioned results into account, the Task Force concludes that TBLC adds important information to MDD, which results in an increase in diagnostic confidence. Diagnostic yield is likely to be somewhat lower than for SLB, although the extent to which this is the case is unclear, with varying results across studies. Overall adverse event rates are difficult to compare between TBLC and SLB, because populations and definitions of complications varied across studies and because pneumothorax is not considered a complication for SLB because all patients require chest tube drainage. Taking these considerations into account, the Task Force put most emphasis on a reduction in serious adverse events (especially mortality) and a shorter period of post-procedural hospitalisation for TBLC. Costs are expected to be reduced in TBLC. Data on patient preferences are unavailable, but the three patient representatives who provided input indicated that they assumed that most patients would opt for TBLC as the initial diagnostic procedure. The Task Force is not aware of major issues in health equity, acceptability or feasibility of implementation of TBLC. The use and availability of TBLC have increased rapidly over the past years, and are likely to further do so in the coming years. Among patients recruited in the European IPF registry (eurIPFreg; an internet-based registry, consisting of IPF patients from a range of European centres), SLB was performed in 32% in 2009 *versus* 8% in 2016, likely due to increased use of TBLC [30]. However, availability varies across countries and not all patients may have easy access to TBLC. A systematic evaluation of ILD diagnostic practice across 457 centres in 64 countries in 2017 showed that around one-third of centres applied TBLC [31].

Overall, the Task Force considers the reduction in serious adverse events to outweigh the reduced diagnostic yield, in centres experienced in performing TBLC. Minimum requirements for safe implementation of TBLC should include elements such as the availability of competent TBLC operators (see PICO Question 4), and the ability to safely apply sedation, promptly manage complications and ensure airway protection. Best practice documents, consensus statements and guidelines on standardisation of the TBLC procedure have been published previously [11–14, 32]. In addition, adequate patient selection in an MDD setting should be ensured. Essential features of an ILD MDD have recently been suggested through an international Delphi survey [33].

Recommendations for monitoring and future research

For quality assurance, healthcare centres that offer TBLC or SLB are advised to keep track of outcomes such as diagnostic yield and complications. Regarding future research, additional direct comparisons between TBLC and SLB are recommended. Ideally, a large randomised trial is performed. In addition to outcomes related to diagnostic accuracy, adverse events and costs, such studies should focus on long-term patient-important outcomes such as disease control and mortality (based on the diagnosis made by either test and the subsequent treatment initiated).

PICO Question 2: In patients with undiagnosed ILD not considered eligible to undergo SLB, does TBLC increase the diagnostic confidence of the multidisciplinary team discussion?

Recommendation

For patients with undiagnosed ILD not considered eligible to undergo SLB, the Task Force suggests TBLC if obtaining histopathological data is indicated (conditional recommendation, "very low" certainty of evidence).

Remark: This recommendation applies to centres experienced in performing TBLC; the advantages of potentially increasing diagnostic certainty by performing TBLC against the disadvantages of potential serious adverse events should be weighed in each individual patient.

Background

Some patients with ILD have severe respiratory or comorbid disease and they may not be able to tolerate SLB. Others may have rapidly progressive ILD and risk of further acceleration may be increased after performing SLB [34]. In these patients, TBLC could provide a less invasive alternative to obtain a histopathological diagnosis.

Evidence summary

GRADE tables and evidence-to-decision tables for PICO Question 2 are provided in supplementary appendix S8. Although it is likely that several of the studies evaluating TBLC included in PICO Question 1 may have also selected patients that were not considered eligible to undergo SLB, this information was rarely explicitly reported. Overall, only two studies were identified that explicitly reported on outcomes in such patients. The only outcomes that could be evaluated were diagnostic yield and adverse events.

Diagnostic yield for a histopathological diagnosis

This is high, based on one non-comparative study (table S1 in supplementary appendix S8) [35]. Matta *et al*. [35] reported on 17 critically ill patients with ILD and acute hypoxaemic respiratory failure, who were considered poor candidates for SLB or refused this. 12 interventions were performed at bedside in the intensive care unit (ICU). Overall, diagnostic yield was 88.2% (95% CI 64–99%) and histopathological data led to management changes in 88.2% (95% CI 64–99%). However, diagnostic yield may be considered inflated by a subset of patients with non-specific patterns inconsistent with their profound respiratory failure and it is unclear whether the reported management changes actually influenced clinical outcomes. The evidence was judged as "very low" (downgraded for risk of bias, indirectness and imprecision).

Adverse events

These vary, based on two non-comparative studies (table S1 in supplementary appendix S8) [35, 36]. This was probably due to considerable differences in disease severity across included patients. In the same study by Matta *et al*. [35], pneumothorax occurred in 35.3% (n=6) and moderate bleeding in 5.9% (n=1), with 30-day ICU mortality of 47.1% (n=8; although, according to the authors, none directly attributable to TBLC). Bondue *et al*. [36] compared adverse events of TBLC in 38 patients with undiagnosed ILD at high risk of SLB (defined as age ⩾75 years, body mass index (BMI) ⩾35 $kg \cdot m^{-2}$, systolic pulmonary arterial pressure (sPAP) by echocardiography ⩾45 mmHg, forced vital capacity (FVC) <50%, diffusing capacity of the lung for carbon monoxide (D_{LCO}) <30% and/or significant cardiac comorbidities with reduced heart ejection fraction) with 58 patients at low risk. Numbers of bleeding, pneumothorax, mortality and hospital stay were equal between both groups (see Narrative Question 2). The evidence was judged as "very low" (downgraded for inconsistency and imprecision).

Justification of the recommendation

Evidence is mostly lacking for answering this PICO question and overall certainty of the evidence is "very low" (tables S1 and S2 in supplementary appendix S8). The Task Force assumes that diagnostic yield is likely to be similar as for patients considered eligible to undergo SLB (PICO Question 1), but there are no data to confirm this. Regarding adverse events, the Task Force acknowledges that the variety of potential patients (and corresponding risk of performing TBLC) in this context is wide, and weighing the advantages and disadvantages of performing TBLC will vary accordingly. Limited evidence suggests safety in high-risk patients described in the study by Bondue *et al*. [36]. However, the risk of accelerating disease in patients who are critically ill or have rapidly progressive ILD, such as in the study by Matta *et al*. [35], may be unacceptably high. If obtaining histopathological data is indicated, TBLC is suggested, but the advantages of potentially increasing diagnostic certainty against the disadvantages of potential adverse events should be carefully weighed in each patient.

Recommendations for monitoring and future research

Healthcare centres that offer TBLC in patients not considered eligible to undergo SLB are advised to collect data on outcomes such as diagnostic yield, complications and patient-important outcomes. Regarding future research, prospective studies evaluating these outcomes of TBLC in high-risk patients not considered eligible to undergo SLB could be initiated in experienced centres, clarifying which patients are at relatively low risk.

PICO Question 3: In patients with undiagnosed ILD and a non-informative TBLC, is step-up SLB or a second TBLC a valid add-on test?

Recommendation

For patients with undiagnosed ILD and a non-informative TBLC, the Task Force suggests performing step-up SLB if obtaining histopathological data is indicated (conditional recommendation, "very low" certainty of evidence). *For patients with undiagnosed ILD and a non-informative TBLC, the Task Force makes no recommendation about performing a second TBLC if obtaining histopathological data is indicated, as there is no evidence.*

Background

As illustrated in PICO Question 1, performing TBLC does not always result in a high confidence diagnosis at MDD and it may be decided that additional efforts to obtain a histopathological diagnosis are warranted. Step-up SLB or a second TBLC could serve as an add-on test in these patients, if diagnostic yield or confidence is sufficiently improved and if the number of adverse events is not unacceptably high.

Evidence summary

Evidence summary tables, results from study quality assessment, GRADE tables and evidence-to-decision tables for PICO Question 3 are provided in supplementary appendix S9. No studies were identified that directly compared outcomes of step-up SLB *versus* a second TBLC (either by randomising patients or by performing both tests in each patient) in patients with undiagnosed ILD and a non-informative initial TBLC. Also, no systematic reviews were found in this specific patient population. Overall, we identified 26 studies that reported on at least one patient with a non-informative initial TBLC and subsequent step-up SLB, and two studies in which at least one patient had a second TBLC. Risk of bias was high in the majority of studies, mainly due to retrospective, non-consecutive inclusion of patients (table S1 in supplementary appendix S9). The outcomes that could be evaluated were diagnostic yield, diagnostic confidence and adverse events.

Diagnostic yield for a histopathological diagnosis

For step-up SLB, this is high, based on a random effects meta-analysis of 26 studies (188 patients) (tables S2 and S3 in supplementary appendix S9), with a summary estimate of diagnostic yield of 92% (95% CI 82–96%) (figure S1a in supplementary appendix S9). However, besides the high risk of bias, it should be noted that only five studies included more than 10 patients; when only including these five studies in the meta-analysis, results are similar, with a summary diagnostic yield of 91% (95% CI 79–97%) (figure S1b in supplementary appendix S9). The evidence for step-up SLB was judged as "very low" (downgraded for risk of bias and indirectness). For a second TBLC, evidence on diagnostic yield is limited: this was 100% (95% CI 39.8–100%) in one study (based on only four patients) [37] and 62.5% (95% CI 24.5–91.5%) in another (based on eight patients) [38]. The evidence for a second TBLC was judged as "very low" (downgraded for risk of bias, indirectness and imprecision).

Diagnostic confidence

For step-up SLB, this seems to increase, based on two studies that explicitly aimed to prospectively evaluate the added value of performing step-up SLB after a non-informative TBLC (tables S2 and S3 in supplementary appendix S9) [39, 40]. Hagmeyer *et al.* [39] evaluated a diagnostic algorithm, proposing TBLC as initial diagnostic with SLB as an optional step-up procedure when findings remained inconclusive. Among 61 patients, a confident diagnosis was reached in MDD after TBLC in 75.4% (n=46). In the remaining 15 cases, step-up SLB was recommended, which was performed in 13, and a conclusive clinical diagnosis could be achieved in 92.3% (n=12) of them (change in histopathological diagnosis: n=3; histopathological diagnosis confirmed with increased confidence leading to increased MDD confidence: n=5; histopathological diagnosis confirmed with same confidence leading to revision of initial MDD working diagnosis: n=4). Bondue *et al.* [40] evaluated a diagnostic algorithm in which patients with ILD initially underwent TBLC, followed by SLB in case of an uncertain histopathological diagnosis or a non-specific interstitial pneumonia (NSIP) pattern after initial TBLC, hypothesising that a coexistent UIP pattern could have been missed. Of 81 patients undergoing TBLC, 16.0% (n=13) had no histopathological diagnosis and 19.8% (n=16) had a pattern suggestive of NSIP. Of these, 14 patients had subsequent SLB, showing a UIP pattern in 78.6% (n=11), hypersensitivity pneumonitis pattern in 14.3% (n=2) and a NSIP pattern in 7.1% (n=1). Of the six patients with a NSIP pattern at TBLC undergoing subsequent SLB, this showed a UIP pattern in five and confirmed a NSIP pattern in only one. The evidence for step-up SLB was judged as "very low" (downgraded for risk of bias and imprecision). For a second TBLC, no evidence on this was identified.

Adverse events

For step-up SLB, only four studies (n=13 patients) reported on this, which occurred in 11.8% (prolonged air leak: n=1; death within 30 days after SLB due to acute exacerbation of lung fibrosis: n=2; overnight stay at the ICU due to prolonged respiratory and cardiovascular instability: n=1) (tables S2 and S3 in supplementary appendix S9). The evidence for step-up SLB was judged as "very low" (downgraded for risk of bias and imprecision). For a second TBLC, no evidence on adverse events was identified.

Justification of the recommendation

Overall certainty of the evidence was considered "very low" (tables S3 and S4 in supplementary appendix S9). Based on our meta-analysis of diagnostic yield, it seems that step-up SLB after an initial non-informative TBLC often results in a histopathological diagnosis. Insufficient evidence was obtained to be able to make similar statements for a second TBLC. Evidence on adverse events in this subgroup of patients is low, but it is likely that overall complication rates of SLB and TBLC in patients with ILD (PICO Question 1) can be extrapolated to patients with a non-informative initial TBLC. No evidence is available on costs. The Task Force is not aware of major issues in health equity, acceptability of either test or feasibility of performing a second procedure. The patient representatives who provided input indicated that it is their opinion that, if initial TBLC is non-informative, most patients would opt for step-up SLB rather than a second TBLC as subsequent diagnostic. In general, the Task Force believes that the potential disadvantages (adverse events and costs) are outweighed by the need to obtain a histopathological diagnosis, if MDD judges that this is indicated. Therefore, the balance is probably in favour of performing an additional test. Yet, this should be decided upon on a case-by-case level, taking into account factors such as (relative) contra-indications (*e.g.* severe lung function or cardiac impairment) to undergo additional testing.

Recommendations for monitoring and future research

Healthcare centres that offer step-up SLB or a second TBLC after a non-informative initial TBLC are advised to collect data on outcomes such as diagnostic yield and complications. Regarding future research, prospective studies should be performed, evaluating the added value (in terms of diagnostic yield, adverse events and costs) of performing step-up SLB or a second TBLC. These can be single-arm studies (*i.e.* step-up SLB or second TBLC only) or two-arm studies (ideally a randomised trial) in which both tests are compared.

PICO Question 4: Is formal training in TBLC recommended to optimise diagnostic yield and minimise adverse events in patients with undiagnosed ILD?

Recommendation

The Task Force suggests that TBLC operators should undergo training (conditional recommendation, "very low" certainty of evidence), *but a recommendation on the optimal type of training cannot be made due to lack of evidence.*

Background

DiBardino *et al.* [41] reported that the introduction of TBLC at a large academic medical centre in the USA was linked to a high rate of complications. It has previously been demonstrated in bronchoscopy and endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) that formal training, often using simulators, can shorten learning curves and assure safe and efficient procedures [42–44]. Similar formal training may increase diagnostic yield and reduce adverse events in TBLC.

Evidence summary

Results from study quality assessment, GRADE tables and evidence-to-decision tables for PICO Question 4 are provided in supplementary appendix S10. Studies directly addressing the PICO question were not identified, but three studies were found that evaluated outcomes in early *versus* late procedures, which was considered as a surrogate for operator experience. Of these, one study reported cumulated sum (CUSUM) scores for some of the prioritised outcomes [45]. However, as these scores were not comprehensible and raw data regarding the outcomes were not published, the Task Force decided to exclude the study from the analyses. The remaining two studies were included [46, 47]. Risk of bias was high or unclear in both of them (table S1 in supplementary appendix S10). The outcomes that could be evaluated were diagnostic yield (including surrogate outcomes sample length and sample area) and adverse events.

Diagnostic yield for a histopathological diagnosis, sample length and sample area

This seems to be positively associated with operator experience, based on one study (table S2 in supplementary appendix S10) [46]. Almeida *et al.* [46] reported a diagnostic yield of 74.0% in the first 50 TBLC procedures performed *versus* 90.0% in the subsequent 50 procedures (p=0.04). Furthermore, sample

area and sample length significantly increased with increasing operator experience. The evidence for these outcomes was judged as "very low" (downgraded for risk of bias and indirectness).

Adverse events

These seem to reduce in late *versus* early procedures, based on two studies (table S2 in supplementary appendix S10) [46, 47]. Almeida *et al*. [46] reported pneumothoraxes in 24.0% (n=12) in the first 50 TBLC procedures performed and 12.0% (n=6) in the subsequent 50 procedures (p=0.12). Kronborg-White *et al*. [47] reported pneumothoraxes in 30.0% (n=6) in the first 20 TBLC procedures performed and in 22.2% (n=4) in the subsequent 18 procedures (p=0.59). Data on bleeding was limited to one study: Almeida *et al*. [46] reported bleeding events in 2.0% (n=1) *versus* 4.0% (n=2) in early *versus* late procedures (p=0.56). The evidence for these outcomes was judged as "very low" (downgraded for risk of bias, indirectness and imprecision).

Justification of the recommendation

Overall certainty of the evidence was considered "very low" (tables S2 and S3 in supplementary appendix S10). The Task Force considers training important to achieve operator competency, as diagnostic yield increases and adverse events decrease with experience. Introducing TBLC in less experienced centres may result in higher rates of complications [41]. For other invasive procedures, it has been shown that formal training programmes can increase operator competency [42–44]. However, formal training in TBLC to shorten the learning curve and improve procedure outcomes in suspected ILD has not yet been evaluated. None of the identified studies assessed learning curves between bronchoscopists that received formal training *versus* those that did not. The included studies could not deliver definite answers about desirable and undesirable effects, the required resources or equity. Comparisons could only be made between earlier and later procedures, but none of them described 1) the bronchoscopists' baseline experience regarding TBLC or other invasive procedures, or 2) the kind of training the bronchoscopists had received before and during the study.

Recommendations for monitoring and future research

The Task Force believes that a certain level of training is needed to perform TBLC in a standardised, safe and effective way. If implemented, the impact of formal TBLC training programmes must be monitored closely. It is strongly recommended that studies evaluating the impact of formal training programmes in TBLC are designed and conducted. First, formal training programmes must be defined and developed. Second, it is recommended that direct comparisons of formal training and apprentice-based training on the prioritised outcomes are performed. This can be done either by performing a randomised trial or by performing observational studies which include bronchoscopists undergoing different types of training.

Narrative Question 1: Are there specific HRCT findings which would lead to TBLC as the first choice for biopsy?

Evidence summary

The Task Force aimed to identify studies that evaluated the performance and safety of TBLC in (subgroups of) patients with specific HRCT findings (*e.g.* areas with increased lung attenuation, areas with decreased lung attenuation, nodular and micronodular patterns, centrilobular distribution, random distribution or reticular pattern). However, no such studies were identified. It is recommended that prospective studies are performed, evaluating diagnostic yield and adverse events of TBLC in patients with specific HRCT findings compared to other methods to obtain histopathological data (*e.g.* TBLB with forceps and SLB).

Narrative Question 2: What are the procedural risks of TBLC in patients with undiagnosed ILD?

Background

Procedural adverse events are frequent in TBLC, although most are minor (see PICO Question 1). The most frequent adverse events are pneumothorax and mild bleeding. Serious adverse events, such as major bleeding, respiratory failure, exacerbation of ILD or mortality, are uncommon in the reported literature. Several previous studies evaluated predictors of adverse events. Aburto *et al*. [48] analysed 257 TBLC procedures, with complications in 15.2%, and 5.4% requiring hospital admission on the day of the procedure. Variables significantly associated with hospital admission were modified Medical Research Council dyspnoea score $\geqslant$2, FVC <50% and Charlson Comorbidity Index $\geqslant$2. To minimise adverse events, it is useful to evaluate which groups of patients are at particularly high procedural risk, so that this risk can be weighed against the added value of increasing diagnostic confidence.

Evidence summary

Evidence summary tables for Narrative Question 2 are provided in supplementary appendix S11. The Task Force aimed to identify which subgroups of patients are at higher procedural risk, specifically focusing on

those with lung function impairment (FVC <50%, D_{LCO} <35%), pulmonary hypertension (sPAP >40 mmHg), advancing age (>65 years), acute exacerbation of ILD (respiratory failure or rapid worsening), major comorbidities or increased bleeding risk. Two types of studies were selected: 1) those evaluating adverse events of TBLC in patients with ILD at high procedural risk only (n=3 studies identified; table S1 in supplementary appendix S11) [35, 37, 49] and 2) those comparing adverse events in patients at high *versus* low procedural risk (n=7 studies identified; table S2 in supplementary appendix S11) [36, 38, 50–54]. Pooling of data was not performed due to heterogeneity in study populations and reported outcomes.

Overall high procedural risk

Bondue *et al*. [36] compared adverse events of TBLC in 38 patients with ILD at high risk of SLB (defined as age ⩾75 years, BMI ⩾35 kg·m^{-2}, sPAP by echocardiography ⩾45 mmHg, FVC <50%, D_{LCO} <30% and/or significant cardiac comorbidities with reduced heart ejection fraction) with 58 patients at low risk with equal numbers of moderate bleeding (28.9% (n=11) *versus* 29.3% (n=17); p=0.969), severe bleeding (2.6% (n=1) *versus* 5.2% (n=3); p=0.542), pneumothorax (13.2% (n=5) *versus* 20.7% (n=12); p=0.419), mortality (2.6% (n=1) *versus* 0% (n=0)) and median hospital stay (1 day *versus* 1 day; p=0.675).

Lung function impairment

Three studies reported on adverse events in patients with lung function impairment, and one study compared adverse events in patients with more and less lung function impairment [35–37, 49]. Matta *et al*. [35] included 17 critically ill patients with acute hypoxaemic respiratory failure, in whom pneumothorax occurred in 35.3% (n=6), moderate haemorrhage in 5.9% (n=1) and 8-day mortality (although not directly related to TBLC) in 47.1% (n=8). Ravaglia *et al*. [37] reported on adverse events in a subgroup of 31 patients with FVC <50% and/or D_{LCO} <35%; pneumothorax occurred in 19.4% (n=6), mild–moderate bleeding in 19.4% (n=6) and empyema in 3.2% (n=1). She *et al*. [49] reported on TBLC in a subgroup of 15 patients with D_{LCO} <40% and identified that no increased rate of complications occurred in these patients, although no further details were provided. Finally, Bondue *et al*. [36] compared a subgroup of 15 patients with severe pulmonary impairment (FVC <50% or D_{LCO} <30%) *versus* 58 low-risk patients, reporting 6.7% (n=1) *versus* 20.7% (n=12) pneumothoraxes (p=0.316), respectively, and no differences in bleeding.

Hospitalised patients

Three studies reported on adverse events in patients that were already hospitalised *versus* non-hospitalised patients, although reasons for hospitalisation (and if these were related to ILD) were unclear in two of these [50, 53, 54]. Cooley *et al*. [50] compared adverse events in 17 hospitalised patients (n=15 due to respiratory failure, n=1 due to fatigue and n=1 due to kidney injury) *versus* 142 outpatients. Pneumothorax occurred in 23.5% (n=4) *versus* 9.9% (n=14; p=0.11), respectively, with persistent air leak in 5.9% (n=1) and 0.7% (n=1; p=0.20). ICU transfer within 48 h after the procedure occurred in 11.8% (n=2) *versus* 2.1% (n=3; p=0.09), and 30-day mortality in 5.9% (n=1) and 1.4% (n=2; p=0.29), respectively. Kropski *et al*. [53] compared adverse events in four hospitalised patients (reason for hospitalisation not reported) *versus* 33 outpatients, but no pneumothoraxes or bleeding occurred. One (25.0%) of the hospitalised patients required ICU admission due to post-procedural hypoxaemia and one (3.0%) of the outpatients required hospitalisation due to haemoptysis. Finally, Pannu *et al*. [54] compared 30-day mortality in eight hospitalised patients (reason for hospitalisation not reported) *versus* 189 outpatients, identifying that this was 25.0% (n=2) and 1.1% (n=2), respectively.

Age

Hetzel *et al*. [52] compared moderate–severe bleeding rates in 189 patients aged ⩾65 years *versus* 160 patients aged <65 years, identifying that this occurred in 20.1% (n=38) and 10.6% (n=17; p=0.018), respectively.

Body mass index

Bondue *et al*. [36] compared a subgroup of 15 patients with BMI ⩾35 kg·m^{-2} *versus* 58 low-risk patients, reporting pneumothoraxes in 6.7% (n=1) *versus* 20.7% (n=12; p=0.206), respectively, and no differences in bleeding.

Anticoagulants

Hetzel *et al*. [52] compared moderate–severe bleeding rates in 51 patients with aspirin use *versus* 303 patients without aspirin use, identifying that this occurred in 25.5% (n=13) and 14.9% (n=45; p=0.067), respectively. Kronborg-White *et al*. [38] did the same in 86 patients with any anticoagulant (n=64 acetyl salicylic acid; n=13 platelet inhibitors; n=15 direct oral anticoagulants; n=18 vitamin K antagonists; all patients ceased individual anticoagulants before the procedure according to national guidelines) *versus* 164

patients without anticoagulants. Moderate–severe bleeding occurred in 22.1% (n=19) *versus* 22.0% (n=36; p=0.98).

Summary and recommendations for future research

Evidence regarding adverse events from TBLC in patients at high procedural risk is limited and most of the aforementioned studies only included a small number of patients, resulting in limited power and wide confidence intervals. Data from high-volume centres suggest that TBLC may be performed relatively safely in patients with advancing age, elevated BMI, cardiac impairment or (non-acute) pulmonary impairment (even at FVC <50% or D_{LCO} <30%). The risk of serious adverse events seems to be particularly high in hospitalised patients with acute hypoxaemic respiratory failure or rapidly progressing ILD. Despite some reassurances from the literature, a conservative approach for patient selection is recommended for centres with less experience in real-world practice. Future prospective studies performed in expert centres are needed to assess in which high-risk patients TBLC can be performed relatively safely.

Discussion

This ERS guideline aimed to establish evidence-based recommendations for the use of TBLC in patients with undiagnosed ILD in clinical practice. The guideline was developed in line with GRADE principles and every guideline question was informed by a thorough systematic review of the published literature.

Several potential limitations should be taken into account. Although a considerable number of clinical experts were involved in the development of the guideline, these mostly included pulmonologists with expertise in TBLC, and only one pathologist and one radiologist. In future updates, we will consider expanding the Task Force to further ensure that all clinical stakeholders involved in the diagnostic process of ILD are sufficiently represented. For several PICO questions, TBLC and SLB were compared. In this comparison, we focused on the *a priori* defined and prioritised outcomes (supplementary appendix S4), and these were carefully weighed in the evidence-to-decision process to arrive at a final recommendation. However, whether or not a test is considered a "valid replacement test" (PICO Question 1) or a "valid add-on test" (PICO Question 4) may depend on many factors, including diagnostic yield, adverse events, downstream consequences of test results, prevalence and costs [55], and others may weigh the relative importance of each outcome differently. Although many studies on TBLC and SLB in patients with ILD have been published, the number of included studies was low for most of the guideline questions. In addition, certainty of the evidence was "very low" for most outcomes. The Task Force formulated recommendations for future research for each guideline question, which may form the basis for studies in upcoming years.

Taking into account the evidence obtained in PICO Questions 1 and 3, the Task Force believes that a step-up strategy is in most situations preferred (figure 1): patients would initially undergo TBLC (at reduced risk of severe adverse events, days of hospitalisation and costs) and if insufficiently informative, this would be followed by SLB. This diagnostic approach was also preferred by the three patient representatives who provided input. The Task Force acknowledges that the recommendations apply to centres that are experienced in performing TBLC. Furthermore, the spectrum of potential patients with ILD (with regard to severity of underlying illness, extent of comorbid disease, level of diagnostic certainty from clinical, laboratory and radiological data, and importance of obtaining a histopathological diagnosis) is broad in the clinical setting. As it is impossible to formulate recommendations that equally apply to every different situation, the advantages and disadvantages of performing invasive testing should be carefully weighed on a case-by-case level in each individual patient.

Acknowledgements: Markus Fally was the recipient of a European Respiratory Society Fellowship in Guidelines Methodology (MTF 2018-01). The Task Force would like to thank Achille Abbondanza, Steve Jones and John Solheim (from the European Lung Foundation's Pulmonary Fibrosis Patient Advisory Group) for their valuable input from a patient's perspective.

This document was endorsed by the ERS Executive Committee on 25 May 2022 and by ERN-LUNG on 5 June 2022.

the patient's caregiver where appropriate and/or necessary, and to verify rules and regulations applicable to drugs and devices at the time of prescription.

Conflict of interest: L. Hagmeyer has received honoraria for lectures and presentations from Boehringer Ingelheim and Roche, and participated in advisory boards for Boehringer Ingelheim and Roche. J. Hetzel has received honoraria for lectures and presentations from Erbe and GlaxoSmithKline, and research support from Boehringer Ingelheim and AstraZeneca. A. Morais has received honoraria for presentations from Boehringer Ingelheim, Roche, Pfizer, AstraZeneca and Sanofi, and research grants from Roche, Boehringer Ingelheim and GlaxoSmithKline. S. Tomassetti has received honoraria for presentations from Roche and Boehringer Ingelheim. L.K. Troy has received honoraria for presentations from Boehringer Ingelheim, has been a member of an advisory board for Roche, and has received research support from Erbe. A.U. Wells has received personal fees from Roche and Boehringer Ingelheim. T. Tonia acts as an ERS methodologist. V. Poletti has received honoraria for lectures and presentations from Boehringer Ingelheim, Roche and Erbe, and participated in advisory boards for Boehringer Ingelheim, Roche and Ambu. The remaining authors have nothing to disclose.

References

1 Walsh SLF, Wells AU, Desai SR, *et al.* Multicentre evaluation of multidisciplinary team meeting agreement on diagnosis in diffuse parenchymal lung disease: a case-cohort study. *Lancet Respir Med* 2016; 4: 557–565.
2 Walsh SLF, Maher TM, Kolb M, *et al.* Diagnostic accuracy of a clinical diagnosis of idiopathic pulmonary fibrosis: an international case–cohort study. *Eur Respir J* 2017; 50: 1700936.
3 Tomassetti S, Ravaglia C, Puglisi S, *et al.* Impact of lung biopsy information on treatment strategy of patients with interstitial lung diseases. *Ann Am Thorac Soc* 2022; 19: 737–745.
4 Costabel U, Guzman J. Bronchoalveolar lavage in interstitial lung disease. *Curr Opin Pulm Med* 2001; 7: 255–261.
5 Poletti V, Patelli M, Poggi S, *et al.* Transbronchial lung biopsy and bronchoalveolar lavage in diagnosis of diffuse infiltrative lung diseases. *Respiration* 1988; 54: Suppl. 1, 66–72.
6 Raghu G, Remy-Jardin M, Myers JL, *et al.* Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018; 198: e44–e68.
7 Sharp C, McCabe M, Adamali H, *et al.* Use of transbronchial cryobiopsy in the diagnosis of interstitial lung disease – a systematic review and cost analysis. *QJM* 2017; 110: 207–214.
8 Hutchinson JP, Fogarty AW, McKeever TM, *et al.* In-hospital mortality after surgical lung biopsy for interstitial lung disease in the United States. 2000 to 2011. *Am J Respir Crit Care Med* 2016; 193: 1161–1167.
9 Van Schil P. Cost analysis of video-assisted thoracic surgery *versus* thoracotomy: critical review. *Eur Respir J* 2003; 22: 735–738.
10 Poletti V, Casoni GL, Gurioli C, *et al.* Lung cryobiopsies: a paradigm shift in diagnostic bronchoscopy? *Respirology* 2014; 19: 645–654.
11 Hetzel J, Maldonado F, Ravaglia C, *et al.* Transbronchial cryobiopsies for the diagnosis of diffuse parenchymal lung diseases: expert statement from the Cryobiopsy Working Group on Safety and Utility and a call for standardization of the procedure. *Respiration* 2018; 95: 188–200.
12 Dhooria S, Agarwal R, Sehgal IS, *et al.* Bronchoscopic lung cryobiopsy: an Indian Association for Bronchology position statement. *Lung India* 2019; 36: 48–59.
13 Maldonado F, Danoff SK, Wells AU, *et al.* Transbronchial cryobiopsy for the diagnosis of interstitial lung diseases: CHEST Guideline and Expert Panel Report. *Chest* 2020; 157: 1030–1042.
14 Guo S, Li Q, Jiang J, *et al.* Chinese expert consensus on the standardized procedure and technique of transbronchial cryobiopsy. *J Thorac Dis* 2019; 11: 4909–4917.
15 McInnes MDF, Moher D, Thombs BD, *et al.* Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018; 319: 388–396.
16 Salameh JP, Bossuyt PM, McGrath TA, *et al.* Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. *BMJ* 2020; 370: m2632.
17 Whiting PF, Rutjes AW, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.
18 Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. Introduction – GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; 64: 383–394.
19 Schunemann HJ, Mustafa RA, Brozek J, *et al.* GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol* 2020; 122: 129–141.
20 Alonso-Coello P, Schunemann HJ, Moberg J, *et al.* GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: introduction. *BMJ* 2016; 353: i2016.

21 Suehs C, Bourdin A, Vachier I, *et al.* Transbronchial cryobiopsy in the diagnosis of interstitial lung diseases: methodologies and perspectives from the Cryo-PID and COLDICE studies. *Ann Transl Med* 2020; 8: 1330.

22 Sethi J, Ali MS, Mohananey D, *et al.* Are transbronchial cryobiopsies ready for prime time? A systematic review and meta-analysis. *J Bronchology Interv Pulmonol* 2019; 26: 22–32.

23 Romagnoli M, Colby TV, Berthet JP, *et al.* Poor concordance between sequential transbronchial lung cryobiopsy and surgical lung biopsy in the diagnosis of diffuse interstitial lung diseases. *Am J Respir Crit Care Med* 2019; 199: 1249–1256.

24 Troy LK, Grainge C, Corte TJ, *et al.* Diagnostic accuracy of transbronchial lung cryobiopsy for interstitial lung disease diagnosis (COLDICE): a prospective, comparative study. *Lancet Respir Med* 2020; 8: 171–181.

25 Ravaglia C, Bonifazi M, Wells AU, *et al.* Safety and diagnostic yield of transbronchial lung cryobiopsy in diffuse parenchymal lung diseases: a comparative study *versus* video-assisted thoracoscopic lung biopsy and a systematic review of the literature. *Respiration* 2016; 91: 215–227.

26 Tomassetti S, Wells AU, Costabel U, *et al.* Bronchoscopic lung cryobiopsy increases diagnostic confidence in the multidisciplinary diagnosis of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2016; 193: 745–752.

27 Tomassetti S, Ravaglia C, Wells AU, *et al.* Prognostic value of transbronchial lung cryobiopsy for the multidisciplinary diagnosis of idiopathic pulmonary fibrosis: a retrospective validation study. *Lancet Respir Med* 2020; 8: 786–794.

28 Hetzel J, Wells AU, Costabel U, *et al.* Transbronchial cryobiopsy increases diagnostic confidence in interstitial lung disease: a prospective multicentre trial. *Eur Respir J* 2020; 56: 1901520.

29 Hernández-González F, Lucena CM, Ramírez J, *et al.* Cryobiopsy in the diagnosis of diffuse interstitial lung disease: yield and cost-effectiveness analysis. *Arch Bronconeumol* 2015; 51: 261–267.

30 Guenther A, Krauss E, Tello S, *et al.* The European IPF registry (eurIPFreg): baseline characteristics and survival of patients with idiopathic pulmonary fibrosis. *Respir Res* 2018; 19: 141.

31 Richeldi L, Launders N, Martinez F, *et al.* The characterisation of interstitial lung disease multidisciplinary team meetings: a global study. *ERJ Open Res* 2019; 5: 00209-2018.

32 Colella S, Haentschel M, Shah P, *et al.* Transbronchial lung cryobiopsy in interstitial lung diseases: best practice. *Respiration* 2018; 95: 383–391.

33 Teoh AKY, Holland AE, Morisset J, *et al.* Essential features of an interstitial lung disease multidisciplinary meeting: an international Delphi survey. *Ann Am Thorac Soc* 2022; 19: 66–73.

34 Lodhi T, Hughes G, Stanel S, *et al.* Transbronchial lung cryobiopsy in idiopathic pulmonary fibrosis: a state of the art review. *Adv Ther* 2019; 36: 2193–2204.

35 Matta A, Gupta E, Swank Z, *et al.* The use of transbronchial cryobiopsy for diffuse parenchymal lung disease in critically ill patients with acute hypoxemic respiratory failure – a case series. *Clin Respir J* 2021; 15: 788–793.

36 Bondue B, Schlossmacher P, Allou N, *et al.* Trans-bronchial lung cryobiopsy in patients at high-risk of complications. *BMC Pulm Med* 2021; 21: 135.

37 Ravaglia C, Wells AU, Tomassetti S, *et al.* Diagnostic yield and risk/benefit analysis of trans-bronchial lung cryobiopsy in diffuse parenchymal lung diseases: a large cohort of 699 patients. *BMC Pulm Med* 2019; 19: 16.

38 Kronborg-White S, Sritharan SS, Madsen LB, *et al.* Integration of cryobiopsies for interstitial lung disease diagnosis is a valid and safe diagnostic strategy – experiences based on 250 biopsy procedures. *J Thorac Dis* 2021; 13: 1455–1465.

39 Hagmeyer L, Theegarten D, Wohlschlager J, *et al.* Transbronchial cryobiopsy in fibrosing interstitial lung disease: modifications of the procedure lead to risk reduction. *Thorax* 2019; 74: 711–714.

40 Bondue B, Leduc D, Froidure A, *et al.* Usefulness of surgical lung biopsies after cryobiopsies when pathological results are inconclusive or show a pattern suggestive of a nonspecific interstitial pneumonia. *Respir Res* 2020; 21: 231.

41 DiBardino DM, Haas AR, Lanfranco AR, *et al.* High complication rate after introduction of transbronchial cryobiopsy into clinical practice at an academic medical center. *Ann Am Thorac Soc* 2017; 14: 851–857.

42 Naur TMH, Nilsson PM, Pietersen PI, *et al.* Simulation-based training in flexible bronchoscopy and endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA): a systematic review. *Respiration* 2017; 93: 355–362.

43 Sehgal IS, Dhooria S, Aggarwal AN, *et al.* Training and proficiency in endobronchial ultrasound-guided transbronchial needle aspiration: a systematic review. *Respirology* 2017; 22: 1547–1557.

44 Vilmann P, Clementsen PF, Colella S, *et al.* Combined endobronchial and oesophageal endosonography for the diagnosis and staging of lung cancer. European Society of Gastrointestinal Endoscopy (ESGE) Guideline, in cooperation with the European Respiratory Society (ERS) and the European Society of Thoracic Surgeons (ESTS). *Eur Respir J* 2015; 46: 40–60.

45 Davidsen JR, Skov IR, Louw IG, *et al.* Implementation of transbronchial lung cryobiopsy in a tertiary referral center for interstitial lung diseases: a cohort study on diagnostic yield, complications, and learning curves. *BMC Pulm Med* 2021; 21: 67.

46 Almeida LM, Lima B, Mota PC, *et al.* Learning curve for transbronchial lung cryobiopsy in diffuse lung disease. *Rev Port Pneumol* 2018; 24: 23–31.

47 Kronborg-White S, Folkersen B, Rasmussen TR, *et al.* Introduction of cryobiopsies in the diagnostics of interstitial lung diseases – experiences in a referral center. *Eur Clin Respir J* 2017; 4: 1274099.

48 Aburto M, Perez-Izquierdo J, Agirre U, *et al.* Complications and hospital admission in the following 90 days after lung cryobiopsy performed in interstitial lung disease. *Respir Med* 2020; 165: 105934.

49 She S, Steinfort DP, Ing AJ, *et al.* Transbronchial cryobiopsy in interstitial lung disease: safety of a standardized procedure. *J Bronchology Interv Pulmonol* 2020; 27: 36–41.

50 Cooley J, Balestra R, Aragaki-Nakahodo AA, *et al.* Safety of performing transbronchial lung cryobiopsy on hospitalized patients with interstitial lung disease. *Respir Med* 2018; 140: 71–76.

51 Gershman E, Fruchter O, Benjamin F, *et al.* Safety of cryo-transbronchial biopsy in diffuse lung diseases: analysis of three hundred cases. *Respiration* 2015; 90: 40–46.

52 Hetzel J, Eberhardt R, Petermann C, *et al.* Bleeding risk of transbronchial cryobiopsy compared to transbronchial forceps biopsy in interstitial lung disease – a prospective, randomized, multicentre cross-over trial. *Respir Res* 2019; 20: 140.

53 Kropski JA, Pritchett JM, Mason WR, *et al.* Bronchoscopic cryobiopsy for the diagnosis of diffuse parenchymal lung disease. *PLoS One* 2013; 8: e78674.

54 Pannu J, Roller LJ, Maldonado F, *et al.* Transbronchial cryobiopsy for diffuse parenchymal lung disease: 30- and 90-day mortality. *Eur Respir J* 2019; 54: 1900337.

55 Korevaar DA, Gopalakrishna G, Cohen JF, *et al.* Targeted test evaluation: a framework for designing diagnostic accuracy studies with clear study hypotheses. *Diagn Progn Res* 2019; 3: 22.