Variable radiologic lung nodule evaluation leads to divergent management recommendations

A. Nair, E.C. Bartlett, S.L.F. Walsh, A.U. Wells, N. Navani, G. Hardavella, S. Bhalla, L. Calandriello, A. Devaraj, J.M. Goo, J.S. Klein, H. MacMahon, C. Schaefer-Prokop, J.B. Seo, N. Sverzellati, S.R. Desai

European Respiratory Journal 2018; **DOI:** 10.1183/13993003.01359-2018

Abstract

Radiologic evaluation of incidentally-detected lung nodules on computed tomography (CT) influences management. We assessed international radiologic variation in (1) pulmonary nodule characterisation; (2) hypothetical guideline-derived management, and (3) radiologists' management recommendations.

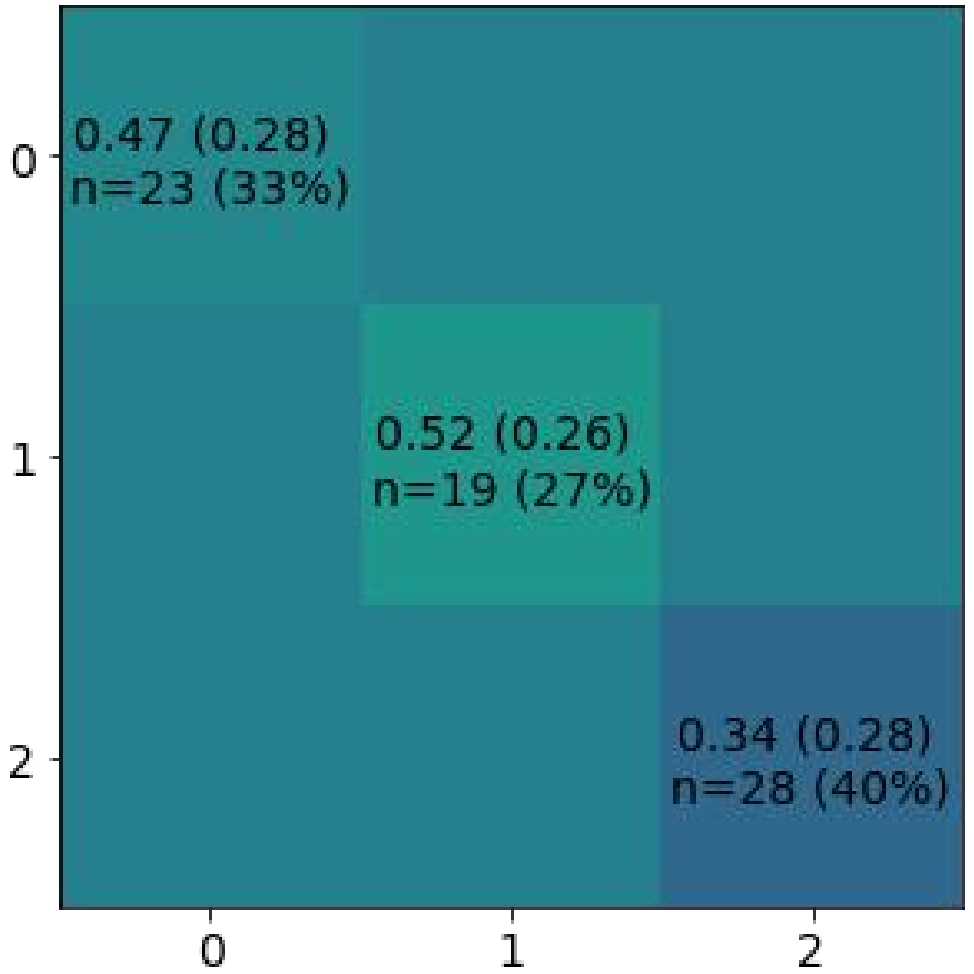
107 radiologists from 25 countries evaluated 69 CT-detected nodules, recording: i) first-choice composition (solid, part-solid [PSN] or ground-glass [GGN], with percentage confidence), ii) morphologic features, iii) dimensions, iv) recommended management and (v) decision-influencing

Compare +100 reviewers and CNN characterization

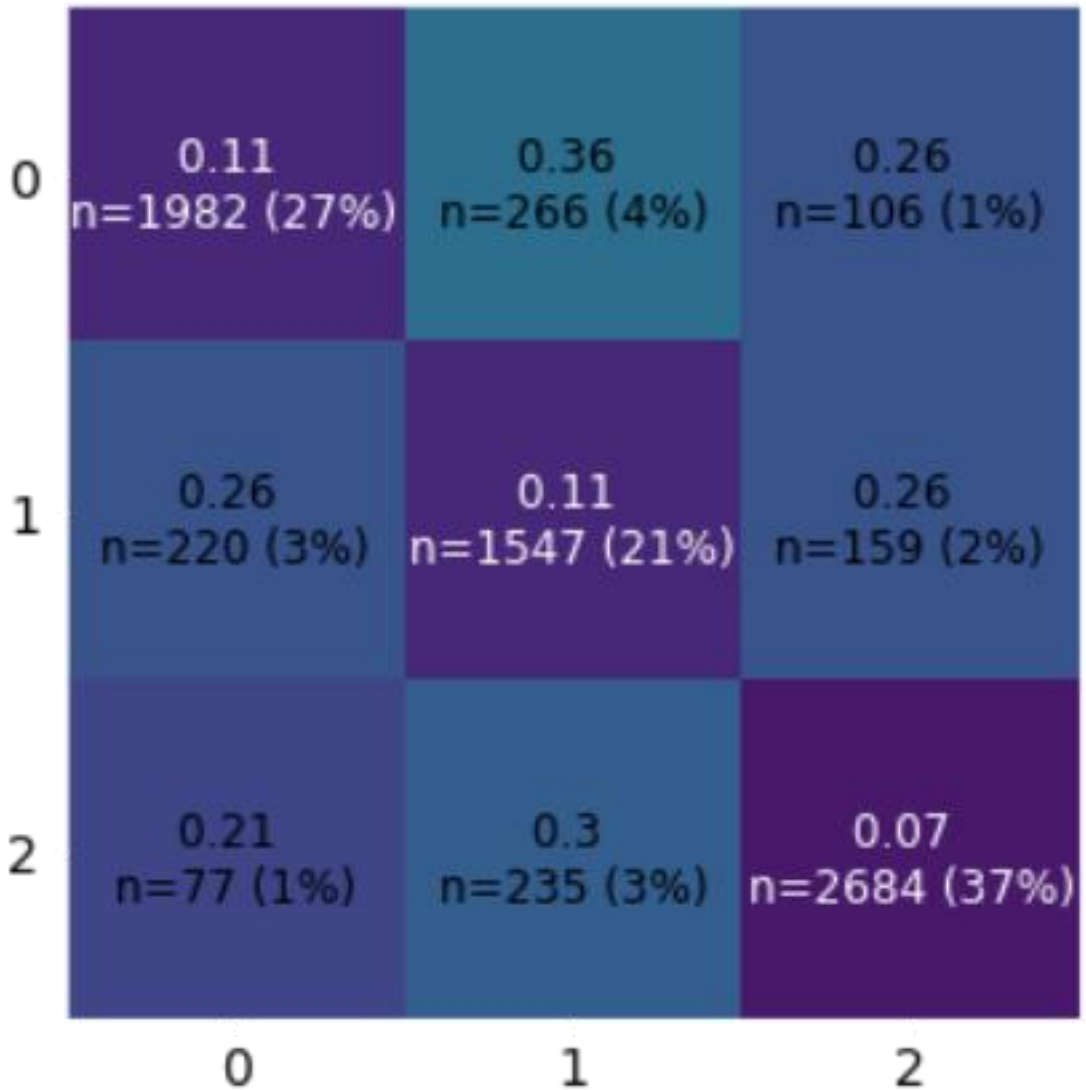
	GGN	PSN	Solid	entropy	majority
0	87	12	8	0.61	GGN
1	83	16	8	0.68	GGN
2	1	5	101	0.24	Solid
3	0	9	98	0.29	Solid
4	33	71	3	0.74	PSN
5	10	23	74	0.81	Solid
6	16	25	66	0.92	Solid

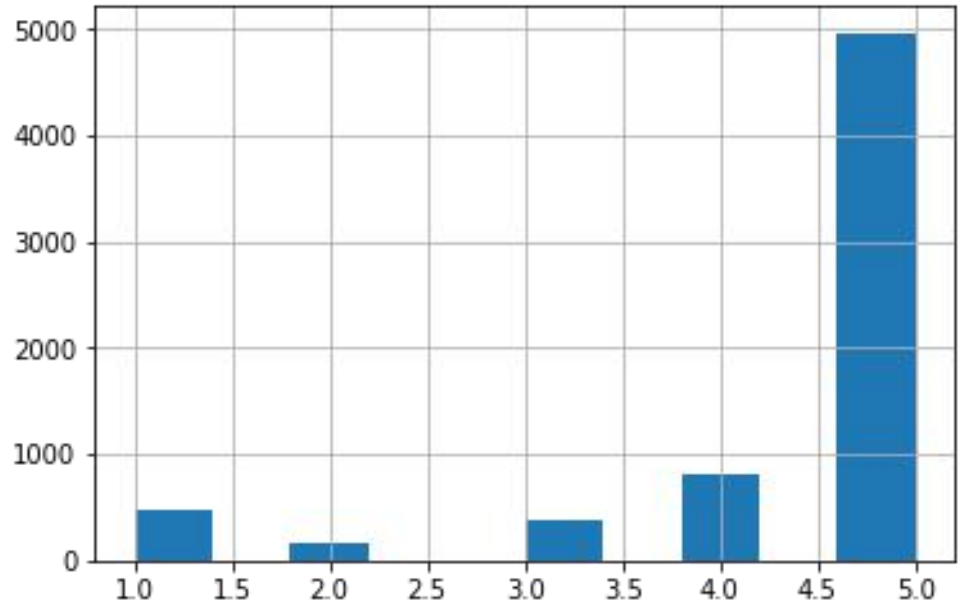
	GGN	PSN	Solid	entropy
14	33	14	60	0.95
6	16	25	66	0.92
29	65	12	30	0.90
17	13	67	27	0.90
68	25	71	11	0.85
5	10	23	74	0.81
25	72	7	28	0.80
23	0	107	0	0.00
65	107	0	0	0.00
39	107	0	0	0.00
62	0	0	107	0.00
10	107	0	0	0.00
20	0	107	0	0.00
37	0	1	106	0.05

	GGN	PSN	Solid	entropy	majority
0	87	12	8	0.61	GGN
1	83	16	8	0.68	GGN
2	1	5	101	0.24	Solid
3	0	9	98	0.29	Solid
4	33	71	3	0.74	PSN
5	10	23	74	0.81	Solid
6	16	25	66	0.92	Solid

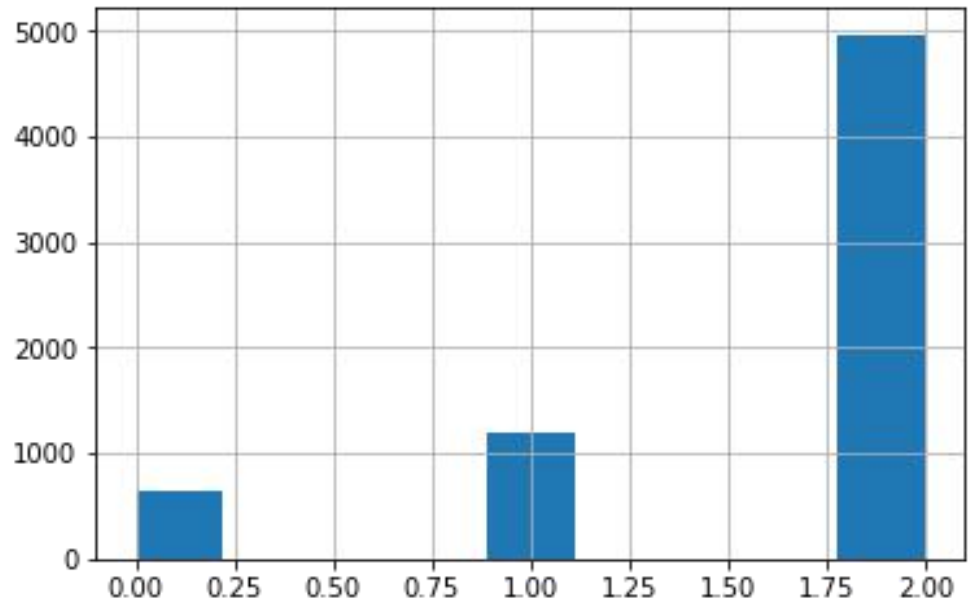


case_id	user_id	dx1	dx2	dx3	con1	con2	con3	Consensus	entropy
1.0	11.0	Pure Ground Glass	0	0	100.0	0.0	0.0	Pure Ground Glass	0.000000
2.0	11.0	Pure Ground Glass	0	0	100.0	0.0	0.0	Pure Ground Glass	0.000000
3.0	11.0	Pure Solid	0	0	100.0	0.0	0.0	Pure Solid	0.000000
4.0	11.0	Pure Solid	0	0	100.0	0.0	0.0	Pure Solid	0.000000
5.0	11.0	Part Solid	Pure Ground Glass	0	70.0	30.0	0.0	Part Solid	0.610864
6.0	11.0	Pure Solid	Part Solid	0	90.0	10.0	0.0	Pure Solid	0.325083
7.0	11.0	Pure Solid	0	0	100.0	0.0	0.0	Pure Solid	0.000000
8.0	11.0	Part Solid	Pure Solid	0	80.0	20.0	0.0	Part Solid	0.500402
9.0	11.0	Pure Solid	Part Solid	0	80.0	20.0	0.0	Pure Solid	0.500402
10.0	11.0	Part Solid	Pure Ground Glass	0	60.0	40.0	0.0	Pure Ground Glass	0.673012

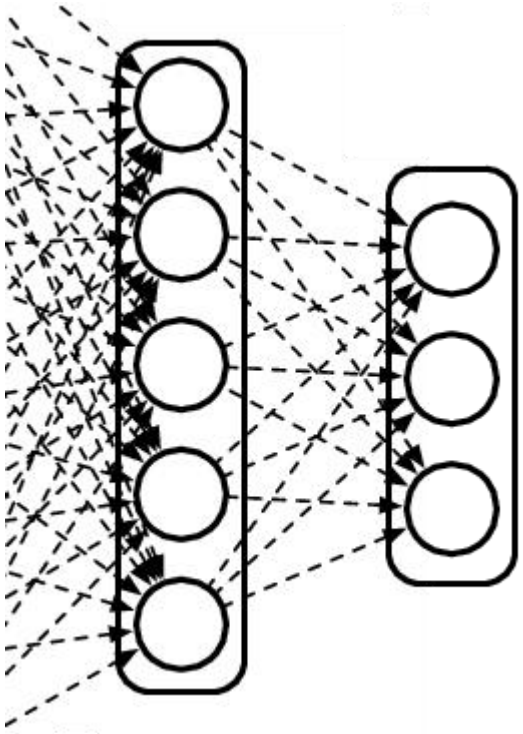

0.11
n=1982 (27%)
0.36
n=266 (4%)
0.26
n=106 (1%)
0.26
n=220 (3%)
0.11
n=1547 (21%)
0.26
n=159 (2%)
0.21
n=77 (1%)
0.3
n=235 (3%)
0.07
n=2684 (37%)
0
1
2
0
1
2

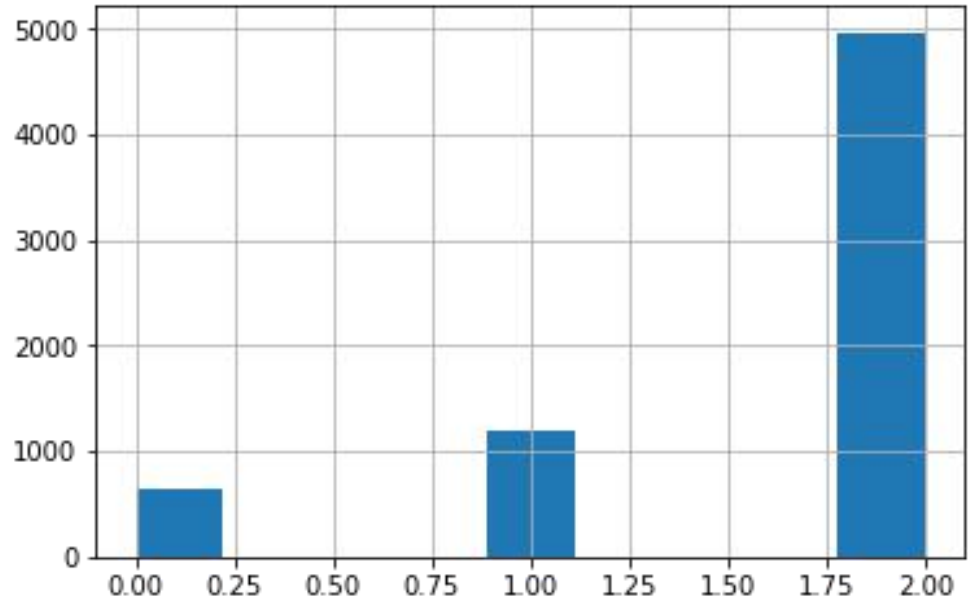


	scores	spiculation	subtlety	internal_structure	calcification	sphericity	margin	lobulation	texture	malignancy
0	1	No	Extreme	Soft tissue	Popcorn	Linear	Poorly defined	No	Non-Solid/GGO	Highly unlikely
1	2	Near no	Moderate	Fluid	Laminated	Ovoid/Linear	Near poorly defined	Nearly no	Non-Solid/Mixed	Moderately unlikely
2	3	Medium	Fairly subtle	Fat	Solid	Ovoid	Medium	Medium	Part Solid/Mixed	Indeterminate
3	4	Near marked	Moderately obvious	Air	Non-central	Ovoid/Round	Near sharp	Near marked	Solid/Mixed	Moderately
4	5	Marked	Obvious	-	Central	Round	Sharp	Marked	Solid	-
5	6	-	-	-	Absent	-	-	-	-	-

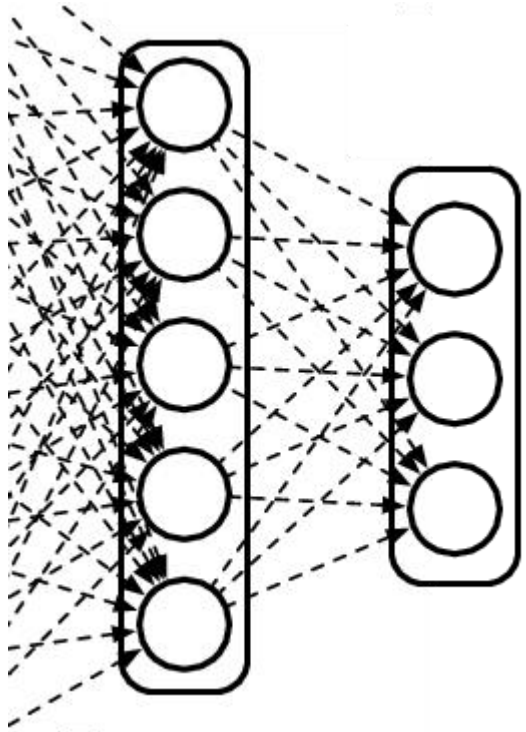


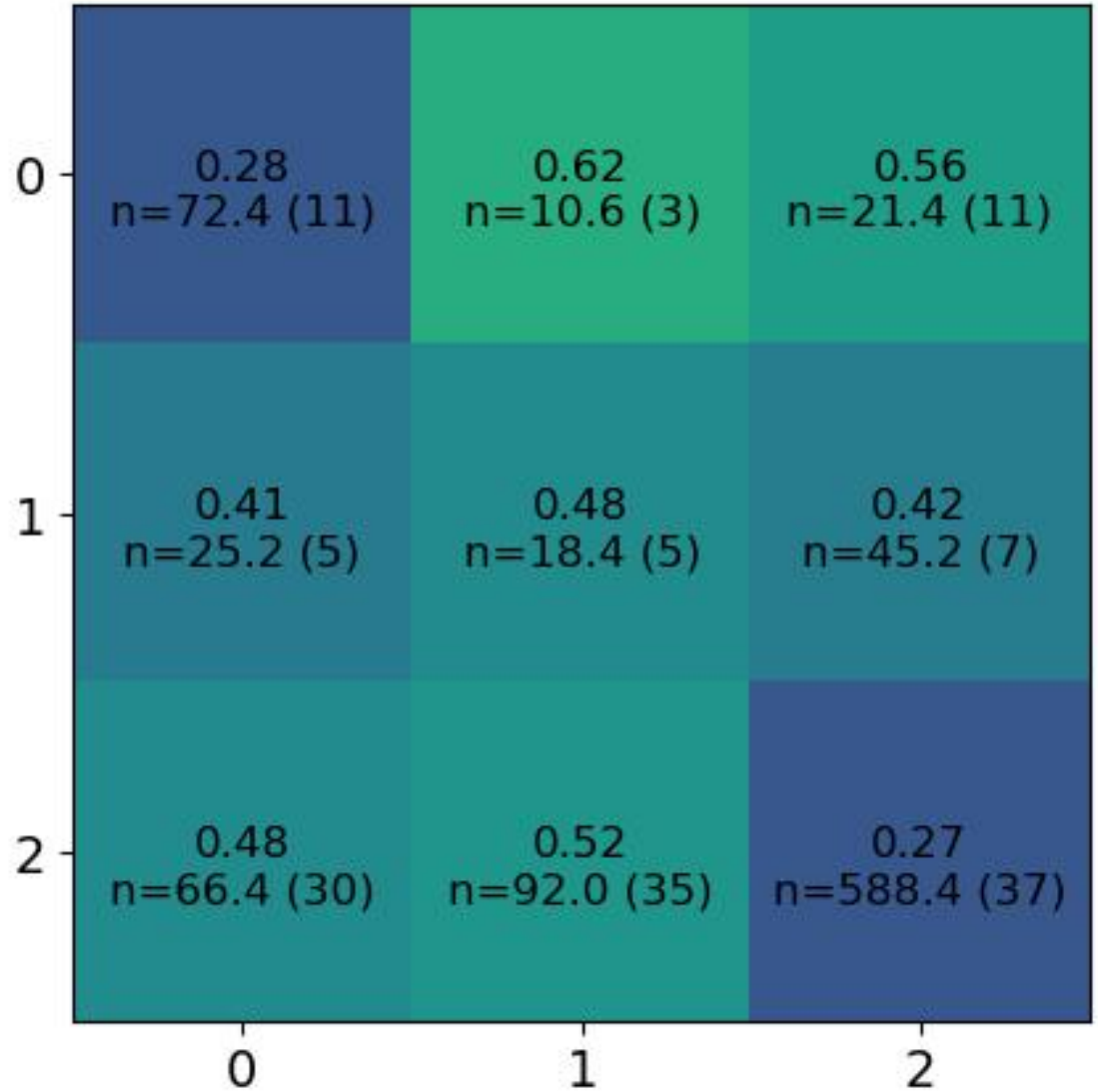
	scores	spiculation	subtlety	internal_structure	calcification	sphericity	margin	lobulation	texture	malignancy
0	1	No	Extreme	Soft tissue	Popcorn	Linear	Poorly defined	No	Non-Solid/GGO	Highly unlikely
1	2	Near no	Moderate	Fluid	Laminated	Ovoid/Linear	Near poorly defined	Nearly no	Non-Solid/Mixed	Moderately unlikely
2	3	Medium	Fairly subtle	Fat	Solid	Ovoid	Medium	Medium	Part Solid/Mixed	Indeterminate
3	4	Near marked	Moderately obvious	Air	Non-central	Ovoid/Round	Near sharp	Near marked	Solid/Mixed	Moderately
4	5	Marked	Obvious	-	Central	Round	Sharp	Marked	Solid	-
5	6	-	-	-	Absent	-	-	-	-	-

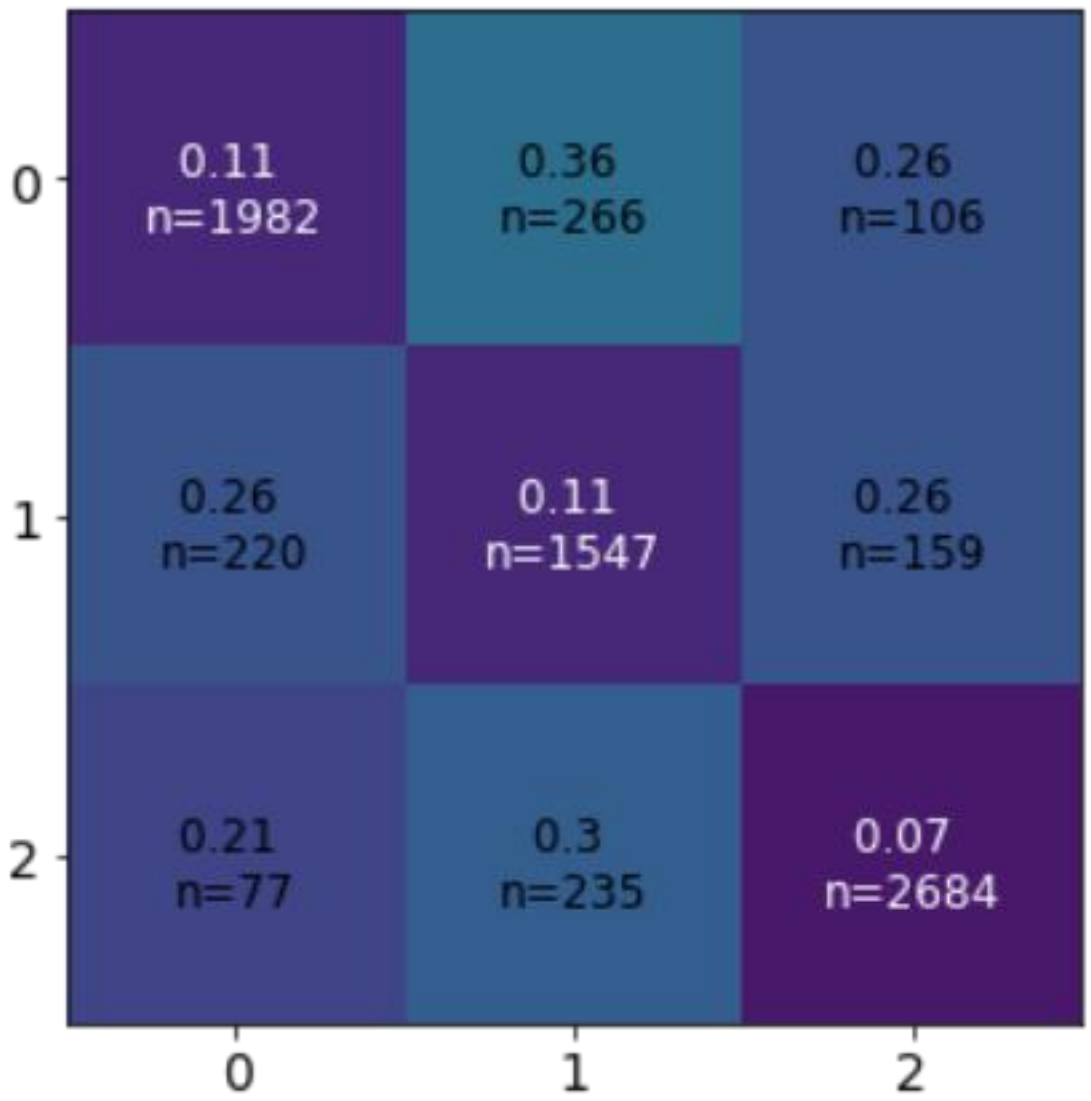


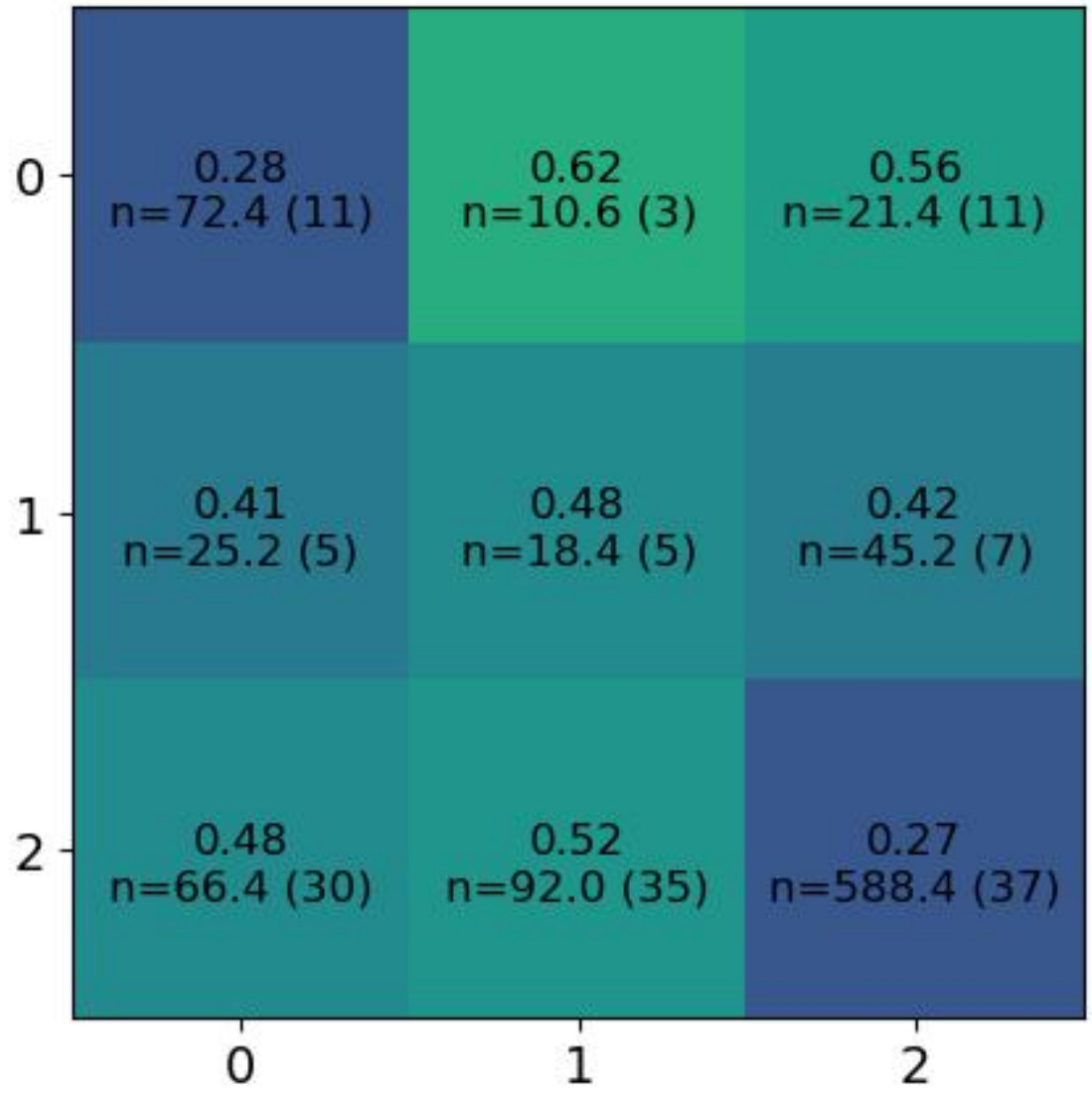


	scores	spiculation	subtlety	internal_structure	calcification	sphericity	margin	lobulation	texture	malignancy
0	1	No	Extreme	Soft tissue	Popcorn	Linear	Poorly defined	No	Non-Solid/GGO	Highly unlikely
1	2	Near no	Moderate	Fluid	Laminated	Ovoid/Linear	Near poorly defined	Nearly no	Non-Solid/Mixed	Moderately unlikely
2	3	Medium	Fairly subtle	Fat	Solid	Ovoid	Medium	Medium	Part Solid/Mixed	Indeterminate
3	4	Near marked	Moderately obvious	Air	Non-central	Ovoid/Round	Near sharp	Near marked	Solid/Mixed	Moderately
4	5	Marked	Obvious	-	Central	Round	Sharp	Marked	Solid	-
5	6	-	-	-	Absent	-	-	-	-	-




0
1
2
0.28
n=72.4 (11)
0.62
n=10.6 (3)
0.56
n=21.4 (11)
0.41
n=25.2 (5)
0.48
n=18.4 (5)
0.42
n=45.2 (7)
0.48
n=66.4 (30)
0.52
n=92.0 (35)
0.27
n=588.4 (37)
0
1
2

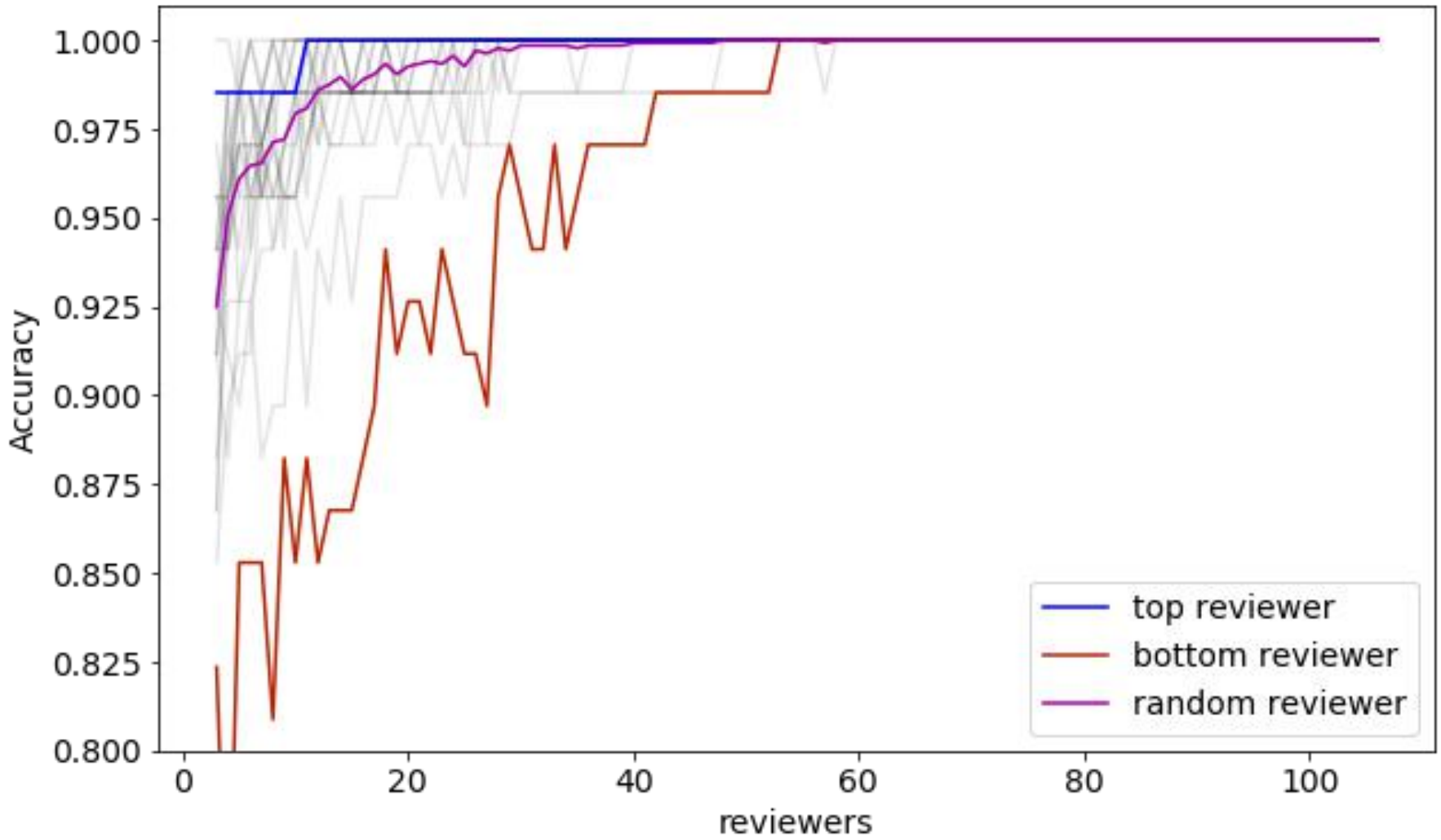

0.11
n=1982
0.36
n=266
0.26
n=106
0.26
n=220
0.11
n=1547
0.26
n=159
0.21
n=77
0.3
n=235
0.07
n=2684
0
1
2
0
1
2


0.28
n=72.4 (11)
0.62
n=10.6 (3)
0.56
n=21.4 (11)
0.41
n=25.2 (5)
0.48
n=18.4 (5)
0.42
n=45.2 (7)
0.48
n=66.4 (30)
0.52
n=92.0 (35)
0.27
n=588.4 (37)
0
1
2
0
1
2

Number of reviewers required for ground truth

user_id	user_acc
175.0	0.985294
197.0	0.955882
125.0	0.955882
113.0	0.941176
118.0	0.941176
...	
33.0	0.735294
43.0	0.735294
92.0	0.720588
47.0	0.691176
114.0	0.661765

	Consensus	reviewer_rank_0	reviewer_rank_1	reviewer_rank_2
0	Pure Ground Glass	Pure Ground Glass	Pure Ground Glass	Pure Ground Glass
1	Pure Ground Glass	Pure Ground Glass	Pure Ground Glass	Pure Ground Glass
2	Pure Solid	Pure Solid	Pure Solid	Pure Solid
3	Pure Solid	Pure Solid	Pure Solid	Pure Solid
4	Part Solid	Part Solid	Part Solid	Pure Ground Glass
5	Pure Solid	Pure Solid	Pure Solid	Pure Solid
6	Pure Solid	Pure Solid	Pure Ground Glass	Pure Solid


1.000
0.975
0.950
0.925
0.900
0.875
0.850
0.825
0.800
Accuracy
0
20
40
60
80
100
reviewers
top reviewer
bottom reviewer
random reviewer

Instructions

– After speaking with Arjun we did not use scan 23 (we were already not using scan 30)
– We try training a CNN based on LIDC and malignancy label (5 classes). From that CNN we get features for the 68 nodules

1) Get the 68 nodules and their dataframe with **Documents/KCL/18Aug29–107_observers/localize nodules on 70 dicoms v5.ipynb**
nodules_five_scans_68_cases_RGB_no23_no29.npy
Nodule Study raw data for KCL 68.csv

2) Train a CNN based on LIDC malignancy with: **/Documents/KCL/18 Oct 9 – ladder on LIDC plus 70 nodules/**
CNN classifier based on LIDC texture v6 – images from nodule consensus classifier again.ipynb:
In part 4 (use the pretrained network on the 70 nodules) load nodules_five_scans_68_cases_RGB_no23_no29.npy to produce
results/results CNN regression based on LIDC texture v6/features128_classification5malignancy_on68ndls.npy

3) Once we have the features for the 68 nodules we run the **Documents/KCL/18 Oct 9 – ladder on LIDC plus 70 nodules/**
regression on 70 nodules v0 – random forests where we load:
nodules_five_scans_68_cases_RGB_no23_no29.npy
Nodule Study raw data for KCL 68.csv

To run the entropy comparison:
– To get model uncertainty **CNN classifier based on LIDC texture v6 – images from nodule consensus classifier again.ipynb**, specially the **Get uncertainty in the prediction** section
– To get uncertainty from the reviewers run **Documents/KCL/18Aug29–107_observers/reviewers uncertainty in 70 nodules v1.ipynb**