| Title  | Source                | Journal/ MICCAI<br>volume             | Publication<br>Year | Authors  | DOI                                       | Diagnosis focus                                    | Classification task      | Test set   | Class balance<br>of test set | Main<br>classification<br>performance<br>metric | Other classification performance metrics REPORTED | Calibration reported or discussed | Calibration<br>metric | Standardized metrics                      | Has<br>minority<br>metric | Calibration<br>measured | Standarized main metric | Balance<br>binary |
|--|-----------------------|---------------------------------------|---------------------|--|---|--|--------------------------|--|------------------------------|---|---|-----------------------------------|-----------------------|---|---------------------------|-------------------------|-------------------------|-------------------|
| A Coherent Cooperative Learning Framework Based on<br>Transfer Learning  |                       | Computer Aided<br>Diagnosis           | 202                 | 1 Xinxin Shan, Ying Wen, Qingli Li, Yue Lu, and Haibin Cai   | 10.1007/978-3-030-87240<br>-3 10          | Tuberculosis                                       | Binary                   | Three public datasets, inlcuding   | Imbalance                    | Accuracy  | Precision, Recall , F1 score                      | No                                |                       | ['vpp', 'sens',<br>'f1score']             | TRUE                      | No                      | Accuracy                | Imbalanced        |
|  | MICCAI                | Outcome/Disease                       | 202                 | Nkechinyere N. Agu, Joy T. Wu, Hanqing Chao, Ismini Lourentzou,  | 10.1007/978-3-030-87240                   | 9 findings   | Multiple binary          | Chest-XRay14<br>Chest ImaGenome  | Imbalance                    | AUC-ROC   | -   | No                                |                       | ["]                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Chest X-Ray RATCHET: Medical Transformer for Chest X-ray Diagnosis and   |                       | Prediction<br>Clinical                |                     | Arjun Sharma, Mehdi Moradi, Pingkun Yan, and James Hendler<br>1 Benjamin Hou, Georgios Kaissis, Ronald M. Summers, and   | -3_77<br>10.1007/978-3-030-87234          | Report generation for                              | tasks<br>Multiple binary | dataset<br>MIMIC CXR   | Imbalance                    | Other   | -   | No                                |                       | m   | FALSE                     | No                      | Other                   | Imbalanced        |
| Reporting CheXRelNet: An Anatomy-Aware Model for Tracking  | Proceedings           | Applications - Lung<br>Heart and Lung |                     | Bernhard Kainz<br>2 Gaurang Karwande, Amarachi B. Mbakwe, Joy T. Wu, Leo A. Celi,  | -2_28<br>10.1007/978-3-031-16431          | multiple findings                                  | tasks<br>Multiple bipany | Chest ImaGenome  | Imhalance                    | Accuracy  | _   | No                                |                       | m   | FALSE                     | No                      | Accuracy                | Imbalanced        |
| Longitudinal Relationships Between Chest X-Ray   | Proceedings           | Imaging                               | 202                 | Mehdi Moradi, and Ismini Lourentzou  | -6 55                                     | pathologies in two<br>cronological chest<br>x-rays | tasks                    | dataset  | iiiibaiance                  | Accuracy  |   | No                                |                       |   | TALSE                     | 140                     | Accuracy                | imbalanceu        |
| Computer-Aided Tuberculosis Diagnosis with Attribute<br>Reasoning Assistance   | MICCAI<br>Proceedings | Heart and Lung<br>Imaging             | 202                 | Chengwei Pan, Gangming Zhao, Junjie Fang, Baolian Qi, Jiaheng Liu,     Chaowei Fang, Dingwen Zhang, Jinpeng Li and Yizhou Yu   | 10.1007/978-3-031-16431<br>-6_59          | Tuberculosis                                       | Binary                   | TBX11K dataset   | Imbalance                    | Accuracy  | F1score   | No                                |                       | [f1score]                                 | TRUE                      | No                      | Accuracy                | Imbalanced        |
| A Comprehensive Study of Modern Architectures and<br>Regularization Approaches on CheXpert5000   | MICCAI<br>Proceedings | Heart and Lung<br>Imaging             | 202                 | Sontje Ihler , Felix Kuhnke, and Svenja Spindeldreier  | 10.1007/978-3-031-16431<br>-6_62          | 5 findings   | Multiple binary tasks    | Chexpert5000   | Imbalance                    | AUC-ROC   | AUC-PR  | Yes                               | ECE                   | ['auc-pr']                                | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Consistency-Based Semi-supervised Evidential Active<br>Learning for Diagnostic Radiograph Classification                                 |                       | Heart and Lung                        | 202                 | Shafa Balaram, Cuong M. Nguyen, Ashraf Kassim,<br>and Pavitra Krishnaswamy   | 10.1007/978-3-031-16431<br>-6 64          | Multiple findings                                  | Multiple binary tasks    | Chest X-Ray14  | Imbalance                    | AUC-ROC   | AUC-PR  | No                                |                       | ['auc-pr']                                | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Self-Rating Curriculum Learning for Localization and<br>Segmentation of Tuberculosis on Chest Radiograph                                 |                       | Heart and Lung                        | 202                 | 2 Kunlei Hong, Lin Guo, and Yuan-ming Fleming Lure   | 10.1007/978-3-031-16431<br>-6 65          | Tuberculosis                                       | Binary                   | Private set  | Imbalance                    | AUC-ROC   | -   | No                                |                       | l'I                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Did You Get What You Paid For? Rethinking Annotation Cost of<br>Deep Learning Based Computer Aided Detection in Chest<br>Radiographs     |                       | Computer Aided                        | 202                 | 2 Tae Soo Kim, Geonwoon Jang, Sanghyup Lee, and Thijs Kooi   | 10.1007/978-3-031-16437<br>-8_25          | Multiple findings                                  | Multiple binary tasks    | Private set  | Imbalance                    | AUC-ROC   | -   | No                                |                       | ["]                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Self-Ensembling Vision Transformer (SEVIT) for Robust<br>Medical Image Classification  | MICCAI<br>Proceedings | Computer Aided                        | 202                 | 2 Faris Almalik , Mohammad Yaqub , and Karthik Nandakumar  | 10.1007/978-3-031-16437<br>-8 36          | Tuberculosis                                       | Binary                   | Rahman et al   | Balance                      | Accuracy  | AUC-ROC   | No                                |                       | ['auc-roc']                               | FALSE                     | No                      | Accuracy                | Balanced          |
| NVUM: Non-volatile Unbiased Memory for Robust Medical  | MICCAI                | Computer Aided                        | 202                 | Fengbei Liu, Yuanhong Chen, Yu Tian, Yuyuan Liu, Chong Wang,   | -8_36<br>10.1007/978-3-031-16437<br>-8_52 | Multiple findings                                  |                          | OpenI and PadChest   | Imbalance                    | AUC-ROC   | -   | Yes                               | Custom loss           | [7]                                       | FALSE                     | Yes                     | AUC-ROC                 | Imbalanced        |
| Image Classification  Dual-Distribution Discrepancy for Anomaly Detection in Chest   |                       | Computer Aided                        | 202                 | Vasileios Belagiannis, and Gustavo Carneiro<br>2 Yu Cai, Hao Chen, Xin Yang, Yu Zhou, and Kwang-Ting Cheng   | 10.1007/978-3-031-16437                   | Abnormality  | tasks<br>Binary          | RSNA Pneumonia,  | Balance                      | AUC-ROC   | -   | No                                | idificion             | l.  | FALSE                     | No                      | AUC-ROC                 | Balanced          |
| X-Rays   |                       | Diagnosis                             |                     |  | -8_56                                     |  |                          | VinBigData, Chest<br>X-ray Anomaly<br>Detection                                  |                              |   |   |                                   |                       |   |                           |                         |                         |                   |
| GazeRadar: A Gaze and Radiomics-Guided Disease<br>Localization Framework   | MICCAI<br>Proceedings | Computer Aided<br>Diagnosis           | 202                 | <ol><li>Moinak Bhattacharya, Shubham Jain, and Prateek Prasanna</li></ol>  | 10.1007/978-3-031-16437<br>-8_66          | Multiple findings                                  | Multiple binary<br>tasks | RSNA Pneumonia,<br>SIIM-FISABIO-RSNA<br>COVID19,<br>Chest X-ray14,<br>VinBigData | Imbalance                    | AUC-ROC   | -   | No                                |                       | r)  | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| TUNA-Net: Task-Oriented UNsupervised Adversarial Network<br>for Disease Recognition in Cross-domain Chest X-rays                         | MICCAI<br>Proceedings | X-Ray Imaging                         | 201                 | 9 Yuxing Tang, Youbao Tang, Veit Sandfort, Jing Xiao, Ronald M. Summers  | 10.1007/978-3-030-32226<br>-7 48          | Pneumonia  | Binary                   | Guangzhou@and<br>Kaggle Pneumonia  | Imbalance                    | AUC-ROC   | Accuracy;sens;spec;f1scor                         | No                                | no                    | ['accuracy', 'sens'<br>'spec', 'f1score'] | TRUE                      | No                      | AUC-ROC                 | Imbalanced        |
| GraphX NET Chest X-Ray Classification Under Extreme Minimal Supervision  |                       | X-Ray Imaging                         | 201                 | 9 Angelica I. Aviles-Rivero, Nicolas Papadakis, Ruoteng Li, Philip<br>Sellars, Qingnan Fan, Robby T. Tan   | 10.1007/978-3-030-32226<br>-7 56          | Several findings                                   | Multilabel               | ChestXray-14   | Imbalance                    | AUC-ROC   |   | No                                | no                    | [7]                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Automated Detection and Type Classification of Central<br>Venous Catheters in Chest X-Rays   |                       | X-Ray Imaging                         | 201                 | Vaishnavi Subramanian, Hongzhi Wang, Joy T. Wu, Ken C. L.     Wong, Arjun Sharma, Tanveer Syeda-Mahmood  | 10.1007/978-3-030-32226<br>-7 58          | Catheter   | Binary/Multiclas         | ChestXray-14   | Imbalance                    | Accuracy  | Precision;recall;AUC-ROC                          | No                                | NO                    | ['vpp', 'sens',<br>'auc-roc']             | FALSE                     | No                      | Accuracy                | Imbalanced        |
| Adaptive Image-Feature Learning for Disease Classification   | MICCAI                | X-Ray Imaging                         | 201                 | Hendrik Burwinkel, Anees Kazi, Gerome Vivar, Shadi Albarqouni,   | 10.1007/978-3-030-32226                   | Several findings                                   | Multilabel               | ChestXray-14   | Imbalance                    | Accuracy  |   | No                                | no                    | ["]                                       | FALSE                     | No                      | Accuracy                | Imbalanced        |
| Using Inductive Graph Networks Quantifying and Leveraging Classification Uncertainty for   | Proceedings<br>MICCAI | X-Ray Imaging                         | 201                 | Guillaume Zahnd, Nassir Navab<br>9 Florin C. Ghesu, Bogdan Georgescu, Eli Gibson, Sebastian  | -7_71<br>10.1007/978-3-030-32226          | Several findings                                   | Multilabel               | ChestXray-14 and   | Imbalance                    | AUC-ROC   | f1score   | Yes                               | Predictive            | [f1score]                                 | TRUE                      | Yes                     | AUC-ROC                 | Imbalanced        |
| Chest Radiograph Assessment Automatic Radiology Report Generation Based on Multi-view  |                       | X-Ray Imaging                         | 201                 | Guendel, Mannudeep K. Kalra, Ramandeep Singh<br>9 Jianbo Yuan, Haofu Liao, Rui Luo, Jiebo Luo  | -7_75<br>10.1007/978-3-030-32226          | Several findings                                   | Multilabel               | PLCO<br>CheXpert and   | Imbalance                    | AUC-ROC   |   | No                                | uncertainty<br>no     | l.  | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Image Fusion and Medical Concept Enrichment  Multi-label Thoracic Disease Image Classification with                                      | Proceedings<br>MICCAI | X-Ray Imaging                         | 201                 | 9 Congbo Ma, Hu Wang, Steven C. H. Hoi   | -7_80<br>10.1007/978-3-030-32226          | Several findings                                   | Multilabel               | ChestXray14<br>ChestXray-14  | Imbalance                    | AUC-ROC   |   | No                                | no                    | ľ"  | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Cross-Attention Networks InfoMask: Masked Variational Latent Representation to   |                       | X-Ray Imaging                         | 201                 | 9 Saeid Asgari Taghanak, Mohammad Havaei, Tess Berthier, Francis   | -7_81<br>10.1007/978-3-030-32226          | Several findings                                   | Multilabel               | ChestXray-14   | Imbalance                    | Accuracy  | AUC-ROC   | No                                | no                    | ['auc-roc']                               | FALSE                     | No                      | Accuracy                | Imbalanced        |
| Localize Chest Disease Longitudinal Change Detection on Chest X-rays Using   | Proceedings<br>MICCAI | X-Ray Imaging                         | 201                 | Dutil, Lisa Di Jorio, Ghassan Hamarneh<br>Dong Yul Oh, Jihang Kim, Kyong Joon Lee  | -7_82<br>10.1007/978-3-030-32226          | Longitudinal change                                | Binary                   | Private set  | Imbalance                    | AUC-ROC   | sens;spec   | No                                | no                    | ['sens', 'spec']                          | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Geometric Correlation Maps Semi-supervised Learning by Disentangling and   | Proceedings<br>MICCAI | X-Ray Imaging                         | 201                 | 9 Prashnna Kumar Gyawali, Zhiyuan Li, Sandesh Ghimire, Linwei  | -7_83<br>10.1007/978-3-030-32226          | Several findings                                   | Multilabel               | CheXpert   | Imbalance                    | AUC-ROC   |   | No                                | NO                    | ľ"  | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Self-ensembling over Stochastic Latent Space Endotracheal Tube Detection and Segmentation in Chest                                       | Proceedings<br>MICCAI | X-Ray Imaging                         | 201                 | Wang<br>9 Maayan Frid-Adar, Rula Amer, Hayit Greenspan   | -7_85<br>10.1007/978-3-030-32226          | Endotracheal tube                                  | Binary                   | ChestXray-14   | Balance                      | AUC-ROC   | sens;spec   | No                                | no                    | ['sens', 'spec']                          | FALSE                     | No                      | AUC-ROC                 | Balanced          |
| Radiographs Using Synthetic Data<br>Learning Interpretable Features via Adversarially Robust   | Proceedings<br>MICCAI | X-Ray Imaging                         | 201                 | 9 Ashkan Khakza, Shadi Albarqouni, Nassir Navab  | -7_87<br>10.1007/978-3-030-32226          | Several findings                                   | Multilabel               | ChestXray-14   | Imbalance                    | AUC-ROC   |   | No                                | NO                    | [7]                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Optimization  Learning Semantics-Enriched Representation via   | Proceedings           | Machine Learning                      |                     | Fatemeh Haghighi1, Mohammad Reza Hosseinzadeh Taher1,  | -7_88<br>10.1007/978-3-030-59710          | Several findings                                   | Multilabel               | ChestXray-14   | Imbalance                    | AUC-ROC   |   | No                                | NO                    | ľ)  | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Self-discovery, Self-classification, and Self-restoration  |                       | Methodologies                         |                     | Zongwei Zhou1,Michael B. Gotway2, and Jianming Liang1  | -8_14                                     | (assumed<br>ChestX-Ray14<br>classes)               |                          |  |                              |   |   |                                   |                       |   |                           |                         |                         |                   |
| DECAPS: Detail-Oriented Capsule Networks   |                       | Machine Learning<br>Methodologies     | 202                 | D Aryan Mobiny, Pengyu Yuan, Pietro Antonio Cicalese and Hien Van<br>Nguyen  | 10.1007/978-3-030-59710<br>-8_15          | Several findings<br>(CheXpert main 5<br>classes)   | Multilabel               | CheXpert and RSNA<br>Pneumonia   | Imbalance                    | AUC-ROC   |   | No                                | NO                    | ניז                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Domain Aware Medical Image Classifier Interpretation by<br>Counterfactual Impact Analysis  | Proceedings           | Machine Learning<br>Methodologies     | 202                 | D Dimitrios Lenis, David Major, Maria Wimmer, Astrid Berg, Gert Sluiter, and Katja B�uhler   | 10.1007/978-3-030-59710<br>-8_31          |  | Binary                   | Private set  | Balance                      | AUC-ROC   |   | No                                | NO                    | [1]                                       | FALSE                     | No                      | AUC-ROC                 | Balanced          |
| Comparing to Learn: Surpassing ImageNet Pretraining on<br>Radiographs by Comparing Image Representations                                 | MICCAI<br>Proceedings | Machine Learning<br>Methodologies     | 202                 | D Hong-Yu Zhou, Shuang Yu, Cheng Bian, Yifan Hu, Kai Ma, Yefeng Zheng  | 10.1007/978-3-030-59710<br>-8 39          | Several findings                                   | Multilabel               | Multiple public datasets   | Imbalance                    | AUC-ROC   |   | No                                | NO                    | l.1                                       | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| SALAD: Self-supervised Aggregation Learning for Anomaly<br>Detection on X-Rays   | MICCAI                | Machine Learning<br>Methodologies     | 202                 | D Behzad Bozorgtaba, Dwarikanath Mahapatra, Guillaume Vray,<br>Jean-Philippe Thiran  | 10.1007/978-3-030-59710<br>-8 46          | Abnormal vs normal                                 | Binary                   | ChestXray-14   | Imbalance                    | AUC-ROC,<br>AUC-PR                              |   | No                                | NO                    | m   | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Semi-supervised Medical Image Classification with Global<br>Latent Mixing  | MICCAI                | Machine Learning<br>Methodologies     | 202                 | Prashnna Kumar Gyawali, Sandesh Ghimire, Pradeep Bajracharya,     Zhiyuan Li, Linwei Wang  | 0_40                                      | Several findings                                   | Multilabel               | CheXpert   | Imbalance                    | AUC-ROC   |   | Yes                               | Reliability plots     | l.1                                       | FALSE                     | Yes                     | AUC-ROC                 | Imbalanced        |
| Characterizing Label Errors: Confident Learning for  | MICCAI                | Machine Learning<br>Methodologies     | 202                 | Minqing Zhang, Jiantao Gao, Zhen Lyu, Weibing Zhao, Qin Wang,  |   | Noisy mask labels                                  | Binary                   | JSRT   | Imbalance                    | F1-score  | Recall;precision                                  | No                                | no                    | ['sens', 'vpp']                           | FALSE                     | No                      | Other                   | Imbalanced        |
| Noisy-Labeled Image Segmentation  Joint Modeling of Chest Radiographs and Radiology Reports for Pulmonary Edema Assessment               | MICCAI                | Machine Learning<br>Applications      | 202                 | Weizhen Ding, Sheng Wang, Zhen Li, Shuguang Cui<br>D Geeticka Chauhan, Ruizhi Liao, William Wells, Jacob Andreas, Xin<br>Wang, Seth Berkowitz, Steven Horng, Peter Szolovits, Polina<br>Golland    |   | Pulmonary edema                                    | Binary                   | MIMIC-CXR  | Imbalance                    | AUC-ROC   | f1score   | No                                | no                    | [f1score]                                 | TRUE                      | No                      | AUC-ROC                 | Imbalanced        |
| Chest X-Ray Report Generation Through Fine-Grained Label Learning  | MICCAI<br>Proceedings | Machine Learning<br>Applications      | 202                 | Tanveer Syeda-Mahmood, Ken C. L. Wong, Yaniv Gur, Joy T. Wu,     Ashutosh Jadhav, Satyananda Kashyap, Alexandros Karargyris,     Anuo Pillai. Ariun Sharma. Ali Bin Sved. Orest Boyko. Mehdi Morad | 10.1007/978-3-030-59713<br>-9_54          | Several findings                                   | Multilabel               | MIMIC-CXR and<br>ChestX-Ray14  | Imbalance                    | AUC-ROC   | f1score;precision                                 | No                                | no                    | ['f1score', 'vpp']                        | TRUE                      | No                      | AUC-ROC                 | Imbalanced        |
| Abnormality Detection in Chest X-Ray Images Using<br>Uncertainty Prediction Autoencoders   | MICCAI<br>Proceedings | Heart and Lung                        | 202                 | Yifan Mao, Fei-Fei Xue, Ruixuan Wang, Jianguo Zhang, Wei-Shi<br>Zheng, Hongmei Liu   | 10.1007/978-3-030-59725<br>-2 51          | Pneumonia  | Binary                   | Kaggle pneumonia<br>and pediatric  | Imbalance                    | AUC-ROC   | f1score;equal error rate                          | Yes                               | Uncertainty           | ['f1score', 'equal<br>error rate']        | TRUE                      | Yes                     | AUC-ROC                 | Imbalanced        |
| Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs |                       | Clin Infect Dis                       | 201                 | Zheng, Hongmei Liu  Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, Goo JM, Aum J, Yim JJ, Park CM; Deep Learning-Based Automatic Detection Algorithm Development and Evaluation Group.         | -2_51<br>10.1093/cid/ciy967               | ТВ   | Binary                   | Institutional+Montgo<br>mery+Shenzhen  | Balance                      | AUC-ROC   | sens;spec   | No                                |                       | ['sens', 'spec']                          | FALSE                     | No                      | AUC-ROC                 | Balanced          |
|  | Pubmed<br>search      | Int J Environ Res<br>Public Health    | 201                 | Heo SJ, Kim Y, Yun S, Lim SS, Kim J, Nam CM, Park EC, Jung I, Yoon JH.   | 10.3390/ijerph16020250                    | ТВ   | Binary                   | Institutional  | Imbalance                    | AUC-ROC   | sens;spec   | No                                |                       | ['sens', 'spec']                          | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
| Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification  | Pubmed<br>search      | Sci Rep                               | 201                 | Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A.  | 10.1038/s41598-019-4229                   | 14 abnormalities                                   | Multilabel               | ChestX-Ray14   | Imbalance                    | AUC-ROC   | sens;spec   | No                                |                       | ['sens', 'spec']                          | FALSE                     | No                      | AUC-ROC                 | Imbalanced        |
|  |                       |                                       |                     |  |   |  |                          |  |                              |   |   |                                   |                       |   |                           |                         |                         |                   |

| Using deep-learning techniques for pulmonary-thoracic segmentations and improvement of pneumonia diagnosis in  | Pubmed search    | Pediatr Pulmonol                       | 2019 | E L, Zhao B, Guo Y, Zheng C, Zhang M, Lin J, Luo Y, Cai Y, Song X,<br>Liang H.  | 10.1002/ppul.24431                               | Pneumonia type                                    | Binary     | Guangzhou WCMC +<br>JSRT       | Imbalance | AUC-ROC   | segmentation metrics                 | No  |  | ['segmentation<br>metrics']                            | FALSE | No  | AUC-ROC  | Imbalanced |
|--|------------------|--|------|---|--|---|------------|--------------------------------|-----------|-----------|--------------------------------------|-----|--|--|-------|-----|----------|------------|
| pediatric chest radiographs  Deep Learning to Assess Long-term Mortality From Chest  Radiographs   | Pubmed<br>search | JAMA Netw Open                         | 2019 | Lu MT, Ivanov A, Mayrhofer T, Hosny A, Aerts HJWL, Hoffmann U.  | 10.1001/jamanetworkopen<br>2019 7416             | Mortality   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec                            | Yes | Calibration  | ['sens', 'spec']                                       | FALSE | Yes | AUC-ROC  | Imbalanced |
| Value of three deep learning systems  Study of three deep learning systems   | Pubmed           | Sci Rep                                | 2019 | Qin ZZ, Sander MS, Rai B, Titahong CN, Sudrungrot S, Laah SN,<br>Adhikari LM, Carter EJ, Puri L, Codlin AJ, Creswell J.   | 10.1038/s41598-019-5150<br>3-3                   | ТВ  | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;accuracy                   | No  |  | ['sens', 'spec',<br>'accuracy']                        | FALSE | No  | AUC-ROC  | Imbalanced |
| Development and Validation of Deep Learning-based<br>Automatic Detection Algorithm for Malignant Pulmonary<br>Nodules on Chest Radiographs   | Pubmed<br>search | Radiology                              | 2019 | Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, Vu TH, Sohn JH, Hwang S, Goo JM, Park CM.   | 10.1148/radiol.201818023<br>7                    | Nodules   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;F1score;JAFRO<br>C-FOM     | No  |  | ['sens', 'spec',<br>'f1score',<br>'jafroc-fom']        | TRUE  | No  | AUC-ROC  | Imbalanced |
| Development and Validation of a Deep Learning-Based<br>Automated Detection Algorithm for Major Thoracic Diseases on<br>Chest Radiographs   | Pubmed<br>search | JAMA Netw Open                         |      | Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, Goo JM, Aum J, Yim JJ, Cohen JG, Ferretti GR, Park CM; DLAD Development and Evaluation Group.  | 10.1001/jamanetworkopen<br>.2019.1095            | Major thoracic disease                            | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;F1score;AUC-J<br>AFROC     | No  |  | ['sens', 'spec',<br>'f1score',<br>'auc-jafroc']        | TRUE  | No  | AUC-ROC  | Imbalanced |
| Application of deep learning-based computer-aided detection<br>system: detecting pneumothorax on chest radiograph after<br>siopsy  | Pubmed<br>search | Eur Radiol                             |      |   | 10.1007/s00330-019-0613<br>0-x                   |   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;accuracy; PPV;<br>NPV      | No  |  | ['sens', 'spec',<br>'accuracy', 'ppv',<br>'npv']       |       | No  | AUC-ROC  | Imbalanced |
| Deep Learning Method for Automated Classification of<br>Anteroposterior and Posteroanterior Chest Radiographs  | Pubmed<br>search | J Digit Imaging                        | 2019 | Kim TK, Yi PH, Wei J, Shin JW, Hager G, Hui FK, Sair HI, Lin CT.  | 10.1007/s10278-019-0020<br>8-0                   | AP/PA   | Binary     | ChestX-Ray14                   | Imbalance | AUC-ROC   | accuracy;sens;spec;PPV;N<br>PV       | No  |  | ['accuracy', 'sens',<br>'spec', 'ppv', 'npv']          | TRUE  | No  | AUC-ROC  | Imbalanced |
| Deep Learning for Chest Radiograph Diagnosis in the<br>Emergency Department  | Pubmed<br>search | Radiology                              |      | Hwang EJ, Nam JG, Lim WH, Park SJ, Jeong YS, Kang JH, Hong<br>EK, Kim TM, Goo JM, Park S, Kim KH, Park CM.  | 5  |   | Binary     | Institutional                  | Imbalance | AUC-ROC   | , , , ,                              | No  |  | ['sens', 'spec',<br>'ppv', 'npv']                      | moL   | No  | AUC-ROC  | Imbalanced |
| Assessment of an ensemble of machine learning models<br>oward abnormality detection in chest radiographs   | Pubmed<br>search | Annu Int Conf IEEE<br>Eng Med Biol Soc | 2019 | Rajaraman S, Sornapudi S, Kohli M, Antani S.  | 10.1109/EMBC.2019.8856<br>715                    | Abnormal  | Binary     | Kaggle                         | Imbalance | Accuracy  | AUC-ROC, f1score; MCC                | No  |  | ['auc-roc',<br>'f1score', 'mcc']                       | TRUE  | No  | Accuracy | Imbalanced |
| Human-machine partnership with artificial intelligence for chest<br>adiograph diagnosis  | Pubmed<br>search | NPJ Digit Med                          |      | Patel BN, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J,<br>Rajpurkar P, Amrhein T, Gupta R, Halabi S, Langlotz C, Lo E,<br>Mammarappallil J, Mariano AJ, Riley G, Seekins J, Shen L, Zucker<br>E. Lunoren M. | 10.1038/s41746-019-0189<br>-7                    | Pneumonia   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;f1score                    | Yes | Brier; MAE   | ['sens', 'spec',<br>'f1score']                         | TRUE  | Yes | AUC-ROC  | Imbalanced |
| Deep learning-based detection system for multiclass lesions on<br>chest radiographs: comparison with observer readings   | Pubmed<br>search | Eur Radiol                             | 2020 | Park S, Lee SM, Lee KH, Jung KH, Bae W, Choe J, Seo JB.   | 10.1007/s00330-019-0653<br>2-x                   | Abnormal  | Binary     | Institutional                  | Balance   | AUC-ROC   | sens;spec;JAFROC-FOM                 | No  |  | ['sens', 'spec',<br>'jafroc-fom']                      | FALSE | No  | AUC-ROC  | Balanced   |
| Optimal matrix size of chest radiographs for computer-aided detection on lung nodule or mass with deep learning  | Pubmed search    | Eur Radiol                             | 2020 | Kim YG, Lee SM, Lee KH, Jang R, Seo JB, Kim N.  | 10.1007/s00330-020-0689<br>2-9                   | Nodules   | Binary     | Institutional                  | Balance   | AUC-ROC   | sens                                 | No  | -  | ['sens']   | FALSE | No  | AUC-ROC  | Balanced   |
| Deep learning for automated classification of<br>tuberculosis-related chest X-Ray: dataset distribution shift<br>limits diagnostic performance generalizability                            | Pubmed<br>search | Heliyon                                | 2020 | Sathitratanacheewin S, Sunanta P, Pongpirul K.  | 10.1016/j.heliyon.2020.e0<br>4614                | ТВ  | Binary     | Institutional+ChestXR<br>ay8   | Balance   | AUC-ROC   | sens;spec                            | No  |  | ['sens', 'spec']                                       | FALSE | No  | AUC-ROC  | Balanced   |
| Deep learning to predict elevated pulmonary artery pressure in<br>patients with suspected pulmonary hypertension using<br>standard chest X ray   | Pubmed<br>search | Sci Rep                                |      |   | 10.1038/s41598-020-7635<br>9-w                   | pressure  | Binary     |                                | Balance   | AUC-ROC   | Prognosis                            | No  |  | ['prognosis']  | FALSE | No  | AUC-ROC  | Balanced   |
| Chest x-ray analysis with deep learning-based software as a<br>riage test for pulmonary tuberculosis: a prospective study of<br>diagnostic accuracy for culture-confirmed disease          | Pubmed<br>search | Lancet Digit Health                    |      |   | 0221-1   |   | Binary     | Institutional                  | Balance   | sens;spec |                                      | No  |  | [1]  |       | No  | Other    | Balanced   |
| dentifying pulmonary nodules or masses on chest radiography<br>using deep learning: external validation and strategies to<br>mprove clinical practice                                      | search           | Clin Radiol                            |      | Liang CH, Liu YC, Wu MT, Garcia-Castro F, Alberich-Bayarri A, Wu<br>FZ.   | 5  |   | Binary     |                                | Balance   | AUC-ROC   | sens;spec;PPV;NPV;accur<br>acy;LR    |     |  | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy', 'lr'] | TRUE  | No  | AUC-ROC  | Balanced   |
| Modality-specific deep learning model ensembles toward<br>mproving TB detection in chest radiographs   | Pubmed<br>search | IEEE Access                            |      |   | 10.1109/access.2020.297<br>1257                  |   | Binary     | Shenzhen                       | Balance   | Accuracy  | AUC-ROC;sens;spec;F1sc<br>ore;MCC    |     |  | ['auc-roc', 'sens',<br>'spec', 'f1score',<br>'mcc']    |       | No  | Accuracy | Balanced   |
| Potential of deep learning in assessing pneumoconiosis depicted on digital chest radiography   | Pubmed<br>search | Occup Environ<br>Med                   |      |   | 10.1136/oemed-2019-106<br>386                    |   | Binary     | Institutional                  | Balance   | AUC-ROC   | sens;spec;PPV;NPV;kappa              |     |  | ['sens', 'spec',<br>'ppv', 'npv',<br>'kappa']          |       | No  | AUC-ROC  | Balanced   |
| A promising approach for screening pulmonary hypertension<br>based on frontal chest radiographs using deep learning: A<br>retrospective study  | Pubmed<br>search | PLoS One                               |      | <u> </u>  | 6378   | pressure  | Binary     | Institutional                  | Balance   | AUC-ROC   | MAE; F1score; sens;spec;<br>PPV;NPV  |     |  | ['mae', 'f1score',<br>'sens', 'spec',<br>'ppv', 'npv'] |       | No  | AUC-ROC  | Balanced   |
| CheXaid: deep learning assistance for physician diagnosis of<br>uberculosis using chest x-rays in patients with HIV  | search           | NPJ Digit Med                          |      | Rajpurkar P, O'Connell C, Schechter A, Asnani N, Li J, Kiani A, Ball RL, Mendelson M, Maartens G, van Hoving DJ, Griesel R, Ng AY, Boyles TH, Lungren MP.   | 2-2  |   | Binary     | Institutional                  | Balance   | Accuracy  |                                      | No  |  | ['sens', 'spec',<br>'auc-roc']                         | FALSE | No  | Accuracy | Balanced   |
| extravalidation and reproducibility results of a commercial deep<br>earning-based automatic detection algorithm for pulmonary<br>lodules on chest radiographs at tertiary hospital         | search           | J Med Imaging<br>Radiat Oncol          |      | Koo YH, Shin KE, Park JS, Lee JW, Byun S, Lee H.  | 10.1111/1754-9485.13105                          |   | Binary     | Institutional                  | Balance   | AUC-ROC   | JAFROC-FOM; sens;spec                |     |  | ['jafroc-fom',<br>'sens', 'spec']                      | FALSE | No  | AUC-ROC  | Balanced   |
| Deep Learning Using Chest Radiographs to Identify High-Risk<br>Smokers for Lung Cancer Screening Computed Tomography:<br>Development and Validation of a Prediction Model                  | search           | Ann Intern Med                         |      | Lu MT, Raghu VK, Mayrhofer T, Aerts HJWL, Hoffmann U.   | 10.7326/M20-1868                                 | Nodules   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;prognosis;                 | Yes |  | ['sens', 'spec',<br>'prognosis', "]                    |       | No  | AUC-ROC  | Imbalanced |
| Smart chest X-ray worklist prioritization using artificial<br>ntelligence: a clinical workflow simulation  | Pubmed<br>search | Eur Radiol                             |      | radin o, raropp 1, taron 1.   | 10.1007/s00330-020-0748<br>0-7                   |   | Multilabel | ChestX-Ray8                    | Imbalance | AUC-ROC   | sens;spec;time to report             | No  |  | ['sens', 'spec',<br>'time to report']                  | FALSE | No  | AUC-ROC  | Imbalanced |
| Performance of a Deep Learning Algorithm Compared with<br>Radiologic Interpretation for Lung Cancer Detection on Chest<br>Radiographs in a Health Screening Population                     | Pubmed<br>search | Radiology                              | 2020 | Lee JH, Sun HY, Park S, Kim H, Hwang EJ, Goo JM, Park CM.   | 10.1148/radiol.202020124<br>0                    | Nodules   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec; PPV;NPV;<br>accuracy      | No  |  | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy']       | TRUE  | No  | AUC-ROC  | Imbalanced |
| Deep learning, reusable and problem-based architectures for<br>detection of consolidation on chest X-ray images  | Pubmed<br>search | Comput Methods<br>Programs Biomed      |      | Behzadi-Khormouji H, Rostami H, Salehi S, Derakhshande-Rishehri<br>T, Masoumi M, Salemi S, Keshavarz A, Gholamrezanezhad A,<br>Assadi M, Batouli A.   | 62   |   | Binary     | Guangzhou WCMC                 | Imbalance | AUC-ROC   | sens;spec;accuracy                   | No  |  | ['sens', 'spec',<br>'accuracy']                        | FALSE | No  | AUC-ROC  | Imbalanced |
| Augmenting Interpretation of Chest Radiographs With Deep<br>Learning Probability Maps  | Pubmed<br>search | J Thorac Imaging                       | 2020 | Hurt B, Yen A, Kligerman S, Hsiao A.  | 10.1097/RTI.0000000000<br>000505                 | Pneumonia   | Binary     | Kaggle                         | Imbalance | AUC-ROC   | sens;spec                            | No  |  | ['sens', 'spec']                                       | FALSE | No  | AUC-ROC  | Imbalanced |
| Deep Learning-based Automatic Detection Algorithm for<br>Reducing Overlooked Lung Cancers on Chest Radiographs   | Pubmed<br>search | Radiology                              | 2020 | Jang S, Song H, Shin YJ, Kim J, Kim J, Lee KW, Lee SS, Lee W,<br>Lee S, Lee KH.   | 10.1148/radiol.202020016<br>5                    | Nodules   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec                            | No  |  | ['sens', 'spec']                                       | FALSE | No  | AUC-ROC  | Imbalanced |
| mpact of Confounding Thoracic Tubes and Pleural Dehiscence<br>Extent on Artificial Intelligence Pneumothorax Detection in<br>Chest Radiographs   | Pubmed<br>search | Invest Radiol                          |      | Rueckel J, Trappmann L, Schachtner B, Wesp P, Hoppe BF, Fink N, Ricke J, Dinkel J, Ingrisch M, Sabel BO.  | 10.1097/RLI.000000000000000000000000000000000000 | Pneumothorax with<br>and without thoracic<br>tube | Binary     | Institutional                  | Imbalance | AUC-ROC   |                                      | No  |  | [1]  | FALSE | No  | AUC-ROC  | Imbalanced |
| Development and validation of a deep learning algorithm<br>letecting 10 common abnormalities on chest radiographs  | Pubmed<br>search | Eur Respir J                           | 2020 | Nam JG, Kim M, Park J, Hwang EJ, Lee JH, Hong JH, Goo JM,<br>Park CM.   | 10.1183/13993003.03061-<br>2020                  | 10 anomalies                                      | Multilabel |                                | Imbalance | AUC-ROC   | time to report; sens; spec; accuracy | No  |  | ['time to report',<br>'sens', 'spec',<br>'accuracy']   | FALSE | No  | AUC-ROC  | Imbalanced |
| mage-based Deep Learning in Diagnosing the Etiology of<br>Pneumonia on Pediatric Chest X-rays  | Pubmed search    | Pediatr Pulmonol                       |      | Longjiang E, Baisong Zhao, Liu H, Zheng C, Song X, Cai Y, Liang<br>H.   |  | Pneumonia type                                    | Binary     | Institutional                  | Imbalance | AUC-ROC   |                                      | No  | -  | ['kappa', 'sens',<br>'spec', 'lr']                     | FALSE | No  | AUC-ROC  | Imbalanced |
| Chest Radiograph Interpretation with Deep Learning Models:<br>Assessment with Radiologist-adjudicated Reference Standards<br>and Population-adjusted Evaluation                            | Pubmed<br>search | Radiology                              |      | Ding A, Corrado GS, Tse D, Shetty S.  | 10.1148/radiol.201919129<br>3                    | Nodules; opacity;<br>pneumothorax;<br>fracture    | Multilabel | Institutional+ChestX-<br>Ray14 | Imbalance | AUC-ROC   | sens;spec;PPV;NPV                    | No  |  | ['sens', 'spec',<br>'ppv', 'npv']                      | TRUE  | No  | AUC-ROC  | Imbalanced |
| Prediction of Pulmonary to Systemic Flow Ratio in Patients<br>With Congenital Heart Disease Using Deep Learning-Based<br>Analysis of Chest Radiographs                                     | Pubmed<br>search | JAMA Cardiol                           |      | Toba S, Mitani Y, Yodoya N, Ohashi H, Sawada H, Hayakawa H,<br>Hirayama M, Futsuki A, Yamamoto N, Ito H, Konuma T, Shimpo H,<br>Takao M.  | 10.1001/jamacardio.2019.<br>5620                 | High pulmonary to<br>systemic flow ratio          | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec;PPV;NPV;accur<br>acy       | Yes | Regression<br>metrics for<br>systemic flow<br>ratio prediction | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy']       | TRUE  | Yes | AUC-ROC  | Imbalanced |
| Artificial Intelligence Algorithm Detecting Lung Infection in<br>Supine Chest Radiographs of Critically III Patients With a<br>Diagnostic Accuracy Similar to Board-Certified Radiologists | Pubmed<br>search | Crit Care Med                          |      | Rueckel J, Kunz WG, Hoppe BF, Patzig M, Notohamiprodjo M,<br>Meinel FG, Cyran CC, Ingrisch M, Ricke J, Sabel BO.  | 10.1097/CCM.000000000<br>0004397                 | Pneumonia and<br>pleural efussion                 | Multilabel | Institutional                  | Imbalance | AUC-ROC   | sens;spec;PPV;NPV;accur<br>acy       |     |  | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy']       |       | No  | AUC-ROC  | Imbalanced |
| Clinical Validation of a Deep Learning Algorithm for Detection<br>of Pneumonia on Chest Radiographs in Emergency<br>Department Patients with Acute Febrile Respiratory Illness             | Pubmed<br>search | J Clin Med                             | 2020 | Kim JH, Kim JY, Kim GH, Kang D, Kim IJ, Seo J, Andrews JR, Park<br>CM.  |  | Pneumonia   | Binary     | Institutional                  | Imbalance | AUC-ROC   | sens;spec; PPV;NPV;                  | No  |  | ['sens', 'spec',<br>'ppv', 'npv', "]                   | TRUE  | No  | AUC-ROC  | Imbalanced |
| An Efficient Method to Predict Pneumonia from Chest X-Rays<br>Using Deep Learning Approach   | Pubmed<br>search | Stud Health<br>Technol Inform          | 2020 | Shah U, Abd-Alrazeq A, Alam T, Househ M, Shah Z.  | 10.3233/SHTI200594                               | Pneumonia   | Binary     | Kaggle                         | Imbalance | Accuracy  | f1score; sens;spec;<br>PPV:NPV       | No  |  | ['f1score', 'sens',<br>'spec', 'ppv', 'npv']           | TRUE  | No  | Accuracy | Imbalanced |

| Automated identification of chest radiographs with referable<br>abnormality with deep learning: need for recalibration   | Pubmed<br>search | Eur Radiol                        | 2020   | -twang E.J, Kim H, Lee JH, Goo JM, Park CM.  | 10.1007/s00330-020-0706<br>2-7         | Abnormalities                                      | Binary                   | Institutional   | Imbalance | Brier                       | AUC-ROC, sens; spec;<br>PPV;NPV;accuracy   | Yes Brier;<br>Spiegelhalter<br>s Z statistics;<br>Max and mean<br>calibration erro | ['auc-roc', 'sens', 'spec', 'ppv', 'npv', 'accuracy']                             | TRUE  | Yes | Other    | Imbalanced |
|--|------------------|-----------------------------------|--------|--|--|--|--------------------------|---|-----------|-----------------------------|--|--|---|-------|-----|----------|------------|
| Deep-Pneumonia Framework Using Deep Learning Models<br>Based on Chest X-Ray Images   | Pubmed search    | Diagnostics (Basel)               | 2020   | Elshennawy NM, Ibrahim DM.   | 10.3390/diagnostics10090<br>649        | Pneumonia  | Binary                   | kaggle  | Imbalance | AUC-ROC                     | NPV;sens;F1score;accurac<br>y  |  | ['npv', 'sens',<br>'f1score',<br>'accuracy']                                      | TRUE  | No  | AUC-ROC  | Imbalanced |
| Validation of a Deep Learning Algorithm for the Detection of<br>Malignant Pulmonary Nodules in Chest Radiographs   | Pubmed search    | JAMA Netw Open                    | 2020   | Yoo H, Kim KH, Singh R, Digumarthy SR, Kalra MK.   | 10.1001/jamanetworkoper<br>.2020.17135 | Nodules  | Binary                   | Institutional   | Imbalance | AUC-ROC                     | SENS;SPEC;ppv;npv;<br>kappa  | No   | ['sens', 'spec',<br>'ppv', 'npv',<br>'kappa']                                     | TRUE  | No  | AUC-ROC  | Imbalanced |
| Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents   | Pubmed<br>search | JAMA Netw Open                    |        | Nu JT, Wong KCL, Gur Y, Ansari N, Karargyris A, Sharma A, Morris<br>M, Saboury B, Ahmad H, Boyko O, Syed A, Jadhav A, Wang H,<br>Pillai A, Kashyan S. Moradi M. Syeda-Mahmood T.     | 10.1001/jamanetworkoper<br>.2020.22779 | 72 findings  | Multilabel               | MIMIC+ChestX-Ray1   | Imbalance | AUC-ROC                     | sens;spec;PPV;kappa;   | No   | ['sens', 'spec',<br>'ppv', 'kappa', "]  | TRUE  | No  | AUC-ROC  | Imbalanced |
| Impact of hybrid supervision approaches on the performance o<br>artificial intelligence for the classification of chest radiographs  |                  | Comput Biol Med                   |        | Ellis R, Ellestad E, Elicker B, Hope MD, Tosun D.  | 10.1016/j.compbiomed.20<br>20.103699   | Abnormal   | Binary                   | Institutional   | Imbalance | AUC-ROC                     | Sens;spec;f1Score;NPP;P<br>PV; accuracy  | Yes  | ['sens', 'spec',<br>'f1score', 'npp',<br>'ppv', 'accuracy']                       | TRUE  | No  | AUC-ROC  | Imbalanced |
| Diagnosing Heart Failure from Chest X-Ray Images Using<br>Deep Learning  | Pubmed search    | Int Heart J                       |        | Matsumoto T, Kodera S, Shinohara H, leki H, Yamaguchi T,<br>-ligashikuni Y, Kiyosue A, Ito K, Ando J, Takimoto E, Akazawa H,<br>Morita H. Komuro I.                                  | 10.1536/ihj.19-714                     | Heart failure                                      | Binary                   | ChestX-Ray14  | Balance   | Accuracy                    | sens;spec  | No no  | ['sens', 'spec']  | FALSE | No  | Accuracy | Balanced   |
| Deep learning-based automated detection algorithm for active<br>bulmonary tuberculosis on chest radiographs: diagnostic<br>performance in systematic screening of asymptomatic<br>ndividuals | Pubmed<br>search | Eur Radiol                        | 2021 I | Lee JH, Park S, Hwang EJ, Goo JM, Lee WY, Lee S, Kim H,<br>Andrews JR, Park CM.  | 10.1007/s00330-020-0721<br>9-4         | ТВ   | Binary                   | Institutional   | Imbalance | AUC-ROC                     | sens;spec;PPV;NPV  | No   | ['sens', 'spec', 'ppv', 'npv']  | TRUE  | No  | AUC-ROC  | Imbalanced |
| Automatically measuring the Cobb angle and screening for<br>scoliosis on chest radiograph with a novel artificial intelligence<br>method   | Pubmed<br>search | Am J Transl Res                   | 2022   | Xie L, Zhang Q, He D, Wang Q, Fang Y, Ge T, Jiang Y, Tian W.   |  | Scoliosis diagnosis                                | Binary                   | Shenzhen  | Imbalance | AUC-ROC                     | ICC/Kappa/Accuracy/Sensi<br>tivity/Specificity   | No   | ['icc', 'kappa',<br>'accuracy', 'sens',<br>'spec']                                | FALSE | No  | AUC-ROC  | Imbalanced |
| Deep Learning Models to Predict Fatal Pneumonia Using<br>Chest X-Ray Images  | Pubmed<br>search | Can Respir J                      | 2022   | Anai S, Hisasue J, Takaki Y, Hara N.   | 10.1155/2022/8026580                   | Fatal pneumonia                                    | Binary                   | Private set   | Balance   | AUC-PR                      | sensitivity, specificity, PPV,<br>negative predictive value<br>(NPV), accuracy, and F1<br>score            | No   | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy',<br>'f1score']                    | TRUE  | No  | AUC-ROC  | Balanced   |
| Chest X-ray-based opportunistic screening of sarcopenia using deep learning  | search           | J Cachexia<br>Sarcopenia Muscle   | 2023   | Ryu J, Eom S, Kim HC, Kim CO, Rhee Y, You SC, Hong N.  | 10.1002/jcsm.13144                     | Sarcopenia   | Binary                   | Private set   | Imbalance | AUC-ROC                     | AUC-PR, sensitivity,<br>specificity, PPV, negative<br>predictive value (NPV),<br>accuracy, F1 score, Brier | Yes  | ['auc-pr', 'sens',<br>'spec', 'ppv', 'npv',<br>'accuracy',<br>'f1score', 'brier'] | TRUE  | No  | AUC-ROC  | Imbalanced |
| Classification of Central Venous Catheter Tip Position on Ches<br>X-ray Using Artificial Intelligence  | Pubmed<br>search | J Pers Med                        | 2022   | Jung S, Oh J, Ryu J, Kim J, Lee J, Cho Y, Yoon MS, Jeong JY.   | 10.3390/jpm12101637                    | Catheter position                                  | Categorical              | Private set   | Imbalance | Accuracy                    | recall, precision, recall, and F1-score  | No   | ['sens', 'vpp',<br>'sens', 'f1score']   | TRUE  | No  | Accuracy | Imbalanced |
| Deep learning-based classification for lung opacities in chest<br>x-ray radiographs through batch control and sensitivity<br>regulation  | Pubmed<br>search | Sci Rep                           | 2022 ( | Chang IY, Huang TY.  | 10.1038/s41598-022-2250<br>6-4         | Lung opacities                                     | Binary                   | ChestX-ray14  | Balance   | Sensitivity/Spe<br>cificity | e sensitivity, specificity, PPV,<br>negative predictive value<br>(NPV), accuracy, and F1<br>score          | No   | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy',<br>'f1score']                    | TRUE  | No  | Other    | Balanced   |
| Validation study of machine-learning chest radiograph software in primary and emergency medicine   | Pubmed<br>search | Clin Radiol                       | 2023 1 | van Beek EJR, Ahn JS, Kim MJ, Murchison JT.  | 10.1016/j.crad.2022.08.12<br>9         | Abnormality and multipathologies                   | Multiple binary tasks    | Private set   | Imbalance | AUC-ROC                     | sensitivity, specificity, accuracy   | No   | ['sens', 'spec',<br>'accuracy']   | FALSE | No  | AUC-ROC  | Imbalanced |
| Al-CenterNet CXR: An artificial intelligence (Al) enabled system for localization and classification of chest X-ray disease  | Pubmed<br>search | Front Med<br>(Lausanne)           | 2022   | Albahli S, Nazir T.  | 10.3389/fmed.2022.95576<br>5           |  | Multiple binary tasks    | ChestX-ray14  | Imbalance | AUC-ROC                     | -  | No   | ['vpp', 'sens',<br>'accuracy',<br>'f1score']                                      | TRUE  | No  | AUC-ROC  | Imbalanced |
| The effect of an artificial intelligence algorithm on chest X-ray interpretation of radiology residents  | Pubmed<br>search | Br J Radiol                       | 2022   | Pekçevîk Y, Orbatu D, Güngör F, Yıldırım O, Yaşar E, Yimer MA,<br>Şişman AR, Emiroğlu M, Dao L, Cohen JP, Sevinç S.  | 10.1259/bjr.20210688                   | Multiple findings                                  | Multiple binary tasks    | Private set   | Imbalance | AUC-ROC                     | sensitivity, specificity,<br>Cohen Kappa, precision,<br>NPV,   | No   | ['sens', 'spec',<br>'kappa', 'vpp',<br>'npv', "]                                  | TRUE  | No  | AUC-ROC  | Imbalanced |
| Association of Artificial Intelligence-Aided Chest Radiograph<br>Interpretation With Reader Performance and Efficiency   | Pubmed<br>search | JAMA Netw Open                    |        | Ahn JS, Ebrahimian S, McDermott S, Lee S, Naccarato L, Di Capua<br>JF, Wu MY, Zhang EW, Muse V, Miller B, Sabzalipour F, Bizzo BC,<br>Dreyer KJ, Kaviani P, Digumarthy SR, Kalra MK. | 10.1001/jamanetworkoper<br>.2022.29289 | Multiple findings                                  | Multiple binary tasks    | MIMIC-CXR + Private set   | Imbalance | AUC-ROC                     | sensitivity, specificiity  | No   | ['sens', 'spec']  | FALSE | No  | AUC-ROC  | Imbalanced |
| Development of deep learning chest X-ray model for cardiac dose prediction in left-sided breast cancer radiotherapy  | Pubmed<br>search | Sci Rep                           | 2022   | Koide Y, Aoyama T, Shimizu H, Kitagawa T, Miyauchi R, Tachibana  | 10.1038/s41598-022-1658<br>3-8         | Risk of cardiac<br>irradiation                     | Binary                   | Private set   | Balance   | AUC-ROC                     | sensitivity, specificity, PPV,<br>negative predictive value<br>(NPV), accuracy                             | No   | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy']                                  | TRUE  | No  | AUC-ROC  | Balanced   |
| A multichannel EfficientNet deep learning-based stacking<br>ensemble approach for lung disease detection using chest<br>X-ray images   | Pubmed<br>search | Cluster Comput                    | 2023   | Ravi V, Acharya V, Alazab M.   | 10.1007/s10586-022-0366<br>4-6         | Pneumonia,<br>tuberculosis and<br>COVID-19         | Multiple binary tasks    | Multiple public datasets  | Imbalance | Accuracy                    | sensitivity, specificity, PPV,<br>negative predictive value<br>(NPV), accuracy, and F1<br>score            | No   | ['sens', 'spec',<br>'ppv', 'npv',<br>'accuracy',<br>'f1score']                    | TRUE  | No  | Accuracy | Imbalanced |
| An efficient deep learning-based framework for tuberculosis detection using chest X-ray images   | Pubmed<br>search | Tuberculosis<br>(Edinb)           | 2022   | qbal A, Usman M, Ahmed Z.  | 10.1016/j.tube.2022.1022<br>34         | Tuberculosis                                       | Binary                   | Montgomery &<br>Shenzen,<br>ChestX-ray14,<br>Belarus              | Balance   | AUC-ROC                     | precision, recall, F1 score, accuracy  | Yes Monte Carlo<br>Dropout   | ['vpp', 'sens',<br>'f1score',<br>'accuracy']                                      | TRUE  | Yes | AUC-ROC  | Balanced   |
| Performance Evaluation of the Deep Learning Based<br>Convolutional Neural Network Approach for the Recognition of<br>Chest X-Ray Images  | Pubmed<br>search | Front Oncol                       | 2022   | Sharma S, Gupta S, Gupta D, Rashid J, Juneja S, Kim J, Elarabawy<br>MM.  | 10.3389/fonc.2022.93249<br>6           | Pneumonia  | Binary                   | ChestX-ray14  | Balance   | Accuracy                    | AUC-ROC, sensitivity,<br>specificity, cross entropy,<br>precision  | No   | ['auc-roc', 'sens',<br>'spec', 'cross<br>entropy', 'vpp']                         | FALSE | No  | Accuracy | Balanced   |
| Use data augmentation for a deep learning classification mode<br>with chest X-ray clinical imaging featuring coal workers'<br>pneumoconiosis   | Pubmed<br>search | BMC Pulm Med                      | 2022 I | Dong H, Zhu B, Zhang X, Kong X.  | 10.1186/s12890-022-0206<br>8-x         | Coal workers' pneumoconiosis                       | Binary                   | Private set   | Balance   | AUC-ROC                     | Accuracy, sensitivity,<br>specificity, F1 score,<br>precision  | No   | ['accuracy', 'sens',<br>'spec', 'f1score',<br>'vpp']                              | TRUE  | No  | AUC-ROC  | Balanced   |
| Deep Learning for Detection of Exercise-Induced Pulmonary<br>Hypertension Using Chest X-Ray Images   | Pubmed<br>search | Front Cardiovasc<br>Med           | 2022   | Kusunose K, Hirata Y, Yamaguchi N, Kosaka Y, Tsuji T, Kotoku J,<br>Sata M.   | 10.3389/fcvm.2022.89170<br>3           | Exercise-induced<br>pulmonary<br>hypertension      | Binary                   | Private set   | Balance   | AUC-ROC                     | precision, recall, f-score<br>values, and confusion<br>matrix  | No   | ['vpp', 'sens',<br>'f-score values',<br>'confusion matrix']                       | FALSE | No  | AUC-ROC  | Balanced   |
| Using Artificial Intelligence to Establish Chest X-Ray Image<br>Recognition Model to Assist Crucial Diagnosis in Elder Patients<br>With Dyspnea  | Pubmed<br>search | Front Med<br>(Lausanne)           | 2022 I | Liong-Rung L, Hung-Wen C, Ming-Yuan H, Shu-Tien H, Ming-Feng<br>T, Chia-Yu C, Kuo-Song C.  | 10.3389/fmed.2022.89320<br>8           | Pneumonia and<br>pulmonary edema                   | Multiple binary<br>tasks | Private set   | Imbalance | AUC-ROC                     | Accuracy, sensitivity,<br>specificity, F1 score,<br>precision  | No   | ['accuracy', 'sens',<br>'spec', 'f1score',<br>'vpp']                              | TRUE  | No  | AUC-ROC  | Imbalanced |
| The Contribution of Chest X-Ray to Predict Extubation Failure<br>in Mechanically Ventilated Patients Using Machine<br>Learning-Based Algorithms  | Pubmed<br>search | Crit Care Explor                  | 2022   | Fukuchi K, Osawa I, Satake S, Ito H, Shibata J, Dohi E, Kasugai D, Miyamoto Y, Ohbe H, Tamoto M, Yamada N, Yoshikawa K, Goto T.  | 10.1097/CCE.0000000000<br>000718       | Extubation failure                                 | Binary                   | MIMIC-IV  | Imbalance | AUC-ROC                     | AUC-PR, sensitivity,<br>specificity, PPV, NPV  | Yes Calibration curves   | ['auc-pr', 'sens',<br>'spec', 'ppv', 'npv']                                       | TRUE  | Yes | AUC-ROC  | Imbalanced |
| Diagnostic accuracy of a commercially available, deep<br>learning-based chest X-ray interpretation software for detecting<br>culture-confirmed pulmonary tuberculosis                        | Pubmed<br>search | Int J Infect Dis                  |        | Tavaziva G, Majidulla A, Nazish A, Saeed S, Benedetti A, Khan AJ, Ahmad Khan F.  | 10.1016/j.ijid.2022.05.037             | Tuberculosis                                       | Binary                   | Private set   | Imbalance | Sensitivity/Spe<br>cificity | AUC-ROC  | Yes Calibration of<br>threshold cut-or   | ['auc-roc']   | FALSE | Yes | Other    | Imbalanced |
| Early severity prediction of BPD for premature infants from<br>chest X-ray images using deep learning: A study at the 28th<br>day of oxygen inhalation                                       | Pubmed<br>search | Comput Methods<br>Programs Biomed |        | King W, He W, Li X, Chen J, Cao Y, Zhou W, Shen Q, Zhang X, Ta D.  | 69                                     | dysplasia  | Categorical              | Private set   | Balance   | Accuracy                    | specificity, and F1 score  | No   | ['vpp', 'sens',<br>'spec', 'f1score']   |       | No  | Accuracy | Balanced   |
| Tuberculosis detection in chest radiograph using convolutional<br>neural network architecture and explainable artificial<br>intelligence   | Pubmed<br>search | Neural Comput<br>Appl             | 2022   | Nafisah SI, Muhammad G.  | 10.1007/s00521-022-0725<br>8-6         | Tuberculosis                                       | Binary                   | Montgomery<br>Shenzhen<br>Belarus                                 | Balance   | AUC-ROC                     | Accuracy, recall, precision,<br>F1 score, Kappa  | No   | ['accuracy', 'sens',<br>'vpp', 'f1score',<br>'kappa']                             | TRUE  | No  | AUC-ROC  | Balanced   |
| Deep Learning in Multi-Class Lung Diseases' Classification on Chest X-ray Images   | Pubmed<br>search | Diagnostics (Basel)               | 2022   | Kim S, Rim B, Choi S, Lee A, Min S, Hong M.  | 10.3390/diagnostics12040<br>915        | Three classes of lung diseases                     | Categorical              | ChestX-ray14 &<br>Cheonan<br>Soonchunhyang<br>University Hospital | Balance   | Accuracy                    | Sensitivity, specificity, cross entropy  | No   | ['sens', 'spec',<br>'cross entropy']  | FALSE | No  | Accuracy | Balanced   |
| Deep learning-based automatic detection of tuberculosis disease in chest X-ray images  | Pubmed<br>search | Pol J Radiol                      | 2022   | Showkatian E, Salehi M, Ghaffari H, Reiazi R, Sadighi N.   | 10.5114/pjr.2022.113435                | Tuberculosis                                       | Binary                   | Montgomery and<br>Shenzhen  | Balance   | AUC-ROC                     | accuracy, sensitivity/recall, precision, F1-score.   | No   | ['accuracy', 'sens',<br>'sens', 'vpp',<br>'f1score']                              | TRUE  | No  | AUC-ROC  | Balanced   |
| Concordance rate of radiologists and a commercialized<br>deep-learning solution for chest X-ray: Real-world experience<br>with a multicenter health screening cohort                         | Pubmed<br>search | PLoS One                          |        | YJ.  | 10.1371/journal.pone.026<br>4383       |  | Categorical              | Private set   | Imbalance | Other                       | Concordance rate   | No   | ['concordance<br>rate']   | FALSE |     | Other    | Imbalanced |
| Detection of Left Ventricular Systolic Dysfunction Using an<br>Artificial Intelligence-Enabled Chest X-Ray   | Pubmed<br>search | Can J Cardiol                     | 2022   | Hsiang CW, Lin C, Liu WC, Lin CS, Chang WC, Hsu HH, Huang<br>3S, Lou YS, Lee CC, Wang CH, Fang WH.   | 10.1016/j.cjca.2021.12.01<br>9         | left ventricular ejection<br>fraction (LVEF) ≤ 35% | Binary                   | Private set   | Imbalance | AUC-ROC                     | Kaplan-Meier survival<br>analysis and the Cox<br>proportional hazards                                      | No   | ['kaplan-meier<br>survival analysis',<br>'the cox<br>proportional<br>hazards']    | FALSE | No  | AUC-ROC  | Imbalanced |

| A lightweight deep learning architecture for the automatic detection of pneumonia using chest X-ray images   | Pubmed<br>search | Multimed Tools<br>Appl                 | 2022 | Trivedi M, Gupta A.  | 10.1007/s11042-021-1180<br>7-x         | Pneumonia  | Binary                   | ChestX-ray14  | Imbalance | Accuracy | recall, specificity, precision,<br>F1 score, AUC-ROC   | No   | ['sens', 'spec',<br>'vpp', 'f1score',<br>'auc-roc']         | TRUE   | No  | Accuracy | Imbalanced |
|--|------------------|--|------|--|--|--|--------------------------|---|-----------|----------|--|--|---|--------|-----|----------|------------|
| An Artificial Intelligence-Based Chest X-ray Model on Human<br>Nodule Detection Accuracy From a Multicenter Study  | Pubmed<br>search | JAMA Netw Open                         |      | Homayounieh F, Digumarthy S, Ebrahimian S, Rueckel J, Hoppe<br>BF, Sabel BO, Conjeti S, Ridder K, Sistermanns M, Mang L, Preuhs<br>A, Glesu F, Mansoor A, Moghel M, Botwin A, Singh R, Cartmell S,<br>Patti J, Huemmer C, Fieselmann A, Joerger C, Mirshahzadeh N,<br>Muse V, Kalfa C, | 10.1001/jamanetworkopen<br>.2021.41096 | Multiple findings                                | Multiple binary tasks    | Private set   | Balance   | AUC-ROC  | Sensitivity, Specificity, and Accuracy   | No   | ['sens', 'spec', 'accuracy']                                | FALSE  | No  | AUC-ROC  | Balanced   |
| Evaluation of an artificial intelligence (AI) system to detect tuberculosis on chest X-ray at a pilot active screening project in Guangdong, China in 2019                               | Pubmed<br>search | J Xray Sci Technol                     | 2022 | Liao Q, Feng H, Li Y, Lai X, Pan J, Zhou F, Zhou L, Chen L.  | 10.3233/XST-211019                     | Tuberculosis                                     | Binary                   | Private set   | Imbalance | AUC-ROC  | specificity, sensitivity, and positive predict value   | No   | ['spec', 'sens',<br>'ppv']                                  | TRUE   | No  | AUC-ROC  | Imbalanced |
| Deep Learning and Binary Relevance Classification of Multiple Diseases using Chest X-Ray images  | Pubmed<br>search | Annu Int Conf IEEE<br>Eng Med Biol Soc | 2021 | Blais MA, Akhloufi MA.   | 10.1109/EMBC46164.202<br>1.9629846     | 5 findings                                       | Multiple binary<br>tasks | CheXpert  | Imbalance | AUC-ROC  | -  | No   | m   | FALSE  | No  | AUC-ROC  | Imbalanced |
| Using Transfer Learning Method to Develop an Artificial<br>Intelligence Assisted Triaging for Endotracheal Tube Position<br>on Chest X-ray   | Pubmed<br>search | Diagnostics (Basel)                    | 2021 | Yuan KC, Tsai LW, Lai KS, Teng ST, Lo YS, Peng SJ.   | 10.3390/diagnostics11101<br>844        | Endotracheal Tube<br>Position                    | Binary                   | Private set   | Balance   | AUC-ROC  | accuracy, recall, precision,<br>f1 score, cross entropy  | No   | ['accuracy', 'sens<br>'vpp', 'f1score',<br>'cross entropy'] | , TRUE | No  | AUC-ROC  | Balanced   |
| Pneumonia detection in chest X-ray images using an ensemble of deep learning models  | Pubmed search    | PLoS One                               | 2021 | Kundu R, Das R, Geem ZW, Han GT, Sarkar R.   | 10.1371/journal.pone.025<br>6630       | Pneumonia  | Binary                   | ChestX-ray14 &<br>Kermany et al   | Imbalance | Accuracy | AUC ROC, recall,<br>precision, f1 score  | No   | ['auc-roc', 'sens',<br>'vpp', 'f1score']                    | TRUE   | No  | Accuracy | Imbalanced |
| Chest X-ray Analysis With Deep Learning-Based Software as a<br>Triage Test for Pulmonary Tuberculosis: An Individual Patient<br>Data Meta-Analysis of Diagnostic Accuracy                | Pubmed<br>search | Clin Infect Dis                        |      | Tavaziva G, Harris M, Abidi SK, Geric C, Breuninger M, Dheda K,<br>Esmail A, Muyoyeta M, Reither K, Majidulla A, Khan AJ, Campbell<br>JR, David PM, Denkinger C, Miller C, Nathavitharana R, Pai M,<br>Benedetti A, Ahmad Khan F.  | 10.1093/cid/ciab639                    | Tuberculosis                                     | Binary                   | Private set   | Imbalance | AUC-ROC  | Sensitivity and specificity  | No   | ['sens', 'spec']  | FALSE  | No  | AUC-ROC  | Imbalanced |
| Effect of a comprehensive deep-learning model on the<br>accuracy of chest x-ray interpretation by radiologists: a<br>retrospective, multireader multicase study                          | Pubmed<br>search | Lancet Digit Health                    |      | Seah JCY, Tang CHM, Buchlak QD, Holt XG, Wardman JB,<br>Almoldin A, Esmaili N, Ahmad H, Pham H, Lambur JF, Bachey B,<br>Hogg SJF, Johnston BP, Bennett C, Oakden-Rayner L, Brotchie P,<br>Jones CM.  | 10.1016/S2589-7500(21)0<br>0106-0      | 127 clinical findings                            | Multiple binary tasks    | I-MED Radiology<br>Network<br>MIMIC-CXR<br>ChestX-ray14<br>CheXpert<br>Padchest | Imbalance | AUC-ROC  | Positive predictive value,<br>sensitivity, specificity, and<br>Matthews correlation<br>coefficient | No   | ['ppv', 'sens',<br>'spec', 'mcc']                           | TRUE   | No  | AUC-ROC  | Imbalanced |
| Chest x-ray automated triage: A semiologic approach designed for clinical implementation, exploiting different types of labels through a combination of four Deep Learning architectures | Pubmed<br>search | Comput Methods<br>Programs Biomed      |      | Mosquera C, Diaz FN, Binder F, Rabellino JM, Benitez SE,<br>Beresñak AD, Seehaus A, Ducrey G, Ocantos JA, Luna DR.   | 10.1016/j.cmpb.2021.1061<br>30         | Abnormality + 4 findings                         | Multiple binary tasks    | Private set &<br>ChestX-ray14   | Balance   | AUC-ROC  | Sensitivity, specificity, PPV, NPV   | No   | ['sens', 'spec',<br>'ppv', 'npv']                           | TRUE   | No  | AUC-ROC  | Balanced   |
| Combining Deep Learning and Knowledge-driven Reasoning for Chest X-Ray Findings Detection  | Pubmed search    | AMIA Annu Symp<br>Proc                 | 2021 | Jadhav A, Wong KCL, Wu JT, Moradi M, Syeda-Mahmood T.  |  | Multiple findings                                | Multiple binary tasks    | MIMIC   | Imbalance | Other    | Precision Recall<br>F1-Score   | No   | ['vpp', 'sens',<br>'f1score', '']                           | TRUE   | No  | Other    | Imbalanced |
| Deep-Learning-Based Diagnosis of Bedside Chest X-ray in<br>Intensive Care and Emergency Medicine   | Pubmed search    | Invest Radiol                          |      | Niehues SM, Adams LC, Gaudin RA, Erxleben C, Keller S,<br>Makowski MR, Vahldiek JL, Bressem KK.  | 10.1097/RLI.00000000000<br>00771       | 8 findings                                       | Multiple binary tasks    | Private set   | Imbalance | AUC-ROC  | sensitivity, specificity, and positive predictive value.   | No   | ['sens', 'spec',<br>'ppv']                                  | TRUE   | No  | AUC-ROC  | Imbalanced |
| Deep Learning for Detection of Elevated Pulmonary Artery<br>Wedge Pressure Using Standard Chest X-Ray  | Pubmed<br>search | Can J Cardiol                          | 2021 | Hirata Y, Kusunose K, Tsuji T, Fujimori K, Kotoku J, Sata M.   | 10.1016/j.cjca.2021.02.00<br>7         | elevated pulmonary<br>arterial wedge<br>pressure | Binary                   | Private set   | Imbalance | AUC-ROC  | Cohen kappa  | No   | [ˈkappaˈ]   | FALSE  | No  | AUC-ROC  | Imbalanced |
| Recalibration of deep learning models for abnormality detection in smartphone-captured chest radiograph  | Pubmed<br>search | NPJ Digit Med                          |      | Kuo PC, Tsai CC, López DM, Karargyris A, Pollard TJ, Johnson<br>AEW, Celi LA.  | 10.1038/s41746-021-0039<br>3-9         | Multiple findings                                | Multiple binary tasks    | MIMIC-CXR and<br>CheXpert   | Imbalance | AUC-ROC  | sensitivity, specificity, f1<br>score, and accuracy  | No They use<br>term cali<br>to refer t<br>domain<br>adaptati | oation 'f1score',<br>'accuracy']                            | TRUE   | Yes | AUC-ROC  | Imbalanced |