

**Supplementary Appendix**

Supplement to: Accuracy and Reproducibility of Intra-Operative Assessment on Tumor Spread Through Air Spaces (STAS) in Stage 1 Lung Adenocarcinomas

**Table of Contents:****Page****1. Supplementary Methods**

Method S1	Frozen section protocol for evaluation of lung cancer resection specimens	3
Method S2	Histologic evaluation by observers	3
Method S3	Tumor grading definition	3
Method S4	Root case analysis, univariate analysis, and multivariable logistic regression models	4

**2. Supplementary Results**

Results S1	Equivocal-STAS cases	5
------------	----------------------	---

**3. Supplementary Figures**

Figure S1	Prevalence of cases with STAS clusters and artefactual clusters by each observer in the first and second rounds of evaluation	6
Figure S2	Intra-observer concordance on interpretations of STAS and artefactual clusters in FS, FSP, and NFP in the first and second rounds of evaluation	7
Figure S3	Inter-observer agreement on interpretations of STAS and artefactual clusters in FS, FSP, and NFP in the first and second rounds of evaluation	8

**4. Supplementary Tables**

Table S1	Categories of explanations given by observers when re-classifying cases as equivocal-STAS	9
Table S2	Diagnostic yield of frozen section for the detection of tumor STAS for each observer in first and second rounds of evaluation	10
Table S3	Pairwise comparisons of inter-observer concordance on STAS and artifact in FS, FSP, and NFP	11-12
Table S4	Univariate analysis for inter-observer agreement in Frozen section slide	13-14
Table S5	Multivariate logistic regression model for inter-observer agreement in Frozen section slide	15

**5. References**

16

## **SUPPLEMENTARY METHODS**

### **Methods S1.**

#### **Frozen section protocol for evaluation of lung cancer resection specimens**

For the assessment on the main tumor in a lobectomy, a section from the largest cross section of the tumor with or without the adjacent lung parenchyma is obtained. In wedge and segmentectomy, a short segment of the staple line closet to the main tumor is removed with as little parenchymal tissue attached to the staple line as possible, the exposed parenchyma is inked, and a section perpendicular to the inked parenchymal margin is obtained to demonstrate the relation of the tumor and parenchymal margin. If ample surgical clearance is grossly identified, a section sampling the main tumor and a section from the parenchymal margin may be separately submitted, or only a section from the main tumor may be submitted without microscopic sampling of the parenchymal margin.

### **Methods S2.**

#### **Histologic evaluation by observers**

The participating observers were four pulmonary pathologists from MGH, and one invited pulmonary pathologist from the Philippines, with a range of 2–19 years of experience in pulmonary pathology. All observers were blinded to clinicopathologic data. The five observers independently reviewed FS and the corresponding FSP and NFP in each of the 100 cases and recorded the presence of STAS and artifact per published criteria in two independent sequential rounds, with a consensus conference four days after the first round. The intervening time interval between both rounds (washout period) was at least 5 weeks. The FS, FSP, and NFP slides were randomly organized and separately evaluated at different times by all observers for each round. Standard forms were used to record the presence of the histological findings. Observers were asked to strictly dichotomize cases as having STAS or not; thus, a binary output was generated from each observer’s evaluation of FS and FSP in both rounds. However, after giving a binary output the observers were given the possibility of re-classifying cases as “equivocal-STAS” if they felt the cases did not meet the published criteria entirely. They were also asked to give a detailed explanation of the reason they would call the case “equivocal-STAS”. The consensus conference was held and consisted of both a discussion of STAS published criteria[1] and a slide examination session of FS, FSP, and NFP from 15 selected cases that yielded low inter-observer concordance rates in round 1. Three additional pathologists participated in the consensus conference. Notes were recorded during the discussion. The consensus conference resulted in overall agreement on specific issues about published criteria,[1] and specific cases, and a consensus note was circulated to reiterate these points.

### **Methods S3.**

#### **Tumor grading definition**

We used a tumor grading system as defined by *Moreira et al*,**Error! Reference source not found.** which was based on the predominant histologic pattern and the extent of high-grade pattern component (solid, micropapillary, cribriform acinar, and other complex glandular [including fused glandular] patterns) present within the tumor. Grade 1 (Low grade) is defined as a lung adenocarcinoma with a predominant lepidic histologic pattern and any high-grade histologic pattern comprising less than 20% of the tumor. Grade 2 (Moderate grade) is defined as a lung adenocarcinoma with predominant acinar or papillary histologic pattern and any high-grade histologic pattern comprising less than 20% of the tumor. Grade 3 (High grade) is defined as a lung adenocarcinoma with any high-grade component comprising 20% or greater of the tumor.

#### **Methods S4.**

##### **Univariate analysis and multivariable logistic regression models to identify variables associated with inter-observer disagreement**

We performed a Root cause analysis (RCA) to identify possible variables associated with inter-observer disagreement. The RCA was conducted by one of the consensus panel pathologists involved in case selection but not in IOA analyses, and thus could provide an unbiased review. The pathologist reviewed the reasons for “equivocal-STAS” (Table S1) and also consulted the observers after both rounds of evaluation to get their insights based on personal experiences with the cases. Unadjusted univariate analyses were conducted to evaluate differences in variables described in Table 1 and those identified by the RCA between cases with an unanimous diagnosis (full-agreement group) and those with discrepant diagnoses (discrepancy group). The variables identified by the RCA included the number of STAS clusters recorded by the consensus panel and presence of artifacts scored by the majority ( $n \geq 3$ ) of the observers along with the consensus panel diagnosis. A Student’s nonpaired  $t$  test was applied for continuous variables and the chi-square test for categorical variables to assess the differences. Subsequently, we built multivariable logistic regression models with variables that showed  $p < 0.1$  in the unadjusted univariate analyses. Of note, the predominant pattern was excluded from the multivariate analyses given the significant overlap in definitions with histologic grade. The multivariable logistic regression analyses were performed with R, version 3.6.1.[3]

## **SUPPLEMENTARY RESULTS**

### **Results S1.**

#### **Equivocal-STAS cases**

A total of 16 (16%) cases were categorized as equivocal-STAS in FS by two or more observers in round 1; 63% (10/16) of them were categorized as STAS positive by the final integrated diagnosis based on review of all the histology slides by the consensus panel. A total of 46% (37/80) of the interpretations of these cases from the five observers were called STAS positive. Half of the cases were categorized as STAS positive by 3 or more observers. Equivocal-STAS cases were more likely to be called STAS positive by any of the observers [OR: 1.96 (95% CI: 1.2-3.2)], and to have high-grade (Grade 3) histology [OR: 5.5 (95% CI: 1.5-20.8)], than non-equivocal-STAS cases.

When re-classifying a case as Equivocal-STAS, the observers were asked to give a detailed explanation of the reasons they would consider the case as equivocal. The observers were allowed to give more than one explanation. We performed a qualitative analysis of the data to identify potential patterns of the explanations given by the observers, and we found similar reasons that were repeatedly used by different observers in different cases. We grouped all the different explanations into the following five major categories:

- 1- Location of the tumor cell clusters within the tissue
- 2- Morphological features of the tumor cell clusters
- 3- Quantity of tumor cell clusters
- 4- Background of artefactual clusters
- 5- Background tumor characteristic

Most observers classified cases as equivocal-STAS in 2 or 3 of the categories described above (range: 2-5). Four out of five observers included explanations that corresponded to the categories 1 and 2. The categories 1 and 4 had the highest number of explanations (25% each; 12/48). Additional information regarding the explanations is displayed in supplemental table S1.

## SUPPLEMENTARY FIGURES

Figure S1.

Prevalence of cases with STAS clusters and artefactual clusters by each observer in the first and second rounds of evaluation.

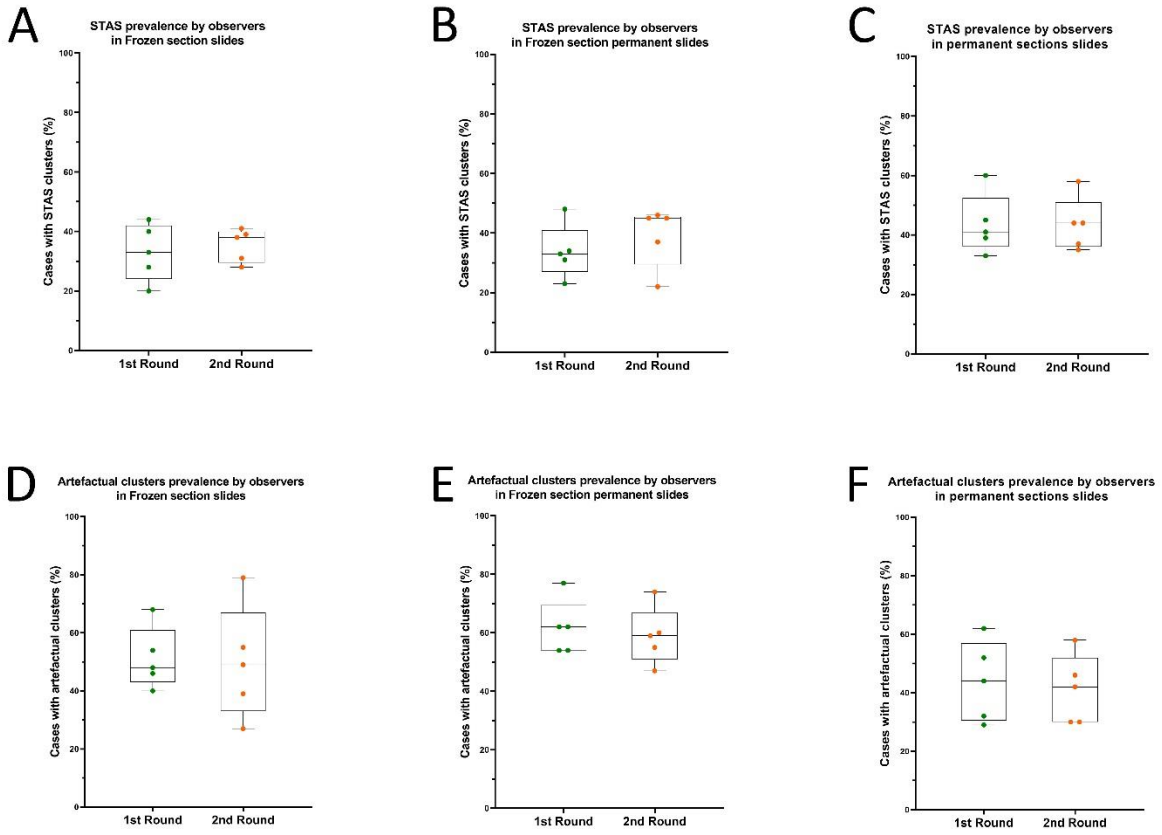
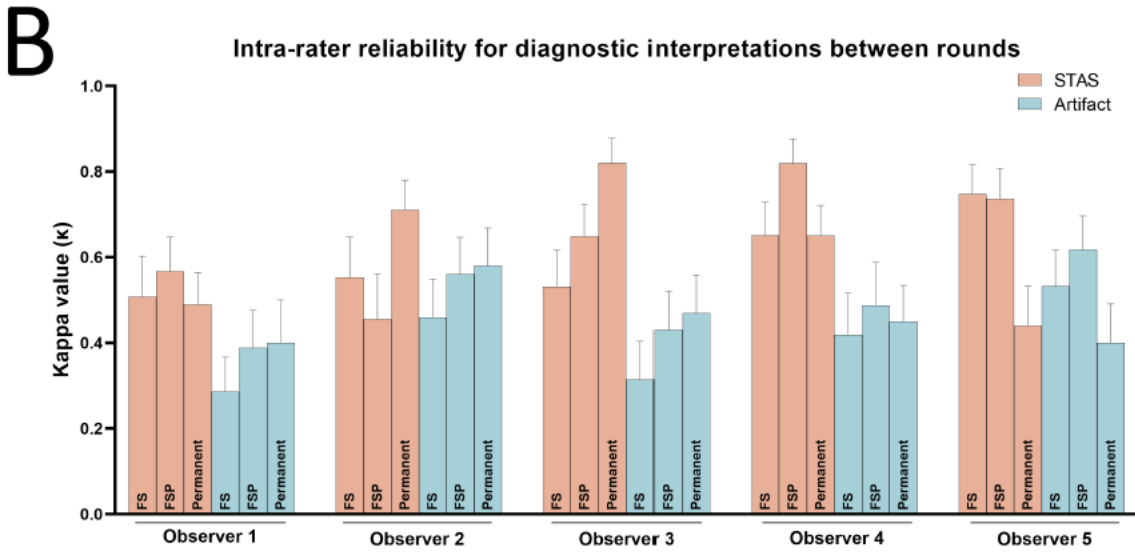
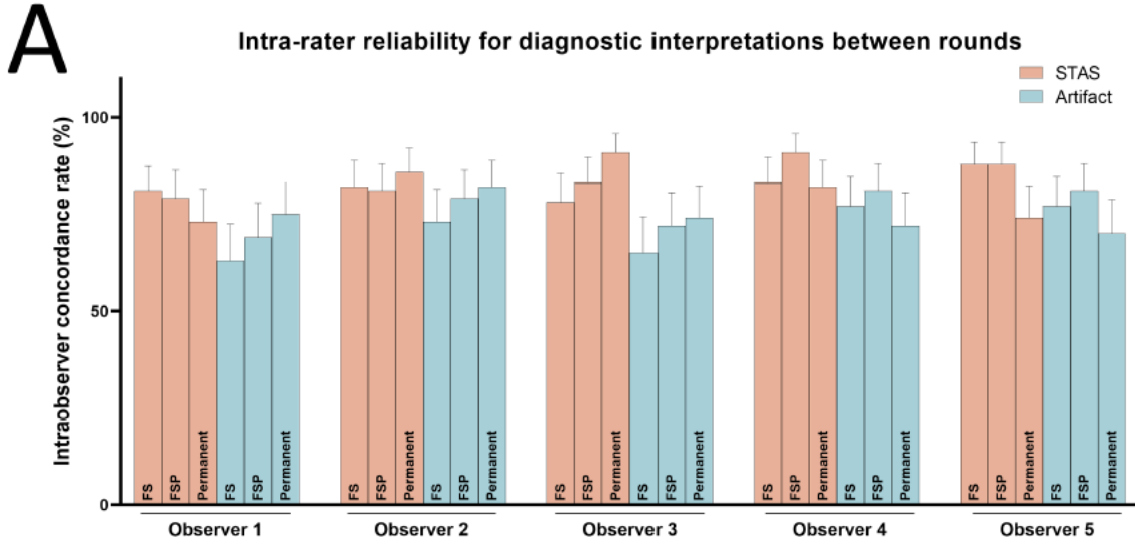
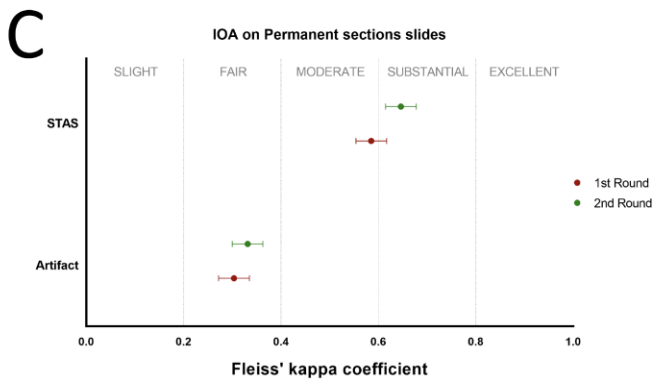
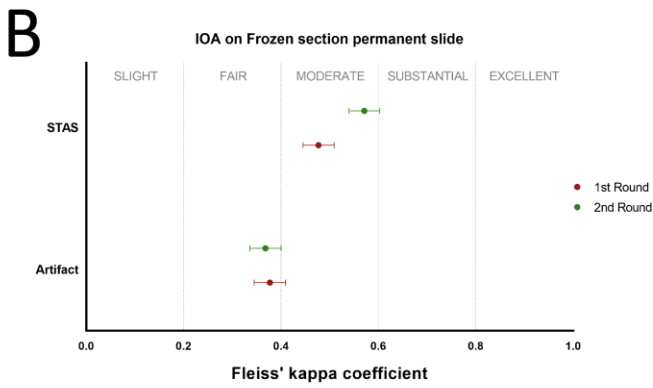
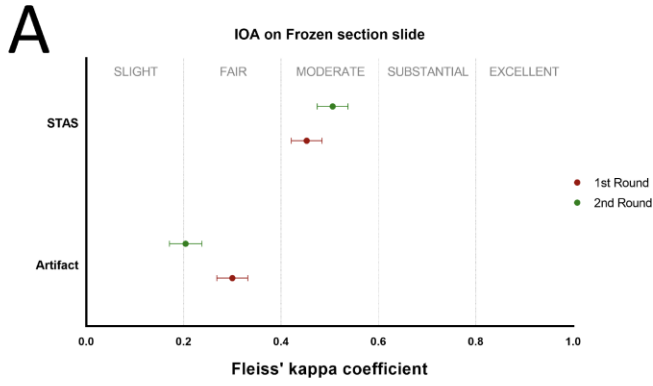


Figure S2.

Intra-observer concordance on interpretations of STAS and artefactual clusters in FS, FSP, and NFP in the first and second rounds of evaluation.



**Figure S3.**  
**Inter-observer agreement on interpretations of STAS and artefactual clusters in FS, FSP, and NFP in the first and second rounds of evaluation.**





## SUPPLEMENTARY TABLES

Table S1.

Categories of explanations given by observers when re-classifying cases as equivocal-STAS in FS

<u>Category</u>	<u>Subcategory</u>	<u>Explanations (%)</u> <u>(n=188)</u>
1.	Location of the tumor cell clusters within the tissue	25.0
	-Tumor cell clusters too close to the tumor edge	20.8
	-Tumor cell clusters too close to the edge of the slide	4.2
2.	Morphological features of the tumor cell clusters	20.8
	-Do not completely meet published criteria for STAS or artefactual clusters	10.4
	-Difficult to differentiate from artefactual cluster morphology	8.3
	-Cell aggregates are not tightly clustered	2.1
3.	Quantity of tumor cell clusters	22.9
	-Only one tumor cell cluster convincing for STAS	10.4
	-Rare tumor cell clusters present but without a trace	4.2
	-Tumor cell clusters with trace are present but are very rare	8.3
4.	Background of artefactual clusters	25.0
	-Extremely abundant artefactual clusters present in the background	18.8
	-Focally convincing STAS clusters in the setting of artefactual clusters	6.2
5.	Background tumor	6.3
	-Tumor with high-grade morphology (e.g. micropapillary pattern) which is commonly associated with STAS, but no STAS seen	6.3

**Supplementary table S2.**

**Diagnostic yield of frozen section for the detection of STAS for each observer in first and second rounds of evaluation**

Pathologists		Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)	Accuracy (95% CI)	ROC AUC (95% CI)
Observer 1	1st Round	34.9 (21.0-50.9)	91.2 (80.7-97.1)	75.0 (54.2-88.4)	65.0 (59.5-70.1)	4.0 (1.6-10.1)	0.7 (0.6-0.9)	67.0 (56.9-76.1)	0.63 (0.51-0.74)
	2nd Round	55.8 (39.9-70.9)	87.7 (76.3-94.9)	77.4 (62.0-87.8)	72.5 (65.0-78.9)	4.5 (2.2-9.6)	0.5 (0.4-0.7)	74.0 (64.3-82.3)	0.73 (0.63-0.83)
Observer 2	1st Round	51.2 (35.5-66.7)	89.5 (78.5-96.0)	78.6 (62.0-89.2)	70.8 (63.9-77.0)	4.9 (2.2-10.9)	0.5 (0.4-0.8)	73.0 (63.2-81.4)	0.70 (0.59-0.80)
	2nd Round	53.5 (37.7-68.8)	91.2 (80.7-97.1)	82.1 (65.6-91.8)	72.2 (65.1-78.4)	6.1 (2.5-14.7)	0.5 (0.4-0.7)	75.0 (65.3-83.1)	0.74 (0.63-0.84)
Observer 3	1st Round	53.5 (37.7-68.8)	82.5 (70.1-91.3)	69.7 (55.1-81.2)	70.1 (62.5-76.8)	3.0 (1.6-5.7)	0.6 (0.4-0.8)	70.0 (60.0-78.8)	0.69 (0.59-0.80)
	2nd Round	69.8 (53.9-82.8)	80.7 (68.1-90.0)	73.2 (60.8-82.8)	78.0 (68.8-85.0)	3.6 (2.1-6.4)	0.4 (0.2-0.6)	76.0 (66.4-84.0)	0.74 (0.64-0.84)
Observer 4	1st Round	76.7 (61.6-88.2)	80.7 (68.1-90.0)	75.0 (63.3-84.0)	82.1 (72.5-88.9)	4.0 (2.3-6.9)	0.3 (0.2-0.5)	79.0 (69.7-86.5)	0.80 (0.71-0.89)
	2nd Round	67.4 (51.5-80.9)	82.5 (70.1-91.3)	74.4 (61.4-84.1)	77.0 (68.2-84.0)	3.8 (2.1-7.0)	0.4 (0.3-0.6)	76.0 (66.4-84.0)	0.74 (0.64-0.84)
Observer 5	1st Round	62.8 (46.7-77.0)	77.2 (64.2-87.3)	67.5 (55.0-77.9)	73.3 (64.5-80.6)	2.8 (1.6-4.7)	0.5 (0.3-0.7)	71.0 (61.1-79.6)	0.69 (0.58-0.80)
	2nd Round	62.8 (46.7-77.0)	80.7 (68.1-90.0)	71.1 (57.9-81.4)	74.2 (65.7-81.2)	3.3 (1.8-5.8)	0.5 (0.3-0.7)	73.0 (63.2-81.4)	0.71 (0.6-0.81)

Supplementary Table 3A: Pairwise comparisons of inter-observer concordance on of STAS in FS, FSP, and NFP

FS	Round	Observer 2		Observer 3		Observer 4		Observer 5		All
		% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	
Observer 1	1st	82	0.511	77	0.422	68	0.310	72	0.364	<b>Mean Interobserver % Agreement (95% CI)</b>
	2nd	83	0.592	80	0.571	76	0.476	77	0.494	
Observer 2	1st			75	0.412	76	0.493	80	0.561	75.8 (72.9-78.7)
	2nd			83	0.631	73	0.402	80	0.553	77.4 (74.5-80.3)
Observer 3	1st					79	0.562	73	0.421	<b>Fleiss Kappa (mean <math>\kappa \pm</math> SE)</b>
	2nd					74	0.459	77	0.519	
Observer 4	1st							76	0.508	0.453 $\pm$ 0.032
	2nd							71	0.388	0.506 $\pm$ 0.032
FSP	Round	Observer 2		Observer 3		Observer 4		Observer 5		All
		% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	
Observer 1	1st	76	0.396	85	0.658	79	0.574	82	0.587	<b>Mean Interobserver % Agreement (95% CI)</b>
	2nd	74	0.456	85	0.698	83	0.657	79	0.571	
Observer 2	1st			69	0.250	67	0.325	74	0.363	76.6 (72.4-80.8)
	2nd			77	0.513	77	0.513	77	0.462	79.6 (76.8-82.4)
Observer 3	1st					76	0.514	83	0.618	<b>Fleiss Kappa (mean <math>\kappa \pm</math> SE)</b>
	2nd					86	0.717	78	0.548	
Observer 4	1st							75	0.493	0.477 $\pm$ 0.032
	2nd							80	0.589	0.571 $\pm$ 0.032
NFP	Round	Observer 2		Observer 3		Observer 4		Observer 5		All
		% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	
Observer 1	1st	82	0.627	86	0.702	73	0.494	80	0.568	<b>Mean Interobserver % Agreement (95% CI)</b>
	2nd	77	0.558	80	0.608	76	0.529	71	0.447	
Observer 2	1st			88	0.756	77	0.549	74	0.468	79.6 (75.9-83.3)
	2nd			91	0.814	87	0.732	86	0.696	82.6 (78.1-87.1)
Observer 3	1st					77	0.556	84	0.667	<b>Fleiss Kappa (mean <math>\kappa \pm</math> SE)</b>
	2nd					88	0.756	85	0.689	
Observer 4	1st							75	0.521	0.585 $\pm$ 0.032
	2nd							85	0.689	0.646 $\pm$ 0.032

FS: frozen section; FSP: frozen section permanent; NFP: non-frozen permanent; CI: confident interval

Supplementary Table 3B: Pairwise comparisons of inter-observer concordance on artifact in FS, FSP, and NFP

FS	Round	Observer 2		Observer 3		Observer 4		Observer 5		All
		% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	
Observer 1	1st	70	0.402	66	0.331	64	0.259	64	0.285	<b>Mean Interobserver % Agreement (95% CI)</b>
	2nd	60	0.193	32	0.082	42	0.084	66	0.243	
Observer 2	1st			70	0.395	64	0.290	68	0.358	65.0 (62.5-67.5)
	2nd			64	0.281	70	0.407	68	0.357	58.2 (49.6-66.8)
Observer 3	1st					60	0.254	60	0.187	<b>Fleiss Kappa (mean <math>\kappa \pm</math> SE)</b>
	2nd					62	0.193	58	0.178	
Observer 4	1st							64	0.300	0.300 $\pm$ 0.032
	2nd							60	0.291	0.204 $\pm$ 0.032

FSP	Round	Observer 2		Observer 3		Observer 4		Observer 5		All
		% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	
Observer 1	1st	78	0.533	72	0.429	75	0.426	68	0.347	<b>Mean Interobserver % Agreement (95% CI)</b>
	2nd	68	0.367	77	0.545	59	0.203	56	0.125	
Observer 2	1st			70	0.388	73	0.380	68	0.347	70.6 (67.3-73.9)
	2nd			77	0.523	79	0.540	74	0.470	69.4 (63.1-75.7)
Observer 3	1st					73	0.436	62	0.235	<b>Fleiss Kappa (mean <math>\kappa \pm</math> SE)</b>
	2nd					78	0.513	59	0.163	
Observer 4	1st							67	0.310	0.377 $\pm$ 0.032
	2nd							67	0.307	0.368 $\pm$ 0.032

NFP	Round	Observer 2		Observer 3		Observer 4		Observer 5		All
		% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa	
Observer 1	1st	79	0.514	72	0.412	62	0.307	61	0.238	<b>Mean Interobserver % Agreement (95% CI)</b>
	2nd	68	0.238	66	0.274	78	0.545	58	0.211	
Observer 2	1st			65	0.270	63	0.316	54	0.098	65.6 (60.6-70.6)
	2nd			70	0.359	74	0.463	62	0.286	67.6 (63.3-71.9)
Observer 3	1st					70	0.417	69	0.384	<b>Fleiss Kappa (mean <math>\kappa \pm</math> SE)</b>
	2nd					70	0.392	62	0.259	
Observer 4	1st							61	0.209	0.303 $\pm$ 0.032
	2nd							68	0.368	0.331 $\pm$ 0.032

FS: frozen section; FSP: frozen section permanent; NFP: non-frozen permanent; CI: confident interval

**Table S4A.****Univariate analysis for inter-observer agreement in Frozen section slide - 1<sup>st</sup> round of evaluation**

	Full agreement group 52	Controversy group 48	P
Age (years; mean (SD))	69.8 (11.3)	68.1 (8.6)	0.393
Sex Female / Male (%)	34 (65) / 18 (35)	33 (69) / 15 (31)	0.724
Total Size (cm; mean (SD))	1.5 (0.6)	1.7 (0.6)	0.221
Invasive Size (cm; mean (SD))	0.9 (0.8)	1.3 (0.7)	0.002
Type of operation (%)			0.158
Wedge resection	30 (58)	17 (35)	
Segmentectomy	4 (7.7)	6 (13)	
Wedge resection + completion lobectomy	7 (14)	8 (17)	
Lobectomy or other anatomical resections	11 (21)	17 (35)	
Dominant pattern (%)			0.059
Lepidic	22 (42)	9 (19)	
Acinar	16 (31)	12 (25)	
Papillary	2 (3.8)	6 (12.5)	
Micropapillary	3 (5.8)	7 (14.5)	
Solid	4 (7.7)	7 (14.5)	
Complex gland	5 (9.6)	7 (14.5)	
Lepidic ≥ 5% (%)	43 (83)	31 (65)	0.067
Micropapillary ≥ 5% (%)	18 (35)	28 (58)	0.030
Solid ≥ 5% (%)	8 (15)	12 (25)	0.342
Complex gland ≥ 5% (%)	15 (29)	22 (46)	0.121
Histologic grade (%)			0.008
1	22 (42)	7 (14.5)	
2	10 (19)	11 (23)	
3	20 (39)	30 (62.5)	
Lymphatic vessel invasion + (%)	8 (15)	6 (12.5)	0.899
Blood vessel invasion + (%)	1 (1.9)	4 (8.3)	0.312
Pleural invasion + (%)	1 (1.9)	5 (10)	0.172
Tumor necrosis + (%)	9 (17)	12 (25)	0.485
# of STAS clusters (%)			0.005
0	39 (75)	21 (44)	
1-4	4 (7.7)	6 (12.5)	
≥ 5	9 (17)	21 (44)	
Artifact present (%)**	15 (29)	35 (73)	<0.001
Consensus diagnosis of STAS (%)	14 (27)	29 (60)	0.001

\*\*The presence of artifact recorded by 3 or more observers

**Table S4B.****Univariate analysis and multivariate logistic regression model for inter-observer agreement in Frozen section slide – 2<sup>nd</sup> round of evaluation**

	Full agreement group 54	Controversy group 46	P
Age (years; mean (SD))	70.3 (11.7)	67.4 (7.7)	0.150
Sex Female / Male (%)	34(63) / 20 (37)	33 (72) / 13 (28)	0.473
Total Size (cm; mean (SD))	1.6 (0.7)	1.5 (0.5)	0.556
Invasive Size (cm; mean (SD))	1.0 (0.8)	1.2 (0.7)	0.218
Type of operation (%)			0.192
Wedge resection	30 (56)	17 (37)	
Segmentectomy	6 (11)	4 (8.7)	
Wedge resection + completion lobectomy	7 (13)	8 (17)	
Lobectomy or other anatomical resections	11 (21)	17 (37)	
Dominant pattern (%)			0.103
Lepidic	21 (39)	10 (22)	
Acinar	18 (33)	10 (22)	
Papillary	2 (3.7)	6 (13)	
Micropapillary	4 (7.4)	6 (13)	
Solid	4 (7.4)	7 (15)	
Complex gland	5 (9.3)	7 (15)	
Lepidic ≥ 5% (%)	44 (82)	30 (65)	0.105
Micropapillary ≥ 5% (%)	20 (37)	26 (57)	0.081
Solid ≥ 5% (%)	8 (15)	12 (25)	0.249
Complex gland ≥ 5% (%)	14 (26)	23 (50)	0.023
Histologic grade (%)			0.031
1	21 (39)	8 (17)	
2	12 (22)	9 (20)	
3	21 (39)	29 (63)	
Lymphatic vessel invasion+ (%)	7 (13)	7 (15)	0.972
Blood vessel invasion+ (%)	1 (1.9)	4 (8.7)	0.312
Pleural invasion + (%)	2 (3.7)	4 (8.7)	0.532
Tumor necrosis + (%)	8 (15)	13 (28)	0.162
# of STAS clusters (%)			0.055
0	38 (74)	22 (48)	
1-4	3 (5.6)	7 (15)	
≥ 5	13 (24)	17 (37)	
Artifact present (%)**	17 (32)	36 (78)	<0.001
Consensus diagnosis of STAS (%)	18 (33)	25 (54)	0.056

\*\*The presence of artifact recorded by 3 or more observers

**Table S5A.****Multivariate logistic regression model for inter-observer agreement in Frozen section slide - 1<sup>st</sup> round of evaluation**

		Odds Ratio* (95% CI)	Univariate P	Odds Ratio* (95% CI)	Multivariate P
Invasive Size (cm)	> 1.0 vs. ≤ 1.0	2.33 (1.32-4.11)	0.004	1.25 (0.47-3.32)	0.649
Lepidic pattern	≥ 5% vs. < 5%	0.38 (0.15-0.96)	0.043	0.93 (0.47-4.54)	0.910
Micropapillary pattern	≥ 5% vs. < 5%	2.64 (1.18-5.94)	0.019	1.45 (0.47-4.54)	0.519
Histologic Grade					
	2 vs. 1	3.46 (1.03-11.56)	0.440	2.26 (0.38-13.51)	0.896
	3 vs. 1	4.71 (1.70-13.09)	0.003	1.58 (0.25-10.20)	0.630
# of STAS clusters					
	1-4 vs. 0	2.79 (0.71-10.98)	0.143	1.13 (0.06-21.70)	0.936
	≥ 5 vs. 0	4.33 (1.69-11.13)	0.002	1.47 (0.09-23.4)	0.783
Artifact**	Present vs. absent		<0.001	5.84 (2.18-15.62)	<0.001
Consensus diagnosis of STAS	Present vs. absent	4.14 (1.78-9.62)	<0.001	1.44 (0.47-3.32)	0.649

\*Odds ratio for discrepant interpretations (controversy group vs. full agreement group); CI: confidence interval, \*\*The presence of artifact recorded by 3 or more observers

**Table S5B.****Multivariate logistic regression model for inter-observer agreement in Frozen section slide – 2<sup>nd</sup> round of evaluation**

		Odds Ratio* (95% CI)	Univariate P	Odds Ratio* (95% CI)	Multivariate P
Micropapillary pattern	≥ 5% vs. < 5%	2.21 (0.99-4.93)	0.053	1.16 (0.41-3.29)	0.786
Complex gland pattern	≥ 5% vs. < 5%	2.85 (1.23-6.61)	0.014	1.25 (0.37-4.24)	0.719
Histologic grade					
	2 vs. 1	1.97 (0.60-6.46)	0.264	1.18 (0.28-4.87)	0.818
	3 vs. 1	3.63 (1.35-9.75)	0.011	2.25 (0.47-10.74)	0.309
# of STAS clusters					
	1-4 vs. 0	4.03 (0.94-17.20)	0.060	4.67 (0.25-86.80)	0.301
	≥ 5 vs. 0	2.25 (0.92-5.52)	0.074	3.56 (0.23-55.09)	0.364
Artifact **	Present vs. absent	7.84 (3.17-19.39)	<0.001	7.71 (2.72-21.8)	<0.001
Consensus diagnosis of STAS	Present vs. absent	2.38 (1.06-5.35)	0.036	0.19 (0.01-3.05)	0.244

\*Odds ratio for discrepant interpretations (controversy group vs. full agreement group); CI: confidence interval, \*\*The presence of artifact recorded by 3 or more observers

## **REFERENCES**

- [1]. Kadota K, Nitadori JI, Sima CS, Ujiie H, Rizk NP, Jones DR, et al. Tumor Spread through Air Spaces is an Important Pattern of Invasion and Impacts the Frequency and Location of Recurrences after Limited Resection for Small Stage I Lung Adenocarcinomas. *J Thorac Oncol*. 2015 May;10(5):806-814. PMID: 25629637
- [2]. Moreira AL, Ocampo PS, Xia Y, Zhong H, Russell PA, Minami Y, et al. A Grading system for invasive pulmonary adenocarcinoma: a proposal from the IASLC pathology committee. *J Thorac Oncol*. 2020;15(10):1599-1610. PMID: 32562873
- [3]. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>. [Accessed on November 7, 2020]