Appendix: Extra information on the LCP-CNN

This section gives a brief overview of the nature and training of the Lung Cancer Prediction Convolutional Neural Network (LCP-CNN). The main paper looks at validating this model on data gathered from clinical UK practice under the IDEAL project.

The use of Convolutional Neural Networks for medical imaging recognition is motivated by their success in computer vision tasks (A), which take 2D photographic images as an input, and classify these into different categories based on what objects are visible. In particular, the DenseNet CNN (B) is widely used for such categorisation tasks, though for the LCP-CNN, the network first had to be scaled up to handle three-dimensional datasets, because while computer vision images are generally three-channel (red, green, blue), two-dimensional grids of pixels, CT scans are typically one-channel (Hounsfield Units) three-dimensional arrays of voxels.

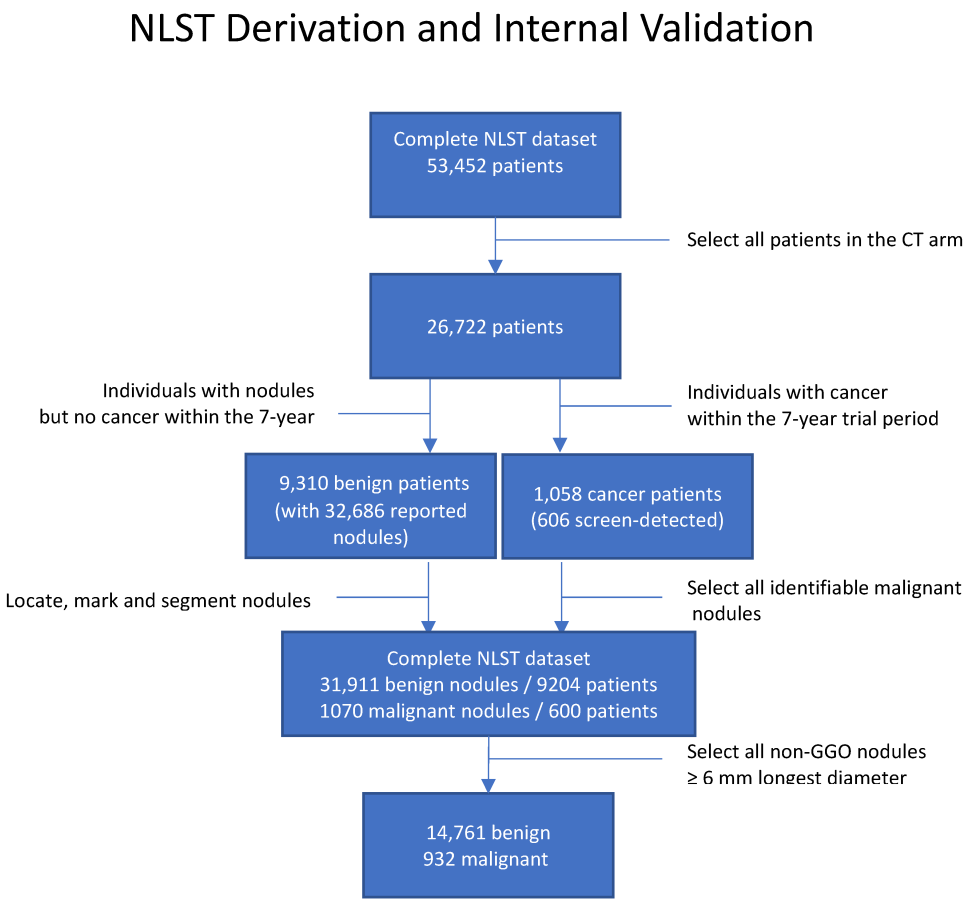
In the first phase of training, the LCP-CNN was primed using hundreds of thousands of images selected and curated carefully to teach the network the kinds of visual discrimination tasks that may form the building blocks for a nodule discrimination task. The second phase comprised full supervised training on a version of the NLST data. This was performed on CT images of all solid and semi-solid nodules of at least 6mm in diameter from the NLST dataset that were not reported as pure ground glass opacities (GGO); all GGOs were excluded because there were too few examples of malignant GGOs to train the system reliably. Working under the supervision of expert thoracic radiologists from OUH, a team of doctors curated the data for machine learning use. Firstly, all patients randomized to the CT screening arm were selected with those who did not develop lung cancer within the 7-year trial window being assigned to the benign group. From the benign group, all studies with reported nodules were examined, and all screen-reported nodules that could be identified on these CTs were marked up. All nodules that could be unambiguously identified as the diagnosed malignancies reported in NLST metadata were marked up using an extensive multi-pass review process, whether or not these corresponded to nodules reported at screening time. The final dataset used to train the CNN contained a total of 14,761 benign nodules from 5,972 patients, and 932 malignant nodules from 575 patients. Note that each patient had three annual images and hence a nodule could be present on up to three CTs.

The derivation and internal validation of the LCP-CNN was performed using eight-fold cross-validation, where all images and nodules for a given patient were assigned to the same fold, and in each of the eight training operations, one eighth of the data were reserved as an auxiliary set for convergence testing and parameter/threshold setting, and one final eighth was kept for internal validation. Because the NLST contains many more benign nodules than cancers, class balancing was used during training, since otherwise the resulting network would be tuned always to predict "benign", since that would be the dominant class. The final result of applying the LCP-CNN to new unseen data (such as IDEAL data) is an regressor that gives a score between 0 and 100 for malignancy.

[A] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44. PubMed PMID

[B] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ, editors. Densely connected convolutional networks. CVPR; 2017.

Below: Overview of data selection and curation process for the NLST dataset used in the LCP-CNN derivation.



Site-wise nodule demographics

Leeds Subset				
Patient sex, counted by nodule:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Male	34 (38.2)	178 (51.0)	212 (48.4)
	Female	55 (61.8)	171 (49.0)	226 (51.6)
Nodule size (clinician-stated):		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	5mm	2 (2.2)	69 (19.8)	71 (16.2)
	>5mm to <=7mm	5 (5.6)	108 (30.9)	113 (25.8)
	>7mm to <=10mm	24 (27.0)	90 (25.8)	114 (26.0)
	>10mm to <=15mm	58 (65.2)	82 (23.5)	140 (32.0)
Patient age, by nodule:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	0 to 49	2 (2.2)	37 (10.6)	39 (8.9)
	50 to 59	4 (4.5)	62 (17.8)	66 (15.1)
	60 to 69	30 (33.7)	128 (36.7)	158 (36.1)
	70 to 79	40 (44.9)	92 (26.4)	132 (30.1)
	80 to 89	12 (13.5)	29 (8.3)	41 (9.4)
	90 to 99	1 (1.1)	1 (0.3)	2 (0.5)
Nodule contrast (auto-detected):		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	0 to <=80HU	58 (65.2)	258 (73.9)	316 (72.1)
	80 to <=300HU	31 (34.8)	91 (26.1)	122 (27.9)
Nodule locations:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Right Upper Lobe	34 (38.2)	98 (28.1)	132 (30.1)
	Right Middle Lobe	6 (6.7)	35 (10.0)	41 (9.4)
	Right Lower Lobe	17 (19.1)	78 (22.3)	95 (21.7)
	Left Upper Lobe	13 (14.6)	67 (19.2)	80 (18.3)
	Lingula lobe	4 (4.5)	12 (3.4)	16 (3.7)
	Left Lower Lobe	15 (16.9)	59 (16.9)	74 (16.9)
Nodule spiculation		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Non-spiculated	29 (32.6)	258 (73.9)	287 (65.5)
	Spiculated	60 (67.4)	91 (26.1)	151 (34.5)

Nottingham Subset				
Patient sex, counted by nodule:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Male	35 (37.6)	187 (56.7)	222 (52.5)
	Female	58 (62.4)	143 (43.3)	201 (47.5)
Nodule size (clinician-stated):		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	5mm	5 (5.4)	68 (20.6)	73 (17.3)
	>5mm to <=7mm	12 (12.9)	153 (46.4)	165 (39.0)
	>7mm to <=10mm	24 (25.8)	80 (24.2)	104 (24.6)
	>10mm to <=15mm	52 (55.9)	29 (8.8)	81 (19.1)
Patient age, by nodule:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	0 to 49	3 (3.2)	44 (13.3)	47 (11.1)
	50 to 59	18 (19.4)	64 (19.4)	82 (19.4)
	60 to 69	34 (36.6)	98 (29.7)	132 (31.2)
	70 to 79	32 (34.4)	92 (27.9)	124 (29.3)
	80 to 89	6 (6.5)	31 (9.4)	37 (8.7)
	90 to 99	0 (0.0)	1 (0.3)	1 (0.2)
Nodule contrast (auto-detected):		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	0 to <=80HU	30 (32.3)	104 (31.5)	134 (31.7)
	80 to <=300HU	63 (67.7)	226 (68.5)	289 (68.3)
Nodule locations:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Right Upper Lobe	30 (32.3)	65 (19.7)	95 (22.5)
	Right Middle Lobe	4 (4.3)	45 (13.6)	49 (11.6)
	Right Lower Lobe	10 (10.8)	102 (30.9)	112 (26.5)
	Left Upper Lobe	28 (30.1)	39 (11.8)	67 (15.8)
	Lingula lobe	1 (1.1)	5 (1.5)	6 (1.4)
	Left Lower Lobe	20 (21.5)	74 (22.4)	94 (22.2)
Nodule spiculation		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Non-spiculated	43 (46.2)	299 (90.6)	342 (80.9)
	Spiculated	50 (53.8)	31 (9.4)	81 (19.1)

Oxford Subset				
Patient sex, counted by nodule:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Male	23 (44.2)	259 (53.5)	282 (52.6)
	Female	29 (55.8)	225 (46.5)	254 (47.4)
Nodule size (clinician-stated):		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	5mm	1 (1.9)	131 (27.1)	132 (24.6)
	>5mm to <=7mm	0 (0.0)	177 (36.6)	177 (33.0)
	>7mm to <=10mm	11 (21.2)	124 (25.6)	135 (25.2)
	>10mm to <=15mm	40 (76.9)	52 (10.7)	92 (17.2)
Patient age, by nodule:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	0 to 49	2 (3.8)	70 (14.5)	72 (13.4)
	50 to 59	5 (9.6)	75 (15.5)	80 (14.9)
	60 to 69	21 (40.4)	133 (27.5)	154 (28.7)
	70 to 79	19 (36.5)	136 (28.1)	155 (28.9)
	80 to 89	5 (9.6)	66 (13.6)	71 (13.2)
	90 to 99	0 (0.0)	3 (0.6)	3 (0.6)
Nodule contrast (auto-detected):		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	0 to <=80HU	16 (30.8)	219 (45.2)	235 (43.8)
	80 to <=300HU	36 (69.2)	265 (54.8)	301 (56.2)
Nodule locations:		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Right Upper Lobe	19 (36.5)	72 (14.9)	91 (17.0)
	Right Middle Lobe	2 (3.8)	105 (21.7)	107 (20.0)
	Right Lower Lobe	9 (17.3)	129 (26.7)	138 (25.7)
	Left Upper Lobe	10 (19.2)	53 (11.0)	63 (11.8)
	Lingula lobe	1 (1.9)	20 (4.1)	21 (3.9)
	Left Lower Lobe	11 (21.2)	105 (21.7)	116 (21.6)
Nodule spiculation		**Number Cancer (%)**	**Number Benign (%)**	**Total (%)**
	Non-spiculated	24 (46.2)	436 (90.1)	460 (85.8)
	Spiculated	28 (53.8)	48 (9.9)	76 (14.2)