

Supplementary Table 1: Showing the distribution of the Yokohama Category with the histopathological category

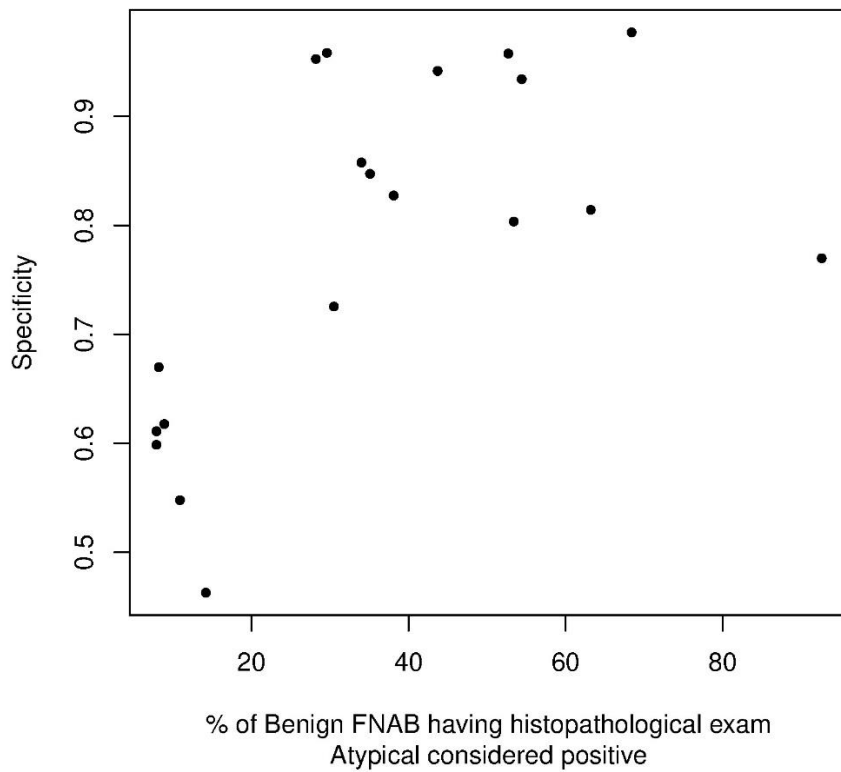
Study 1 st Author	year	Histopathology	FNAB Yokohama Category				
			Inadequate	Benign	Atypical	Suspicious	Malignant
Montezuma D [2]	2019	Benign	19	286	235	1	0
		Malignant	1	4	35	34	160
Chauhan V [3]	2019	Benign	1	234	3	2	0
		Malignant	0	0	1	8	82
Agarwal A [4]	2021	Benign	9	45	24	6	0
		Malignant	14	3	5	21	195
McHugh KE [5]	2019	Benign	20	91	12	13	5
		Malignant	0	12	4	11	51
De Rosa F [6]	2020	Benign	65	231	222	35	11
		Malignant	64	12	58	129	918
Wong S [7]	2019	Benign	38	170	98	7	0
		Malignant	1	3	16	43	199
Agrawal S [8]	2021	Benign	7	71	38	6	1
		Malignant	3	3	10	15	333
Dixit N [9]	2021	Benign	4	205	13	1	0
		Malignant	2	1	2	5	52
Marabi M [10]	2021	Benign	104	563	161	7	0
		Malignant	10	3	47	58	59
Wong YP [11]	2021	Benign	102	282	15	2	0
		Malignant	16	1	5	12	86
Tejeswini V [12]	2021	Benign	7	89	14	0	0
		Malignant	2	5	5	31	73
Sundar PM [13]	2022	Benign	13	152	25	0	2
		Malignant	8	1	7	18	61
Nigam JS [14]	2021	Benign	1	52	9	1	1
		Malignant	1	3	6	5	39
Agrawal N [15]	2021	Benign	21	280	56	2	0
		Malignant	4	2	17	32	199
Sarangi S [16]	2021	Benign	12	263	10	1	0
		Malignant	6	1	6	24	77
Nargund A [17]	2021	Benign	24	161	27	6	6
		Malignant	2	29	51	30	733
Ahuja S [18]	2021	Benign	19	129	19	2	0
		Malignant	1	2	4	9	57
Kamatar PV [19]	2019	Benign	2	90	2	1	1
		Malignant	0	4	4	5	70
Apuroopa M [20]	2020	Benign	19	326	70	4	0
		Malignant	1	4	10	59	116
Niaz M [21]	2022	Benign	6	104	34	16	1
		Malignant	5	12	15	77	151
Deshpande SA [22]	2021	Benign	2	96	33	3	0
		Malignant	0	2	4	18	47
Joshee A [23]	2021	Benign	5	214	8	1	0
		Malignant	1	7	4	21	102

Supplementary Table 2:

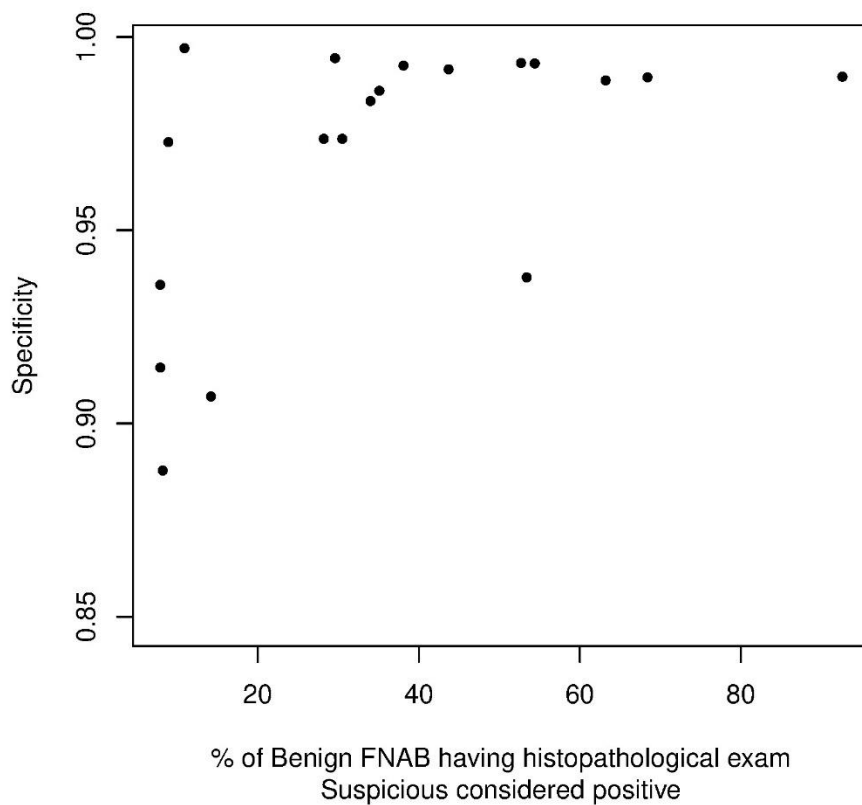
Author	year	Cutoff Atypical considered positive				Cutoff Suspicious considered positive				Cutoff malignant considered positive			
		TP	FN	FP	TN	TP	FN	FP	TN	TP	FN	FP	TN
Montezuma D [2]	2019	229	4	236	286	194	39	1	521	160	73	0	522
Chauhan V [3]	2019	91	0	5	234	90	1	2	237	82	9	0	239
Agarwal A [4]	2021	221	3	30	45	216	8	6	69	195	29	0	75
McHugh KE [5]	2019	66	12	30	91	62	16	18	103	51	27	5	116
De Rosa F [6]	2020	1105	12	268	231	1047	70	46	453	918	199	11	488
Wong S [7]	2019	258	3	105	170	242	19	7	268	199	62	0	275
Agrawal S [8]	2021	358	3	45	71	348	13	7	109	333	28	1	115
Dixit N [9]	2021	59	1	14	205	57	3	1	218	52	8	0	219
Marabi M [10]	2021	164	3	168	563	117	50	7	724	59	108	0	731
Wong YP [11]	2021	103	1	17	282	98	6	2	297	86	18	0	299
Tejeswini V [12]	2021	109	5	14	89	104	10	0	103	73	41	0	103
Sundar PM [13]	2022	86	1	27	152	79	8	2	177	61	26	2	177
Nigam JS [14]	2021	50	3	11	52	44	9	2	61	39	14	1	62
Agrawal N [15]	2021	248	2	58	280	231	19	2	336	199	51	0	338
Sarangi S [16]	2021	107	1	11	263	101	7	1	273	77	31	0	274
Nargund A [17]	2021	814	29	39	161	763	80	12	188	733	110	6	194
Ahuja S [18]	2021	70	2	21	129	66	6	2	148	57	15	0	150
Kamatar PV [19]	2019	79	4	4	90	75	8	2	92	70	13	1	93
Apuroopa M [20]	2020	185	4	74	326	175	14	4	396	116	73	0	400
Niaz M [21]	2022	243	12	51	104	228	27	17	138	151	104	1	154
Deshpande SA [22]	2021	69	2	36	96	65	6	3	129	47	24	0	132
Joshee A [23]	2021	127	7	9	214	123	11	1	222	102	32	0	223

Supplementary figure 1: Showing the relation between specificity reported by a study with percentage of benign FNAB having histopathology report in the same study. Studies with low % of FNAB with histological correlation (left side of the graphs also show a low specificity.

Relation between proportion of cases having histological follow up and specificity



Relation between proportion of cases having histological follow up and specificity



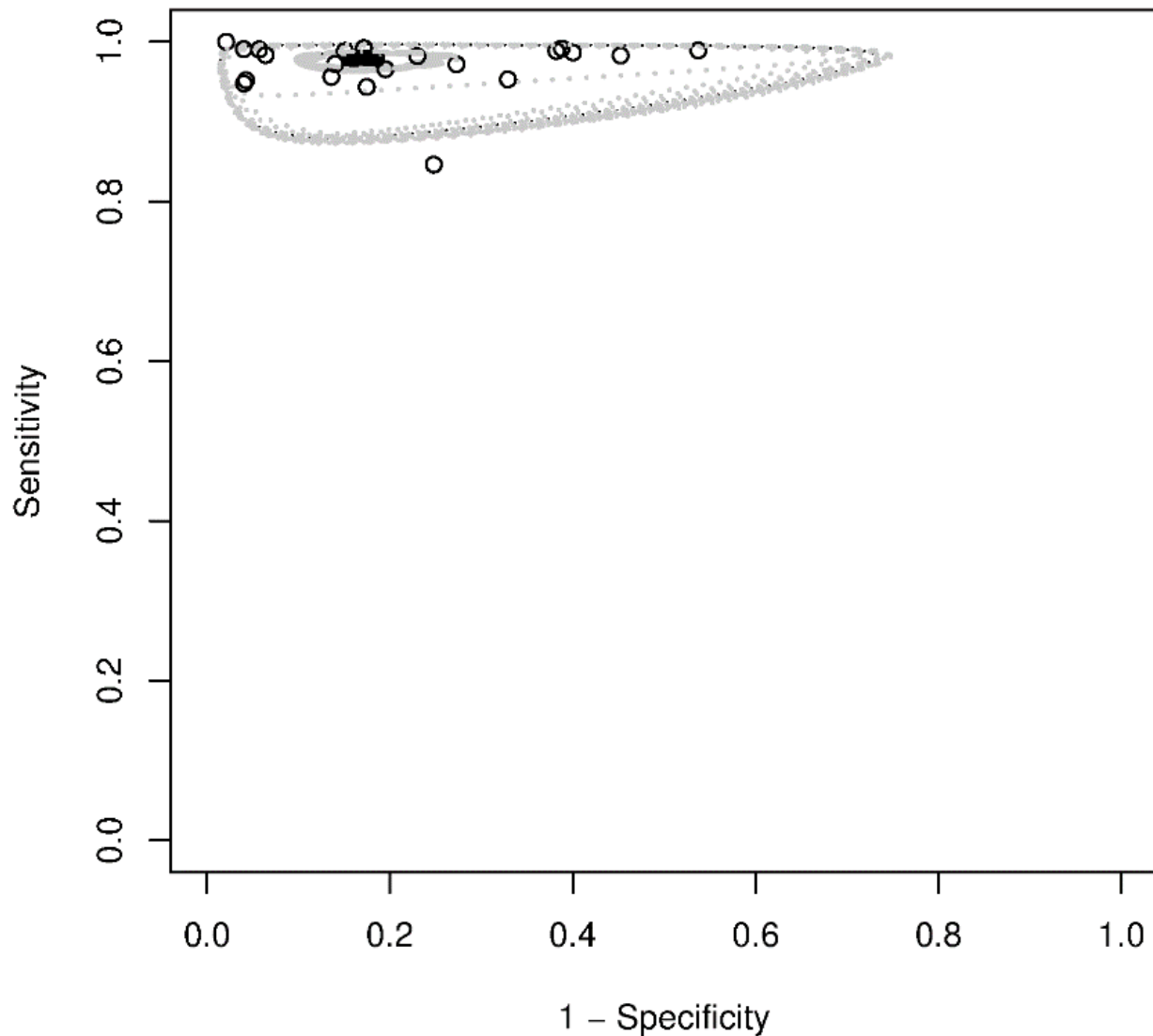
Supplementary table 3: Showing the parameters estimated by the Bivariate models for meta-analysis of the diagnostic accuracy of FNAB at Atypical, Suspicious and Malignant thresholds

Threshold	Parameter	Estimate	Lower 95% CI	Upper 95% CI
Atypical Considered positive	Logit Sens	3.813	3.416	4.210
	Logit Spec	1.601	1.153	2.049
	Sensitivity	0.978	0.968	0.985
	Specificity	0.832	0.760	0.886
	FPR	0.168	0.114	0.240
	DOR	224.576	126.193	399.660
	LR+	5.830	4.018	8.457
	LR-	0.026	0.018	0.038
	Correlation	-0.097		
	Theta	1.638		
	Lambda	5.929		
	beta	0.376		
	sigma2_theta	0.410		
	sigma2_alpha	1.348		
Suspicious Considered positive	Logit Sens	2.392	2.113	2.672
	Logit Spec	4.065	3.485	4.645
	Sensitivity	0.916	0.892	0.935
	Specificity	0.983	0.970	0.990
	FPR	0.017	0.010	0.030
	DOR	637.444	336.527	1207.434
	LR+	54.304	30.702	96.050
	LR-	0.085	0.066	0.110
	Correlation	-0.027		
	Theta	0.325		
	Lambda	6.271		
	beta	0.738		
	sigma2_theta	0.362		
	sigma2_alpha	1.372		
Malignant Considered positive	Logit Sens	1.170	0.878	1.462
	Logit Spec	6.808	5.132	8.484
	Sensitivity	0.763	0.706	0.812
	Specificity	0.999	0.994	1.000
	FPR	0.001	0.000	0.006
	DOR	2915.540	555.785	15294.340
	LR+	691.508	130.799	3655.868
	LR-	0.237	0.190	0.296
	Correlation	-0.297		
	Theta	-0.828		
	Lambda	5.882		
	beta	1.183		
	sigma2_theta	0.931		
	sigma2_alpha	2.019		

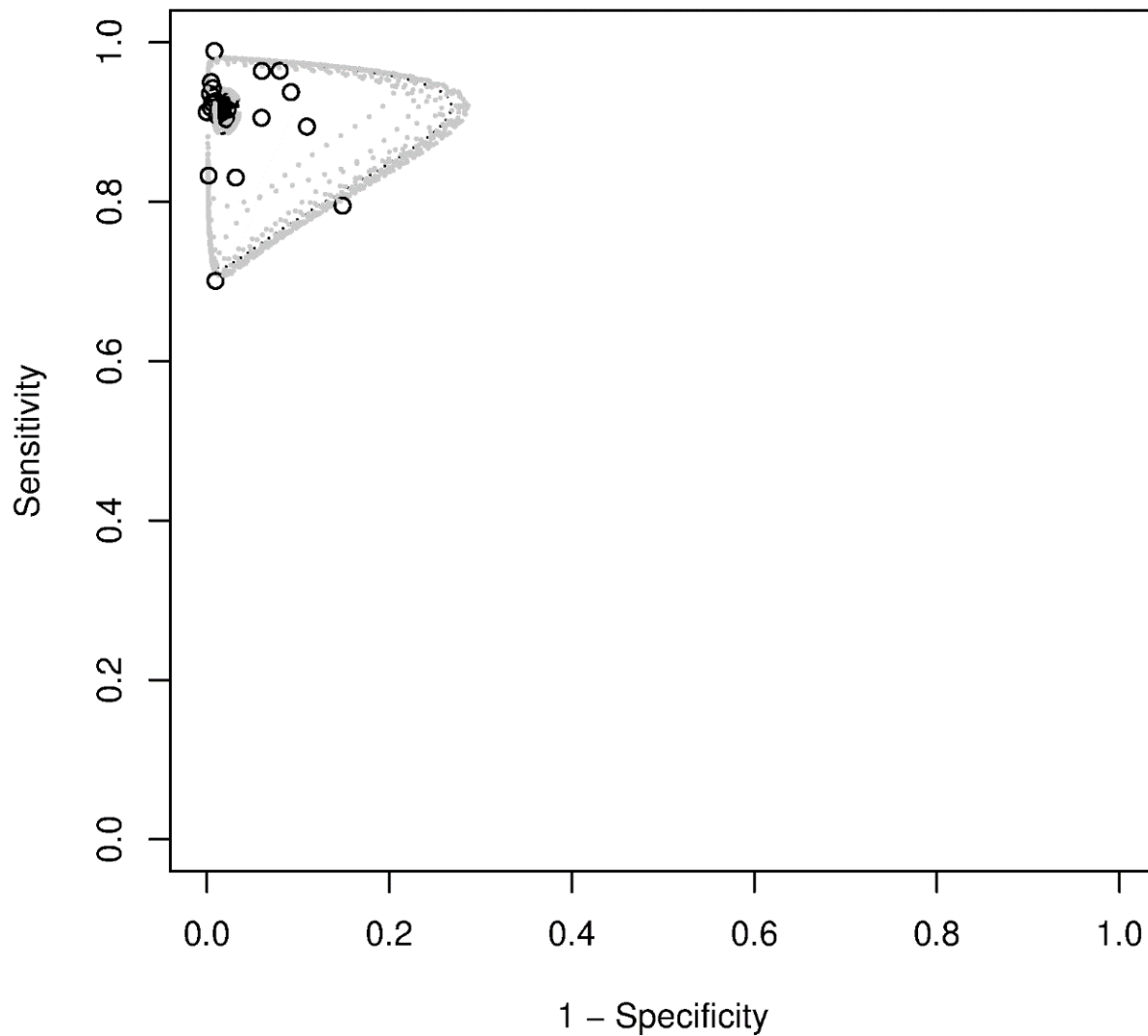
Supplementary table 4: Leave One out meta-analysis of AUC

Study left out	AUC	Lower 95% C.I.	Upper 95% C.I.	Lower 95% Predictive Interval	Upper 95% Predictive Interval
Montezuma D	0.976	0.962	0.985	0.857	0.996
Chauhan V	0.971	0.96	0.979	0.899	0.992
Agarwal A	0.975	0.96	0.984	0.849	0.996
McHugh KE	0.977	0.965	0.985	0.895	0.995
De Rosa F	0.976	0.962	0.985	0.855	0.997
Wong S	0.975	0.961	0.985	0.849	0.996
Agrawal S	0.975	0.96	0.984	0.849	0.996
Dixit N	0.975	0.96	0.984	0.854	0.996
Marabi M	0.976	0.962	0.985	0.857	0.996
Wong YP	0.974	0.96	0.984	0.856	0.996
Tejeswini V	0.976	0.961	0.985	0.851	0.996
Sundar PM	0.975	0.96	0.985	0.85	0.996
Nigam JS	0.976	0.962	0.985	0.859	0.996
Agrawal N	0.974	0.959	0.984	0.853	0.996
Sarangi S	0.974	0.96	0.984	0.857	0.996
Nargund A	0.976	0.962	0.985	0.856	0.996
Ahuja S	0.975	0.961	0.985	0.851	0.996
Kamatar PV	0.976	0.961	0.985	0.852	0.996
Apuroopa M	0.975	0.96	0.985	0.849	0.996
Niaz M	0.977	0.963	0.985	0.865	0.996
Deshpande SA	0.976	0.961	0.985	0.851	0.996
Joshee A	0.976	0.961	0.985	0.851	0.996
None	0.975	0.962	0.984	0.859	0.996

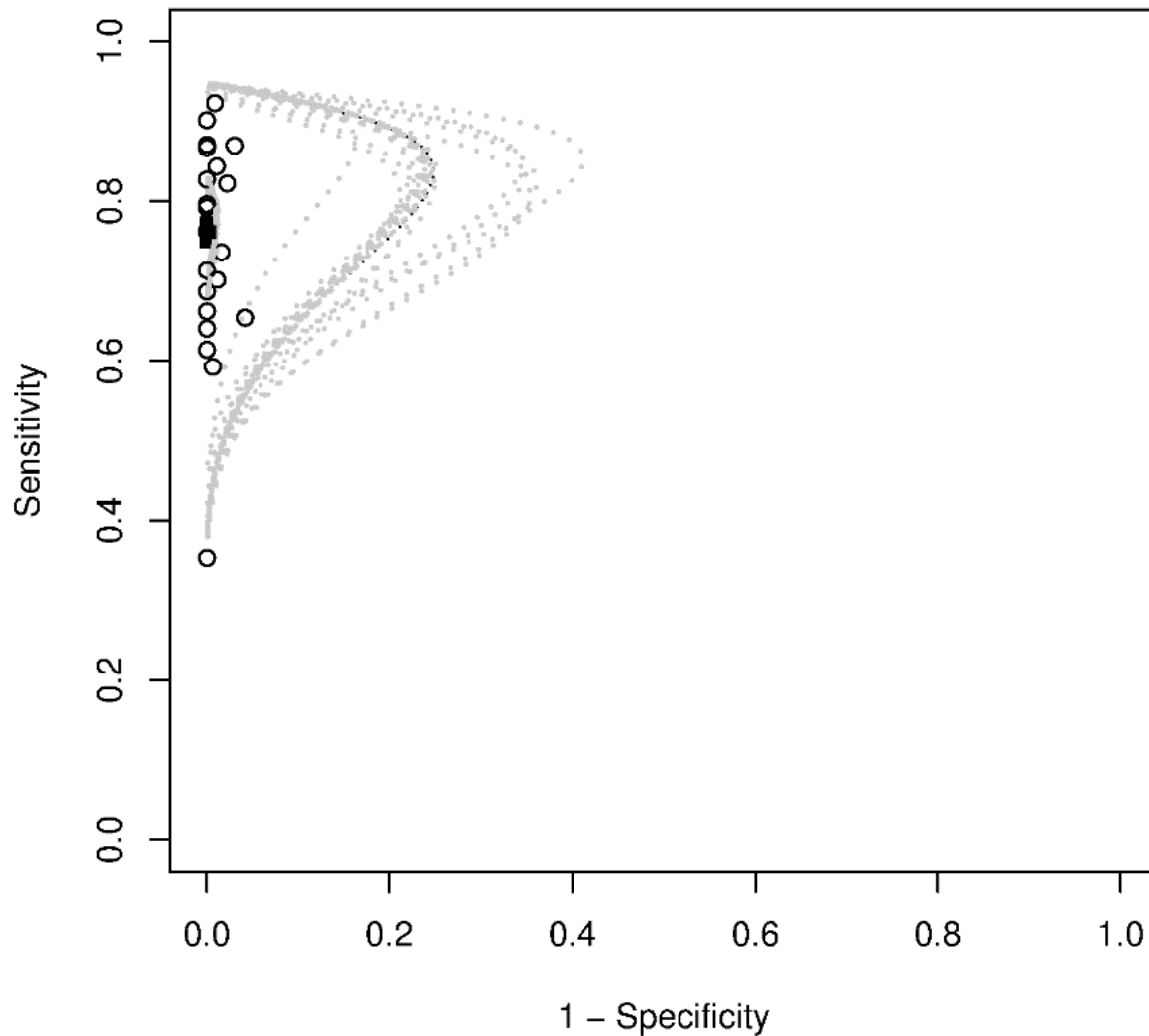
Supplementary Figure 2a: Graphical representation of the leave one out meta-analysis for the sensitivity and specificity for the analysis considering the Atypical category as positive for malignancy. All bivariate plots of 22 analyses leaving one study out have been represented on top of the other. The solid gray lines enclose the 95% Confidence Interval, the dotted lines enclose the 95% predictive intervals. All the solid gray lines from 22 leave one out analyses overlaid on the same plot shows little variation, but the 95% prediction areas do show variation especially for one study, which on further examination had a low load of FNAB performed per year.



Supplementary Figure 2b: Graphical representation of the leave one out meta-analysis for the sensitivity and specificity for the analysis considering the Suspicious category as positive for malignancy. All bivariate plots of 22 analyses leaving one study out have been represented on top of the other. The solid gray lines enclose the 95% Confidence Interval, the dotted lines enclose the 95% predictive intervals. All the solid gray lines from 22 leave one out analyses overlaid on the same plot shows little variation, but the 95% prediction areas do show variation especially for two studies, both of which had low load of FNAB performed per year.



Supplementary Figure 2c: Graphical representation of the leave one out meta-analysis for the sensitivity and specificity for the analysis considering the Suspicious category as positive for malignancy. All bivariate plots of 22 analyses leaving one study out have been represented one on top of the other. The solid gray lines enclose the 95% Confidence Interval, the dotted lines enclose the 95% predictive intervals. All the solid gray lines from 22 leave one out analyses overlaid on the same plot shows little variation, but the 95% prediction areas do show variation.

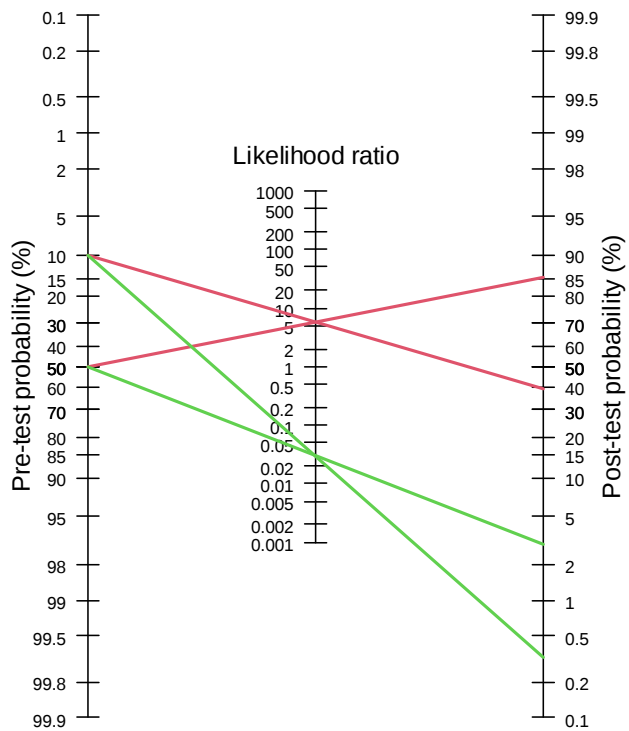


Supplementary table 5: Sensitivity analysis of the meta-analysis for the different thresholds excluding studies whose likelihood ratios differ markedly from the meta-analytic estimate of the likelihood ratios (as shown in Fig 5 of the main manuscript)

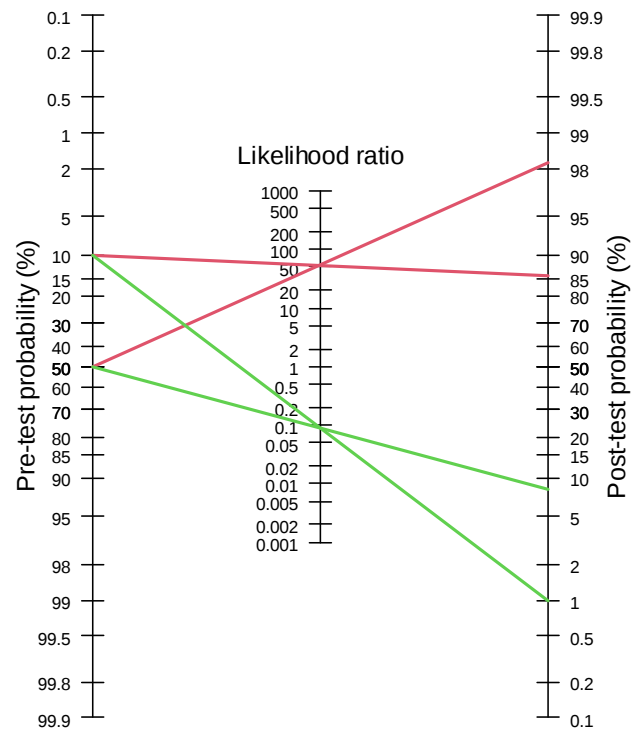
Threshold	Studies excluded	Sensitivity	95% CI	Specificity	95% CI
Atypical considered positive	McHugh KE, 2019	0.979	[0.972, 0.985]	0.836	[0.761, 0.890]
	Wong YP, 2021 & Sarangi S, 2021	0.977	[0.966, 0.984]	0.811	[0.734, 0.870]
Suspicious considered positive	Chauhan V, 2019	0.912	[0.888, 0.931]	0.982	[0.969, 0.990]
	McHugh KE, 2019	0.919	[0.896, 0.937]	0.985	[0.974, 0.991]
Malignant considered positive	McHugh KE, 2019	0.767	[0.709, 0.816]	0.999	[0.995, 1.000]

Supplementary figure 3: Showing Fagan nomograms with estimation of the post-test probability when testing at a pre-test probability of 10% and 50% for the three different diagnostic cut-offs

Fagan - Nomogram for cut-off Atypical considered positive



Fagan - Nomogram for cut-off Suspicious considered positive



Fagan - Nomogram for cut-off Malignant considered positive

