Appendix 3: Selection criteria for study inclusion

General selection criteria (applying to each guideline question):

Studies were only considered for inclusion if they were performed in human subjects with (suspected) ILD. In case a study corresponded to multiple study reports with overlapping study periods, we included the most recent report. However, if a study corresponded to multiple study reports but each addressed another guideline question or different outcomes within a PICO, all relevant reports were included. Both prospective and retrospective studies were eligible for inclusion. We included both comparative studies (e.g. randomized trials comparing two types of tests, or comparative cross-sectional type studies in which multiple tests were applied and compared in the same group of patients), as well as non-comparative studies (e.g. diagnostic accuracy studies that evaluated the accuracy or yield of only one test). However, in case comparative studies were available for a specific PICO outcome, the non-comparative studies were only used as additional/supportive evidence.

We excluded:

- Case reports and studies including 10 patiënts or less
- Studies reported in languages other than English
- Studies only reported as conference abstract
- Commentaries, editorials and letters not reporting original research data
- Studies in children only
- Studies published before 2000

Specific selection criteria for PICO question 1:

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

Studies were included if they compared the diagnostic yield, diagnostic accuracy, diagnostic confidence, complication rate, costs and/or other patient important outcomes of TBLC and SLB in patients with undiagnosed ILD. We included studies in which TBLC and SLB were directly compared, either by applying both tests in the same group of patients (comparative cross-sectional type study), or by randomly assigning a group of patients to undergo TBLC versus SLB (randomized trial). However, we also included studies that evaluated the abovementioned outcomes by performing only TBLC or only SLB, or studies in whom a group of patients undergoing TBLC was compared with a group of patients undergoing SLB, but in whom the tests were not randomly assigned.

Specific selection criteria for 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?

Studies were included if they evaluate the diagnostic yield, diagnostic accuracy, diagnostic confidence, complication rate, costs and/or other patient important outcomes of TBLC in patients with undiagnosed ILD.

We included studies in patients explicitly not eligible to undergo SLB, but excluded studies in which patients were explicitly eligible to undergo SLB, or in whom this information was not explicitly reported. We included comparative studies in which a group of patients was randomly assigned to undergo TBLC versus MDD (without any intervention). However, we also included studies that evaluated the abovementioned outcomes by only performing TBLC in patients ineligible to undergo SLB.

Specific selection criteria for PICO question 3:

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

Studies were included if they evaluated the diagnostic yield, diagnostic accuracy, diagnostic confidence, complication rate, costs and/or other patient important outcomes of TBLC and/or SLB in patients with undiagnosed ILD who have already had a TBLC with inconclusive results. We only included studies in which patients already had a TBLC with inconclusive results, and excluded studies for which this was not the case or unclear. We included studies that evaluated these outcomes by performing only TBLC or only SLB in a group of patients (non-comparative cross-sectional type study), studies that evaluated these outcomes by performing both TBLC and SLB in a group of patients (comparative cross-sectional type study), and studies that randomly assigned a group of patients to undergo TBLC versus SLB or MDD (without intervention) (randomized trial).

Specific selection criteria for PICO question 4:

Is formal training in TBLC recommended to optimize diagnostic yield and minimize adverse events in patients with undiagnosed ILD?

Studies were included if they compared diagnostic yield, diagnostic accuracy, diagnostic confidence, complications rate and/or other patient important outcomes in operators undergoing specific training in the TBLC procedure with those undergoing no specific training. Both randomized and non-randomized comparative studies were eligible for inclusion. We will also include non-comparative studies on learning curves for these outcomes.

Specific selection criteria for narrative question 1:

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

Studies were included if they compared the diagnostic yield, diagnostic accuracy, diagnostic confidence, complication rate, costs and/or other patient important outcomes of TBLC and SLB or conventional forceps TBB in patients with undiagnosed ILD with the specific HRCT findings described in the PICO question. We included studies both in patients considered eligible and not considered eligible to undergo SLB. We included studies in which TBLC is directly compared with SLB, conventional forceps or MDD (without any intervention), either by applying two of these tests in the same group of patients (comparative cross-sectional type study), or by randomly assigning a group of patients to undergo TBLC versus SLB, conventional forceps or MDD (without any intervention) (randomized trial). However, we included studies that evaluated the abovementioned outcomes by performing only TBLC in patients with the reported specific HRCT findings.

Specific selection criteria for narrative question 2:

What are the procedural risks of TBLC in patients with undiagnosed ILD?

Studies were included if they evaluated complication or mortality rates of TBLC in patients at high procedural risk (as described in the narrative question; e.g. only patients included with an FVC <50% or age >65 years), or if they compared complication or mortality rates of TBLC in specific subgroups of patients with suspected ILD at higher versus lower procedural risk (e.g. a subgroup of patients with FVC <50% or age >65 years, is compared with a subgroup of patients with FVC ≥50% or age ≤65 years) within a single study. In the narrative question, clear cutoffs for subgroups of specific interest were defined (i.e., FVC <50%, DLCO <35%, PAPs >40 mmHg, or age >65 years-old), but studies using slightly different thresholds were also eligible; roughly a 20% deviation was considered acceptable. We included studies both in patients who were considered eligible and were not considered eligible to undergo SLB. Both studies directly comparing TBLC with SLB, as well as non-comparative studies in which only TBLC was applied (with or without a separate group of patients undergoing SLB) were included, as long as they compared complication rates or mortality rates in these specific subgroups. We excluded studies only reporting on SLB but not TBLC.