Supplemental Table: **Properties of the Training and Validation Data used for Development of the AI Lung and Airspace Disease Segmentation on CT and CXR, compared to our study cohort.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | AI System Training CT | AI system Training CXR | AI System Validation CT | Study Cohort CT |
| Datasets | Total: 1929, COVID-19: 1005, ILD: 267, Pneumonia: 147, Normal: 510, | Total: 727COVID-19: 0Control: 727 | Total: 182,COVID-19: 86,Control: 96 | Total: 86, COVID-19: 86 |
| Data Origin | Multiple sites including USA, Spain, Switzerland, Germany, France, Denmark, Canada and Belarus,  | Multiple sites including USA and Germany,  | Multiple sites including USA, Spain and Czech Republic,  | USA |
| Sex | Female: 628,Male: 827, Unknown: 474 | Female: 142Male: 123Unknown: 461 | Female: 66,Male: 101,Unknown: 15 | Female: 42,Male: 44,Unknown: 0 |
| Age (years) | Median: 61IQR: 56-66Unknown: 938 | Median: 58IQR: 31Unknown: 461 | Median: 62IQR: 53-72Unknown: 54 | Median: 59 IQR: 25-93Unknown: 0 |
| Scanner Manufacturer | GE: 450,Siemens: 1258,Philips: 41,Toshiba: 23, Other/Unknown: 156 | 'FUJIFILM': 111, 'Carestream': 130, 'Agfa': 25,Other/Unknown: 461 | GE: 60,Siemens: 58,Philips: 24,Toshiba: 27, Other/Unknown: 13 | GE: 29,Siemens: 57,Philips: 0,Toshiba: 0 Other/Unknown: 0 |
| Slice Thickness[mm] | ≤ 1.5: 1632(1.5, 3.0]: 282,>3.0: 12 | N/A | ≤ 1.5: 51,(1.5, 3.0]: 116,>3.0: 15  | ≤ 1.5: 82,(1.5, 3.0]: 4,>3.0: 0 Unknown: 0 |
| Reconstruction Kernel | Soft: 691Hard: 1035 Unknown 203 | N/A | Soft: 86 Hard: 71 Unknown: 25 | Soft: 32Hard: 54 Unknown: 0 |

We used the following selection criteria to build the the AI System Validation CT dataset:

For early-stopping model selection and validating the performance of the proposed system without being exposed to the test dataset, we constructed a validation DRR dataset generated from 86 CT scans from COVID-19 patients and 96 control patients. We initially reserved 100 COVID-19 positive patients and 100 control patients as in a prior work focusing on the COVID-19 quantification on CT [19]. We then manually excluded the CT scans that have lower than standard dose, incomplete lungs or a whole-body acquisition. Those excluded images were not suitable for CXR simulation.