Appendix 7: PICO question 1 evidence synthesis

Tables included in this appendix:

Table 1: QUADAS-2 assessment of prospective studies directly comparing TBLC and SLB in patients with ILD

First author	Q1a.1	Q1a.2	Q1a.3	**Could the selection of patients have introduced bias?**	**Are there concerns that the included patients do not match the review question?**	Q2a.1	**Could the conduct or interpretation of the index test have introduced bias?**	**Are there concerns that the index test, its conduct, or its interpretation differ from the review question?**	Q3a.1	Q3a.2	**Could the reference standard, its conduct or its interpretation have introduced bias?**	**Are there concerns that the target condition as defined by the reference standard does not match the review question?**	Q4a.1	Q4a.2	Q4a.3	Q4a.4	**Could the patient flow have introduced bias?**
Romagnoli, M	Yes	Yes	Yes	**No**	**No**	No	**Yes**	**No**	Yes	No	**Yes**	**No**	Yes	Yes	Yes	Yes	**No**
Troy, L	Yes	Yes	Yes	**No**	**No**	No	**Yes**	**No**	Yes	No	**Yes**	**No**	Yes	Yes	Yes	Yes	**No**

Legend:

- **Q1a.1:** Was a consecutive or random sample of patients enrolled?
- **Q1a.2:** Was a case-control design avoided?
- **Q1a.3:** Did the study avoid inappropriate exclusions?
- **Q2a.1:** Was the index test (assumed to be (MDD of) TBLC) performed without knowledge of the results of the reference standard (assumed to be (MDD of) SLB)?
- **Q3a.1:** Is the reference standard likely to correctly classify the target condition?
- **Q3a.2:** Were the reference standard results (assumed to be (MDD of) SLB) interpreted without knowledge of the results of the index test (assumed to be (MDD of) TBLC)***?***
- **Q4a.1:** Was there an appropriate interval between index tests and reference standard?
- **Q4a.2:** Did all patients included in the 2x2 table receive a reference standard (partial verification bias)?
- **Q4a.3:** Did all patients in the 2x2 table receive the same reference standard (differential verification bias)?
- **Q4a.4:** Were all patients included in the analysis (2x2 table)?

Abbreviations: ILD = interstitial lung disease. MDD = multidisciplinary discussion. SLB = surgical lung biopsy. TBLC = transbronchial lung cryobiopsy.

Table 2: Histopathological and diagnostic agreement in prospective studies directly comparing TBLC and SLB in patients with ILD

First author Year Country	**Tests performed**	**Numer of patients undergoing both tests**		**Diagnostic pattern**	**Histopathological agreement between TBLC and SLB for specific pattern**	**Diagnostic agreement between TBLC and final MDD***	**Diagnostic agreement between MDD TBLC and MDD SLB****	**Deemed helpful at MDD*****	**High or definite confidence diagnosis at MDD**	**Complications**
Romagnoli, M 2019 Italy	TBLC and SLB	21		*TBLC:* 17 (81%) *SLB:* 21 (100%)	*Percentage agreement (for specific pattern):* 38% (95%CI 18-62) *Kappa agreement (for specific pattern):* 0.22 (95%CI 0.01-0.44)	*Percentage agreement:* 48% (95%CI 26-70) *Kappa agreement:* 0.31 (95%CI 0.06-0.56)	-	-	-	*Serious adverse events TBLC:* -n=2: pneumothorax *Serious adverse events SLB:* -n=0
Troy, L 2020 Australia	TBLC and SLB	65		*TBLC:* 59 (91%) *SLB:* 63 (97%)	*Percentage agreement (for specific pattern):* 69.2% *Kappa agreement (for specific pattern):* 0.47 (95%CI 0.30-0.64) *Percentage agreement (for guideline-refined pattern):* 70.8% *Weighted Kappa agreement (for guideline-refined pattern):* 0.70 (95%CI 0.55-0.86)	-	*Percentage agreement:* 76.9% *Kappa agreement:* 0.62 (95%CI 0.47-0.78)	*TBLC:* 48 (74%) *SLB:* 50 (77%) p=0.55	*MDD+TBLC:* 39 (60%) *MDD+SLB:* 48 (74%) p=0.090 *Additional:* 37/39 (95%) of MDD+TBLC high or definite confidence diagnoses were concordant with MDD+SLB diagnoses 6/26 (23%) of MDD+TBLC low confidence or unclassifiable diagnoses were reclassified to alternative high or definite confidence diagnosis in MDD+SLB	*Adverse events TBLC:* -n=14: mild airway bleeding -n=1: pneumothorax *Serious adverse events TBLC:* -n=0 *Adverse events SLB:* -n=1: chest wall wound infection *Serious adverse events SLB:* -n=1: rehospitalization due to chest pain -n=1: bleed requiring intervention *Adverse events either TBLC or SLB:* -n=1: hypotension from anaesthetic -n=1: desaturation during procedure -n=1: bronchospasm *Serious adverse events either TBLC or SLB:* -n=2: acute exacerbation of IPF -n=1: death within 90 days -n=1: rehospitalization due to mild hypoxia

Legend:

*In the Romagnoli study, MDD was informed by both the TBLC and SLB results.

**In the Troy study, two separate MDDs were undertaken: one informed by TBLC results, and one informed by SLB results.

***The addition of biopsy information was deemed helpful if it changed the diagnosis from low to high confidence or definite, or provided an unanticipated diagnosis (as compared to MDD that only included clinical details and imaging findings).

Abbreviations: CI = confidence interval. ILD = interstitial lung disease. MDD = multidisciplinary discussion. SLB = surgical lung biopsy. TBLC = transbronchial lung cryobiopsy.

Table 3: Diagnostic accuracy of TBLC for UIP/IPF in prospective studies directly comparing TBLC and SLB in ILD patients

First author Year Country	**Index test**	**Reference standard for UIP/IPF**	**Total number of patients**	**Number of patients with UIP/IPF according to the reference standard**	**Agreement between TBLC and SLB for definite or probable UIP versus indeterminate for UIP or other diagnosis**	**Sensitivity of index test for diagnosing UIP/IPF**	**Specificity of index test for diagnosing UIP/IPF***	**PPV of index test for diagnosing UIP/IPF**	**NPV of index test for diagnosing UIP/IPF**
Romagnoli, M 2019 Italy	TBLC histology	SLB histology (specific pattern)	21	8	NR	UIP: 63% (5/8) (95%CI 26-90)	UIP: 69% (9/13) (95%CI 39-9%)	UIP: 56% (5/9) (95%CI 23-85)	UIP: 75% (9/12) (95%CI 43-93)
	TBLC histology	MDD after TBLC and SLB**	9	9	NR	IPF: 67% (6/9) (95%CI 31-91)	IPF: 75% (9/12) (95%CI 43-93)	IPF: 67% (6/9) (95%CI 31-91)	IPF: 75% (9/12) (95%CI 43-93)
Troy, L 2020 Australia	TBLC histology	SLB histology (specific pattern)	65	39	*Percentage agreement:* 70.8% *Kappa agreement:* 0.70 (0.55-0.86)	UIP: 87% (34/39) (95%CI 72-95)	UIP: 73% (19/26) (95%CI 52-88)	UIP: 83% (34/41) (95%CI 67-92)	UIP: 79% (19/24) (95%CI 57-92)
	MDD after TBLC	MDD after SLB***	65	35	NR	IPF: 91% (32/35) (95%CI 76-98)	IPF: 80% (24/30) (95%CI 61-92)	IPF: 84% (32/38) (95%CI 68-93)	IPF: 89% (24/27) (95%CI 70-97)

Legend:

*Specificity was calculated as the number of patients with a 'non-UIP/IPF' diagnosis according to the index test, divided by the total number of patients with a 'non-UIP/IPF' diagnosis according to the reference standard. This implies that patients that were considered as 'true negatives' may still have had an index test result that was discrepant from the reference standard result (i.e. different 'non-UIP/IPF' diagnoses).

**Both the TBLC and SLB result were take into account in the MDD.

***Two separate MDDs were undertaken in this study: one including the TBLC results, and one including the SLB results. In this study, IPF diagnosis in MDD was categorized as 'definite', 'high probability' and 'low probability'; in the calculation of sensitivity, these subcategories were al considered as 'IPF positive'.

Abbreviations: CI = confidence interval. ILD = interstitial lung disease. IPF = idiopathic lung fibrosis. MDD = multidisciplinary discussion. NPV = negative preditive value. NR = not reported. PPV = positive predictive value. SLB = surgical lung biopsy. TBLC = transbronchial lung cryobiopsy. UIP = usual interstitial pneumonia.

Table 4: Studies performing indirect comparisons between TLBC and SLB in patients with ILD

First author Year Country	**Inclusion**	**Number of patients undergoing tests**	**Number of patients in whom a specific diangosis was obtained by the test**	**Diagnostic yield** Proportion of diagnostic tests	**Increase in IPF diagnostic confidence in MDD after addition of the test**	**Complications** Proportion of patients with a complication	**Other outcomes** Other outcomes related to testing
Ravaglia, C 2016 Italy	Patients with ILD in whom a diagnosis could not be achieved noninvasively	TBLC: -n=297 SLB: -n=150	TBLC: -n=246 SLB: -n=148	TBLC: -82.8% (246/297) SLB: -98.7% (148/150) -p=0.013	NR/NA	*Pneumothorax:* -TBLC: n=60 (20.2%) -SLB: NA *Pneumothorax requiring drainage:* -TBLC: n=46 (15.5%) -SLB: NA *Severe bleeding:* -TBLC: n=0 -SLB: n=0 *Mortality due to adverse event:* -TBLC: n=1 (0.3%) -SLB: n=4 (2.7%) -p=0.045	*Mean time of hospitalization:* -TBLC: 2.6 days (range 0-17) -SLB: 6.1 days (range 3-48) -p<0.0001
Tomassetti, S 2016 Italy	Patients with fibrotic ILD, without a typical UIP pattern on HRCT *All patients in this study were also included in Tomassetti 2020, which reports on other outcomes in a wider group of patients*	TBLC: -n=58 SLB: -n=59	NR/NA	NR/NA	TBLC: -From 29% to 63% p=0.0003 SLB: -From 30% to 65% -p=0.0016	*Pneumothorax:* -TBLC: n=19 (32.8%) -SLB: NA *Pneumothorax requiring drainage:* -TBLC: n=15 (25.9%) -SLB: NA *Severe bleeding:* -TBLC: n=0 -SLB: n=0 *Mortality:* -TBLC: n=1 (1.7%) -SLB: n=2 (3.4%)	*Mean time of hospitalization:* -TBLC: 3 days (range 0-9) -SLB: 6 days (range 3-17) -p-value NR
Tomassetti, S 2020 Italy	Patients with suspected ILD, without a definite UIP pattern on HRCT	TBLC: -n=266 SLB: -n=160	NR/NA	NR/NA	NR/NA	NR	*Mortality in MDD diagnosis of IPF versus other ILD:* -TBLC: adjusted HR 2.98 (95%CI 1.19-7.47; p=0.02) -SLB: adjusted HR 4.07 (95%CI 2.01-8.24; p<0·0001) *Mortality in UIP pattern versus other patterns:* -TBLC: adjusted HR 2.64 (95%CI 1.11-6.36; p=0.03) -SLB: adjusted HR 4.87 (95%CI 2.27-10.42; p=0.002)

Legend:

Abbreviations: CI = confidence interval. HR = hazard ratio. HRCT = high resolution comuted tomography. ILD = interstitial lung disease. IPF = idiopathic lung fibrosis. MDD = multidisciplinary discussion. NPV = negative preditive value. NA = not applicable. NR = not reported. PPV = positive predictive value. SLB = surgical lung biopsy. TBLC = transbronchial lung cryobiopsy. UIP = usual interstitial pneumonia.

Table 5: Recent systematic reviews on the diagnostic yield and complication rate of TBLC and SLB in patients with ILD

First author Year Country	**Test**	**Selection criteria**	**Searching details**	**Number of studies and patients included**	**Meta-analysis results: Diagnostic yield** Proportion of patients with a diagnostic test	**Meta-analysis results: Diagnostic yield in subgroups** Proportion of patients with a diagnostic test in subgroups	**Meta-analysis results: Complications** Proportion of patients with a complication	**Study designs and study quality assessment**
Sethi, J 2019 USA	TBLC	*Inclusion criteria:* -TBLC in patients with suspected DPLD -Diagnosis confirmed based on characteristic histopathologic findings or after MDD -Data provided on diagnostic yield or complications *Exclusion criteria:* -<10 patients included -TBLC performed for pulmonary nodules -Review articles -No language restrictions	*Sources searched:* -Medline -Embase -Google scholar -Reference lists -Conference abstract proceedings *Date of searching:* -12-2016 *Unique search results:* -n=252	*Studies included in systematic review:* -n=31 (n=18 full-texts; n=13 abstracts) -Published between 2009 and 2017 *Studies included in meta-analysis of diagnostic yield:* -n=27 *Patients included in meta-analysis of diagnostic yield:* -n=1443 -Range of patients across studies: 10-300	*Summary diagnostic yield (n=27 studies):* -72.9% (95%CI 67.9-77.7) *Range of diagnostic yield across studies:* -40.0% to 95.1%	*Summary diagnostic yield based on study design:* *-Retrospective (n=16 studies):* 71.8% (95%CI 65.8-77.5) *-Prospective (n=11 studies):* 74.3% (95%CI 64.9-82.8) *Summary diagnostic yield based on publication type:* *-Abstract (n=12 studies):* 71.4% (95%CI 63.9-78.3) *-Full-text (n=15 studies):* 74.0% (95%CI 67.2-80.3) *Summary diagnostic yield based on probe size:* *-1.9mm only (n=7 studies):* 70.4% (95%CI 58.8-80.8) *Summary diagnostic yield based on QUADAS-2:* *-Low risk of bias only (n=6 studies):* 73.1% (95%CI 63.0-82.1)	*Overall complication rate (n=31 studies):* -23.1% *Summary incidence of pneumothorax (n=30 studies):* -9.4% (95%CI 6.7-12.5%) *Summary incidence of moderate-severe bleed (n=27 studies):* -14.2% (95%CI 7.9-21.9%) *Summary incidence of mortality within 30 days (n=33 studies):* -0.3% (6 events in total)	*Study design:* -Prospective: n=11 (35.5%) -Retrospective: n=20 (64.5%) *QUADAS-2 assessment:* -High or unclear risk of bias: n=25 (80.6%)
Sharp, C 2017 UK	VATS	*Inclusion criteria:* -VATS-biopsy in patients with ILD *Exclusion criteria:* -No language restrictions	*Sources searched:* -Medline -Embase *Date of searching:* -02-2016 *Unique search results:* -n=166	*Studies included in systematic review and meta-analysis:* -n=24 -Published between 1992 and 2015 *Patients included in meta-analysis of diagnostic yield:* -n=2665 -Range of patients across studies: 30-432	*Summary diagnostic yield (n=24 studies):* -91.1% (95%CI 86.9-93.2) *Range of diagnostic yield across studies:* -NR	NR	*Summary incidence of surgical morbidity (n=18 studies):* -12.9% (95%CI 9.3-16.9) *Summary incidence of mortality within 30 days (n=21 studies):* -2.3% (95%CI 1.3-3.6)	*Study design:* -Prospective: n=3 (12.5%) -Retrospective: n=21 (87.5%) *Cochrane Collaboration risk of bias tool assessment:* -High risk of selection bias: n=24 (100%)

Legend:

Abbreviations: CI = confidence interval. DPLD: diffuse parenchymal lung disease. ILD = interstitial lung disease. MDD = multidisciplinary discussion. NR = not reported. SLB = surgical lung biopsy. TBLC = transbronchial lung cryobiopsy. VATS = video-assistend thoracic surgery.

Table 6: Studies reporting on MDD diagnostic confidence before and after TBLC in patients with ILD

First author Year Country	**Test**	**Patients**	**Patient included**	**Increase in diagnostic confdience at MDD**
Hetzel, J 2020 Germany	TBLC	Suspected IIP	128	*Percentage increase in confident diagnosis (likelihood ⩾90%) or provisional diagnosis with high confidence (likelihood ⩾70%):* -50.0% after clinicoradiological discussion -60.2% after BAL -81.2% after TBLC -p<0.0001 (TBLC vs BAL) *Percentage increase in confident diagnosis (likelihood ⩾90%):* -11.7% after clinicoradiological discussion -22.7% after BAL -53.9% after TBLC -p=0.001 (TBLC vs BAL)
Tomassetti, S 2015 Italy	TBLC SLB	Fibrotic ILD	117 *58 TBLC* *59 SLB*	*Percentage increase in IPF diagnosis made with high level of confidence in MDD:* *TBLC:* -29% after clinicoradiological discussion -63% after TBLC -p=0.0003 *SLB:* -30% after clinicoradiological discussion -65% after SLB -p=0.0016

Legend:

Abbreviations: BAL = bronchoalveolar lavage. IIP = idiopathic interstitial pneumonia. ILD = interstitial lung disease. MDD = multidisciplinary discussion. SLB = surgical lung biopsy.

TBLC = transbronchial lung cryobiopsy.

Table 7: GRADE tables for PICO question 1

PICO question:

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

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Diagnostic agreement between TBLC and final MDD									
2[a,1,2]	non-randomised trials	serious[b]	not serious	serious[c]	serious[d]	none	**Romagnoli et al** (n=21; diagnostic agreement between TBLC and TBLC+SLB+MDD): *Percentage agreement:* 48% (95%CI 26-70). *Kappa agreement:* 0.31 (95%CI 0.06-0.56).[e] **Troy et al** (n=65; diagnostic agreement between TBLC+MDD and SLB+MDD): *Percentage agreement:* 76.9% (95%CI NR). *Kappa agreement:* 0.62 (95%CI 0.47-0.78).[f]	⨁◯◯◯ Very low	CRITICAL[s]
High confidence final diagnosis at TBLC+MDD versus SLB+MDD									
1[a,2]	non-randomised trials	serious[b]	not serious	serious[c]	serious[d]	none	**Troy et al** (n=65): TBLC+MDD: 60% (39/65); TBLC+MDD: 74% (48/65); p=0.090. Also, 95% (37/39) of TBLC+MDD high or definite confidence diagnoses were concordant with SLB+MDD diagnoses. And 23% (6/26) of MDD+TBLC low confidence or unclassifiable diagnoses were reclassified to alternative high or definite confidence diagnosis in MDD+SLB.[f]	⨁◯◯◯ Very low	CRITICAL
Increase in MDD diagnostic confidence									
2[3]	observational studies	serious[g]	not serious	not serious	serious[h]	none	**Hetzel et al** (n=128): increase in confident diagnosis or provisional diagnosis with high confidence in MDD from 50.0% to 81.2% (p<0.0001) after TBLC. **Tomassetti et al** (n=117, 58 TBLC, 59 SLB): increase in IPF diagnosis with high level of confidence in MDD from 29% to 63% (p=0.0003) for TBLC, and from 30% to 65% (p=0.0016) for SLB.	⨁◯◯◯ Very low	CRUCIAAL
Diagnostic yield of TBLC versus SLB									
2[a,1,2,4]	non-randomised trials	serious[b]	not serious	serious[c]	serious[i]	none	**Romagnoli et al** (direct comparison of TBLC versus SLB; n=21): *Percentage agreement:* 38% (95%CI 18–62). *Kappa agreement*: 0.22 (95%CI 0.01-0.44). *Diagnostic pattern:* 81% for TBLC, and in 100% for SLB. **Troy et al** (direct comparison of TBLC versus SLB; n=65): *Percentage agreement:* 70.8% (95%CI NR). Weighted Kappa agreement (for guideline-refined pattern):0.70 (95%CI 0.55.-0.86). *Diagnostic pattern:* 91% for TBLC, and 97% for SLB. **Ravaglia et al** (n=447, indirect comparison of TBLC and SLB): *Diagnostic yield:* 82.8% for TBLC and 98.7% for SLB (p=0.013).	⨁◯◯◯ Very low	CRITICAL
Diagnostic yield of TBLC									

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
27[5]	observational studies	serious[k]	not serious	serious[c]	not serious	none	*Summary diagnostic yield* after meta-analysis: 72.9% (95%CI 67.9-77.7).[l]	⨁◯◯◯ Very low	CRITICAL
Diagnostic yield of SLB									
24[a,6]	observational studies	serious[m]	not serious	serious[c]	not serious	none	*Summary diagnostic yield* after meta-analysis: 91.1 (95%CI 86.9–93.2).[n]	⨁◯◯◯ Very low	CRITICAL
Diagnostic accuracy of TBLC for diagnosing IPF									
2[a,1,2]	non-randomised trials	serious[b]	not serious[i]	serious[c]	serious[j]	none	**Romagnoli et al** (n=21, accuracy of TBLC histology, against MDD informed by TBLC and SLB as reference standard): *Sensitivity:* 67% (95%CI 31-91). *Specificity*: 75% (95%CI 43-93).[o] **Troy et al** (n=65, accuracy of MDD informed by TBLC, against MDD informed by SLB as reference standard): *Sensitivity:* 91% (95%CI 76-98). *Specificity*: 80% (95%CI 61-92).[f,p]	⨁◯◯◯ Very low	CRITICAL
Survival after IPF diagnosis									
1[7]	observational studies	serious[q]	not serious	serious[r]	not serious	none	**Tomassetti et al** (indirect comparison of TBLC (n=266) versus SLB (n=160): an MDD diagnosis 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)).	⨁◯◯◯ Very low	CRITICAL
Adverse events of TBLC versus SLB: mortality									
1[4]	observational studies	serious[q]	not serious	serious[r]	not serious	none	**Ravaglia et al** (indirect comparison of TBLC (n=297) versus SLB (n=150)): *Mortality:* 0.3% (n=1) in TBLC versus 2.7% (n=4) in SLB (p=0.045). **Tomassetti et al** (indirect comparison of TBLC (n=58) and SLB (n=59)): *Mortality:* 1.7% (n=1) in TBLC versus 3.4% (n=2) in SLB.	⨁◯◯◯ Very low	CRITICAL
Adverse events of TBLC: mortality									
33[5]	observational studies	serious[b]	not serious	not serious	not serious	none	*Summary incidence of 30-day mortality*: 0.3%.[l]	⨁◯◯◯ Very low	CRITICAL
Adverse events of SLB: mortality									
21[6]	observational studies	serious[b]	not serious	not serious	not serious	none	*Summary incidence of 30-day mortality*: 2.3% (95%CI 1.3-3.6).[n]	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Adverse events of TBLC versus SLB: time of hospitalization									
2[4,8]	observational studies	serious[q]	not serious	serious[r]	not serious	none	**Ravaglia et al** (indirect comparison of TBLC (n=297) versus SLB (n=150)): *Mean time of hospitalization:* 2.6 days (range 0-17) for TBLC and 6.1 days (range 3-48) for SLB (p<0.0001). **Tomassetti et al** (indicrect comparison of TBLC (n=58) and SLB (n=59): *Mean time of hospitalization:* 3 days (range 0-9) for TBLC and 6 days (range 3-17) for SLB.	⨁◯◯◯ Very low	CRITICAL
Adverse events of TBLC versus SLB: other									
2[a,1,2]	non-randomised trials	serious[b]	not serious	serious[r]	serious[j]	none	**Romagnoli et al** (direct comparison of TBLC versus SLB (n=21)): *Serious adverse events:* 9.5% for TLBC (n=2 with pneumothorax), and 0% for SLB. **Troy et al** (direct comparison of TBLC versus SLB (n=65)): *Serious adverse events:* 0% for TBLC (additionally n=1 with pneumothorax was not considered as serious adverse event), and 3.1% for SLB (n=1 with rehospitalisation due to chest pain, and n=1 with bleeding requiring intervention). **Ravaglia et al** (indirect comparison of TBLC (n=297) versus SLB (n=150)): *Pneumothorax:* 15.5% for TBLC. *Severe bleeding:* 0% for TBLC, and 0% for SLB. **Tomassetti et al** (indicrect comparison of TBLC (n=58) and SLB (n=59): *Pneumothorax:* 25.9% for TBLC. *Severe bleeding:* 0% for TBLC, and 0% for SLB.	⨁◯◯◯ Very low	CRITICAL
Adverse events of TBLC: other									
31[5]	non-randomised trials	serious[k]	not serious	not serious	not serious	none	*Overall complication rate:* 23.1%, 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).[l]	⨁◯◯◯ Very low	CRITICAL
Adverse events of TBLC: other									
18[6]	observational studies	serious[m]	not serious	not serious	not serious	none	*Summary incidence of surgical morbidity:* 12.9% (95%CI 9.3-16.9, based on 18 studies).[n]	⨁◯◯◯ Very low	CRITICAL

CI: confidence interval

Explanations

a. In the GRADE approach, appropriately designed test accuracy studies start as high certainty evidence.

b. Risk of bias was unclear in the index test domain for both studies, because it is unclear if TBLC may have been performed differently (e.g. taking less time for the procedure) with the knowledge that SLB would also be performed in the same patient. Risk of bias was high in the reference standard domain, because MDD was not blinded to TBLC results (for Romagnoli et al), or likely to be not completely blinded to TBLC results (Troy et al).

c. Unclear if a histopathological diagnosis, agreement, diagnostic accuracy or diagnostic yield sufficiently correspond to the final MDD-diagnosis and to patient-important outcomes .

d. Only one study; small number of included patients.

e. For Romagnoli et al, both the TBLC and SLB result were taken into account in the MDD.

f. Two separate MDDs were undertaken: one including the TBLC results, and one including the SLB results.

g. High risk of incorporation bias in both studies.

h. No confidence intervals reported around increase in diagnostic confidence.

i. Although the results substantially differ between the two included studies, no downgrading for inconsistency was done as we already downgraded for risk of bias and imprecision, which could explain the inconsistency.

j. Studies not pooled; small number of included patients.

k. In the systematic review by Sethi et al on TBLC, risk of bias according to QUADAS-2 was high or unclear in 25 studies (80.6%).

l. Results from the systematic review by Sethi et al.

m. In the systematic review by Sharp et al on SLB, risk of selection bias according to the Cochrane Collaboration risk of bias tool assessment was high in 24 studies (100%).

n. Results from the systematic review by Sharp et al.

o. These accuracy estimates were not reported by Troy et al and Romagnoli et al, but could be recalculated.

p. These accuracy estimates were not reported by Troy et al, but could be recalculated.

q. High risk of selection bias, as no randomization was performed.

r. Indirect comparison of TBLC and SLB.

s. The outcome 'agreement' was not prespecified and addressed in the survey of assessment of outcome importance within the TF members , but was considered a surrogate of 'diagnostic accuracy', which was considered 'critical'

References

1. 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. American Journal of Respiratory & Critical Care Medicine. 2019;199(10):1249-1256.

2. Troy LK, Grainge C, Corte TJ, et al. Diagnostic accuracy of transbronchial lung cryobiopsy for interstitial lung disease diagnosis (COLDICE): a prospective, comparative study. The Lancet Respiratory Medicine. 2020;8(2):171-181.

3. Hetzel J, Wells AU, Costabel U, et al. Transbronchial cryobiopsy increases diagnostic confidence in interstitial lung disease: a prospective multicentre trial. European Respiratory Journal. 2020;56(6):12.

4. Ravaglia C, Bonifazi M, Wells AU, Tomassetti S, Gurioli C, Piciucchi S, Dubini A, Tantalocco P, Sanna S, Negri E, Tramacere I, Ventura VA, Cavazza A, Rossi A, Chilosi M, La Vecchia C, Gasparini S, Poletti V. 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(3): 215-227.

5. Sethi J, Ali MS, Mohananey D, Nanchal R, Maldonado F, Musani A. Are Transbronchial Cryobiopsies Ready for Prime Time?: A Systematic Review and Meta-Analysis. J Bronchology Interv Pulmonol. 2019;26(1):22-32.

6. Sharp C, McCabe M, Adamali H, Medford AR. Use of transbronchial cryobiopsy in the diagnosis of interstitial lung disease-a systematic review and cost analysis. Qjm. 2017;110(4):207-214.

7. Tomassetti S, Ravaglia C, Wells AU, Cavazza A, Colby TV, Rossi G, Ley B, Ryu JH, Puglisi S, Arcadu A, Marchi M, Sultani F, Martinello S, Donati L, Gurioli C, Gurioli C, Tantalocco P, Hetzel J, Dubini A, Piciucchi S, Klersy C, Lavorini F, Poletti V. Prognostic value of transbronchial lung cryobiopsy for the multidisciplinary diagnosis of idiopathic pulmonary fibrosis: a retrospective validation study. The Lancet Respiratory Medicine 2020: 8(8): 786-794.

8. Tomassetti S, Wells AU, Costabel U, Cavazza A, Colby TV, Rossi G, Sverzellati N, Carloni A, Carretta E, Buccioli M, Tantalocco P, Ravaglia C, Gurioli C, Dubini A, Piciucchi S, Ryu JH, Poletti V. Bronchoscopic Lung Cryobiopsy Increases Diagnostic Confidence in the Multidisciplinary Diagnosis of Idiopathic Pulmonary Fibrosis. American Journal of Respiratory & Critical Care Medicine 2016: 193(7): 745-752.

Table 8: Evidence to decision framework for PICO question 1

PICO question:

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

Problem

Is the problem a priority?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	The prevalence of ILD is estimated to be 6.3-76.0 per 100,000 people in Europe, and 74.3 per 100,000 in the USA. Of these 13-40% are estimated to develop progressive fibrosing ILD, with an overall prevalence estimate of 2.2-20.0 per 100,000 in Europe, and 28.0 per 100,000 in the USA. This reresents a considerable fraction of chronic respiratory disorders (Olson et al. Advances in Therapy 2021: 38:854-867). For the majority of patients with ILD, a MDD of clinical and radiological data results in a diagnosis. However, for around one third of these, MDD indicates that histopathological interpretation of a lung biopsy is needed. Currently, SLB is often performed in these patients, with high costs and high complication rates: Summary incidence of surgical morbidity (n=18 studies): 12.9% (95%CI 9.3-16.9%). Summary incidence of mortality within 30 days (n=21 studies): 2.3% (95%CI 1.3-3.6%).	

Desirable Effects

How substantial are the desirable anticipated effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Trivial ○ Small ● Moderate ○ Large ○ Varies ○ Don't know	Severe complications are anticipated to be lower in TBLC than SLB: -Overall mortatlity rate is lower: 0.3% versus 2.3% (based on the included meta-analyses of studies only focusing on TBLC or only focussing on SLB), and 0.3% versus 2.7% (based on one study indirectly comparing both tests). -Mean time of hospitalization is shorter: 2.6 days for TBLC and 6.1 days for SLB (based on one study indirectly comparing both tests), and 3 days for TBLC and 6 days for SLB (based on a second study indirectly comparing both tests). -Overall complication rate is higher: 23.1% versus 12.9% (based on the included meta-analyses of studies only focusing on TBLC or only focussing on SLB).	-Complication rates are difficult to compare considering the fact that (a) definitions of complications varied and (b) populations varied (e.g. the TBLC population may have also included patients not considered eligible to undergo SLB). -The Task Force put most emphasis on a potential reduction in serious adverse events (especially mortality). -Reported overall complication rate between TBLC and SLB cannot be compared: inTBLC-studies, pneumothorax is considered an adverse event, while in SLB-studies, it is not because all patients require chest tube drainage. -Complications are likely to be influenced by operator experience (see PICO question 4).

		-TBLC complications are generally lower in 'later' studies, where endobronchial balloons were used.

Undesirable Effects

How substantial are the undesirable anticipated effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Large ● Moderate ○ Small ○ Trivial ○ Varies ○ Don't know	Diagnostic accuracy is anticipated to be lower for TBLC than SLB: -Diagnostic agreement between TBLC+MDD and SLB+MDD is 76.9% (based on one study directly comparing both tests). -95% of TBLC+MDD high or definite confidence diagnoses are concordant with SLB+MDD diagnoses; 23% of MDD+TBLC low confidence or unclassifiable diagnoses were reclassified to alternative high or definite confidence diagnosis in MDD+SLB (based on one study directly comparing both tests). -Increase in diagnostic confidence of MDD after adding TBLC is: from 60% to 81% (based on one study only performing TBLC). -Increase in IPF diagnosis made with high level of confidence in MDD is similar for TBLC and SLB: from 29% to 63% for TBLC, and from 30% to 65% for SLB (based on one study indirectly comparing both tests). -Histopathological agreement between TBLC and SLB is between 38% and 69.2% (based on two studies directly comparing both tests). -Diagnostic yield of TBLC is lower: 72.9% versus 91.1% (based on the included meta-analyses of studies only focusing on TBLC or only focussing on SLB). -Diagnostic accuracy of TBLC+MDD for diagnosing IPF is: sensitivity 91% and specificity 80% (based on one study).	Troy and colleagues and Romagnoli and collagues are both indirect comparisons of TBLC versus SLB, yet the first is considered to be at lower risk of bias, and has a much larger sample size, and therefore more relative weight was put to its results in the Task Force discussion.

Certainty of evidence

What is the overall certainty of the evidence of effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
● Very low ○ Low ○ Moderate ○ High ○ No included studies	Overall certainty of the evidence was ‘very low’.	

Values

Is there important uncertainty about or variability in how much people value the main outcomes?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability		-Some may favor a more accurate test. Others may favor a test with less adverse events and lower costs. -Most patients are unlikely to choose SLB, if TBLC (i.e. a less invasive test) is an alternative, especially taking into account that a step-up strategy may be proposed where patients could still undergo SLB after a non-diagnostic initial TBLC. -Summary of patient feedback (one patient who underwent TBLC, one who underwent SLB): “The evidence indicates that SLB is more likely to give an accurate answer than TBLC but is associated with higher risks. Given the data on the scale of these benefits and risks, we consider that most patients would opt for a TBLC but, if that does not work, would then prefer to have a SLB, rather than a second TBLC.” -Summary of patient feedback (one patient undergoing both TBLC and SLB): “I truly believe that TLBC should be the first technique to be proposed in case the diagnosis requires it.”

Balance of effects

Does the balance between desirable and undesirable effects favor the intervention or the comparison?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Favors the comparison ○ Probably favors the comparison ○ Does not favor either the intervention or the comparison ● Probably favors the intervention ○ Favors the intervention ○ Varies ○ Don't know		-Some may favor a more accurate test. Others may favor a test with less adverse events and lower costs. -In centers with sufficient experience in TBLC, the balance of effects probably leans towards performing TBLC instead of SLB.

Resources required

How large are the resource requirements (costs)?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Large costs ○ Moderate costs ○ Negligible costs and savings ● Moderate savings ○ Large savings ○ Varies ○ Don't know	Data on costs are limited. Two studies were identified that report some information.-Hernández-González et al (n=33): estimated that the systematic use of TBLC (followed by SLB if inconclusive) overall reduced costs up to 59846 euro (33 patients over a 3-year period), compared to systematically performing SLB. -Sharp et al (theoratical cost-analysis): estimated that the systematic use of TBLC (followed by SLB if inconclusive) reduced costs up to 647 pound per patient per year.	-It is generally accepted that TBLC results in lower costs than SLB. -A major cost driver is considered to be the number of days in the hospital, which is considered to be higher in SLB.

Certainty of evidence of required resources

What is the certainty of the evidence of resource requirements (costs)?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
● Very low ○ Low ○ Moderate ○ High ○ No included studies		-It is generally accepted that TBLC results in lower costs than SLB. -A major cost driver is considered to be the number of days in the hospital, which is considered to be higher in SLB.

Cost effectiveness

Does the cost-effectiveness of the intervention favor the intervention or the comparison?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Favors the comparison ○ Probably favors the comparison ○ Does not favor either the intervention or the comparison ○ Probably favors the intervention ○ Favors the intervention ○ Varies ● No included studies		-Studies on cost-effectiveness are not available. -It is generally accepted that TBLC results in lower costs than SLB. -It is unknown to which extent reduced diagnostic accuracy for TBLC results in higher costs down the line, compared to SLB.

Equity

What would be the impact on health equity?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know		Due to the anticipated lower proportion of serious adverse events of TBLC compared to SLB, also patients who are no candidates for SLB (e.g. due to poor respiratory status) can now be offered a diagnostic approach.

Acceptability

Is the intervention acceptable to key stakeholders?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know		Overall, diagnostic accuracy of the intervention test (TBLC) is considered lower than for the comparator test (SLB) which it aims to replace, at expected reduced costs and serious adverse events. These are likely to be the most important arguments for or against replacing SLB by TBLC. Some physicians or patients may weigh these advantages and disadvantages in favor of TBLC, others in favor of SLB.

Feasibility

Is the intervention feasible to implement?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know		TBLC has been implemented in many healthcare centers worldwide, as illustrated by the large number of studies evaluating diagnostic yield and/or complications of TBLC in patients with ILD (n=59) identified in our searches. It does require well-trained endoscopists (see PICO question 4) and pathologists, and TBLC-equipment.

SUMMARY OF JUDGEMENTS

	JUDGEMENT						
PROBLEM	No	Probably no	Probably yes	**Yes**		Varies	Don't know
DESIRABLE EFFECTS	Trivial	Small	**Moderate**	Large		Varies	Don't know
UNDESIRABLE EFFECTS	Large	**Moderate**	Small	Trivial		Varies	Don't know
CERTAINTY OF EVIDENCE	**Very low**	Low	Moderate	High			No included studies
VALUES	Important uncertainty or variability	Possibly important uncertainty or variability	**Probably no important uncertainty or variability**	No important uncertainty or variability			
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	**Probably favors the intervention**	Favors the intervention	Varies	Don't know
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	**Moderate savings**	Large savings	Varies	Don't know
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	**Very low**	Low	Moderate	High			No included studies
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	**No included studies**
EQUITY	Reduced	Probably reduced	Probably no impact	**Probably increased**	Increased	Varies	Don't know
ACCEPTABILITY	No	Probably no	**Probably yes**	Yes		Varies	Don't know
FEASIBILITY	No	Probably no	Probably yes	**Yes**		Varies	Don't know

TYPE OF RECOMMENDATION

Strong recommendation against the intervention	Conditional recommendation against the intervention	Conditional recommendation for either the intervention or the comparison	**Conditional recommendation for the intervention**	Strong recommendation for the intervention
○	○	○	●	○

CONCLUSIONS

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 centers experienced in performing TBLC.

Justification

Compared to SLB, it is expected that TBLC results in lower serious adverse events and costs, at the expense of lower diagnostic accuracy. These advantages and disadvanteges should be weighed in each individual patient. Overall, the Task Force considers the reduction in serious adverse events to outweigh the reduced diagnostic accuracy. This especially applies to patients considered at higher risk of surgical adverse events.

Subgroup considerations

Although evidence of safety of TBLC in high-risk groups was limited (PICO question 4), no considerable differences seem to exist in terms of adverse events in high- versus low-risk groups.

Implementation considerations

TBLC has already been implemented by many specialised clinics worldwide. TBLC does not need to be offered in any healthcare center monitoring or treating patients with ILD; patients can be referred for TBLC to a specialised clinic.

Monitoring and evaluation

For quality assurance, healthcare centers that offer TBLC or SLB are advised to keep track of important outcomes such as diagnostic yield and complications.

Research priorities

Additional direct comparisons between TBLC and SLB are recommended. Ideally, a large randomized trial is performed. In addition to outcomes related to diagnostic accuracy, complications 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).