Current Diagnostic Models Utilize Only One Data Type

- Inadequate Utilization of EHR Data Systems
- Barriers to Accessing Complex Insights For Physicians
- Limitations in AI for Comprehensive Diagnostics
Introducing MedFusion Analytics – The Future of Diagnostic Precision

A pioneering multi-modal model to predict top pathological findings in chest X-rays

- **Combines** patient data, clinician notes, and radiology images
- **Tailored** for healthcare researchers and attending physicians
- **Empowers** users to harness AI-based diagnostic insights
- **Easy-to-use** tool allows for seamless integration
Data Processing Pipeline

- **MIMIC-IV**
  - Tabular Data (Patient data, vitals, etc.)
  - TextData (Clinical Notes)
  - Image Data (Chest X-rays)

  **Data Preprocessing**
  - Remove outliers, remove null values, change data types, etc.
  - Extract section, remove outliers, additional NLP handling, etc.
  - Scaling, normalization, remove outlier, etc.

  **Data Filtering**
  - Studies conducted in the ED
  - Studies with a single label (finding)
  - Selection of top 4 findings (atelectasis, cardiomegaly, lung opacity, pleural effusion)

  **Data Split**
  - Train
  - Validation
  - Test
Individual Model Performance

Tabular Model (XGBoost)

Text Model (Bio_ClinicalBERT)

Image Model (EfficientNet-B3)
Revolutionizing Diagnostics with Multi-Model Integration

**Late Fusion**

Most multi-modal models today are good at dealing with missing data. Aggregation functions need to be empirically determined.

**Early Fusion**

A final model determines the final prediction. Able to model interactions between modalities. Potential for better performance.
MedFusion Analytics Web Demo

https://uc-berkeley-i-school.github.io/mids-210-medfusion-analytics-spring24/
Table inference

Result

Pre-existing evidence in ‘being homeless’ with a source described as 57%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%

Pre-existing evidence in ‘being education’ with a source described as 67%
Early Fusion Model Pipeline - From Data to Diagnosis

Data Preprocessing

- MIMIC-IV

Individual Model Development

- Tabular Data (Patient data, vitals, etc.)
- Text Data (Clinical Notes)
- Image Data (Chest X-rays)

Early Fusion Approach

- Fine-Tuned XGBoost
  - Classifying:
    - Atelectasis
    - Cardiomegaly
    - Lung Opacity
    - Pleural Effusion
    - No Findings

- 18 fusion features

- Tabular Data: 8 features
- Text Data: 5 features (final softmax outputs)
- Image Data: 5 features (final softmax outputs)
Early Fusion Outperforms Individual Models

Atelectasis
- Early Fusion: 0.7319
- Image: 0.7361
- Notes: 0.5042
- Tabular: 0.5184

Cardiomegaly
- Early Fusion: 0.8350
- Image: 0.8080
- Notes: 0.6811
- Tabular: 0.6140

Lung Opacity
- Early Fusion: 0.7713
- Image: 0.7450
- Notes: 0.5994
- Tabular: 0.6204

Pleural Effusion
- Early Fusion: 0.8705
- Image: 0.8548
- Notes: 0.7654
- Tabular: 0.6164
Redefining Standards with Early Fusion

Compared to an average late fusion aggregator, our early fusion model has:

- Superior AUC
- Sharper Detection
- Robust Across Pathologies
- Sets New Norms
Early Fusion Approach: Our pioneering multi-modal model integrates patient data, clinician notes, and radiology images for accurate diagnosis.

Key Benefits:

- Real-time predictions in seconds
- Intuitive user interface
- Comprehensive documentation to promote model transparency
Our Mission: Revolutionize patient care by harnessing the power of multi-modal data

Our Mission Pillars

- Pioneering Novelty
- Precision & Personalization
- Empowering Physicians
- Enhancing Patient Outcomes
- Shaping the Future of Healthcare
References


Appendix
“[current clinical decision support systems] increased the workload ... more steps, discrete ... instead of writing a free-text reason for your study”

MARC KOHLI, M.D.
Radiologist, Professor of Radiology, UCSF

“...we need a decision support tool to provide an opportunity to cut down on time required for physicians to make a diagnosis... as a consequence we improve the patient experience...”

JOSEPH NGUYEN, M.D.
Radiologist, Synergy Radiology Associates

“...our problem is resources... radiologists may take 2 to 6 hours to return a reading...”

MOHIT BANSAL, M.D.
Family Physician, Lifeline Urgent Care

“Image analysis is very hot in radiology right now ... [it takes] long time until you incorporate the rest of the things in the chart...”

RONALD CRANDALL, M.D.
Radiology Resident, Richmond University Medical Center
EDA: Combining Notes, CXR (Images) with other MIMIC Modules

<table>
<thead>
<tr>
<th>source</th>
<th>total_count</th>
<th>notes_overlap</th>
<th>notes_proportion_of</th>
<th>cxr_overlap</th>
<th>cxr_proportion_of_table</th>
<th>both_overlap</th>
<th>both_proportion_of_table</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIMICIV_Derived_Age</td>
<td>180733</td>
<td>145914</td>
<td>0.807345642467...</td>
<td>51299</td>
<td>0.283838590628...</td>
<td>45935</td>
<td>0.2541594506813919</td>
</tr>
<tr>
<td>MIMICIV_ECG</td>
<td>161352</td>
<td>111647</td>
<td>0.691946799543...</td>
<td>54362</td>
<td>0.336915563488...</td>
<td>42511</td>
<td>0.263467450047102</td>
</tr>
<tr>
<td>MIMICIV_ED_Edstay</td>
<td>205504</td>
<td>100470</td>
<td>0.488895593273...</td>
<td>61856</td>
<td>0.300996574275...</td>
<td>45922</td>
<td>0.22346037060105886</td>
</tr>
<tr>
<td>MIMICIV_Hosp_Admissions</td>
<td>180733</td>
<td>145914</td>
<td>0.807345642467...</td>
<td>51299</td>
<td>0.283838590628...</td>
<td>45935</td>
<td>0.2541594506813919</td>
</tr>
<tr>
<td>MIMICIV_ICU_Edstay</td>
<td>50920</td>
<td>50496</td>
<td>0.991673212882...</td>
<td>19264</td>
<td>0.378318931657...</td>
<td>19227</td>
<td>0.37759230164964652</td>
</tr>
<tr>
<td>MIMICIV_Note_Discharge</td>
<td>145914</td>
<td>145914</td>
<td>1.0</td>
<td>45935</td>
<td>0.314808722946...</td>
<td>45935</td>
<td>0.31480872294639306</td>
</tr>
<tr>
<td>MIMIC_CXR</td>
<td>65379</td>
<td>45935</td>
<td>0.702595634683...</td>
<td>65379</td>
<td>1.0</td>
<td>45935</td>
<td>0.70259563468392</td>
</tr>
</tbody>
</table>
Data Processing - MIMIC-IV Modules

```
mimiciv_note.discharge
subject_id INTEGER NN
hadm_id INTEGER NN
text VARCHAR

mimiciv_car.study
subject_id INTEGER NN
study_id INTEGER NN
study_datetime TIMESTAMP
path VARCHAR

mimiciv_car_jpg.mimic-car-2.0.0-chcipert
subject_id INTEGER NN
study_id INTEGER NN
atelectasis VARCHAR
cardiomegaly VARCHAR
consolidation VARCHAR
edema VARCHAR
enlarged_cardiomegaly VARCHAR
fracture VARCHAR
lung_lesion VARCHAR
lung_opacity VARCHAR
no_finding VARCHAR
pleural_lesion VARCHAR
pleural_peritumoral VARCHAR
pleural_other VARCHAR
pneumonia VARCHAR
pneumothorax VARCHAR
support_devices VARCHAR

mimiciv_hosp.transfers
subject_id INTEGER NN
hadm_id INTEGER
transfer_id INTEGER NN
event_type VARCHAR
encounterid VARCHAR
inime TIMESTAMP
outtime TIMESTAMP

mimiciv_ed.triage
subject_id INTEGER NN
stay_id INTEGER NN
temperature INTEGER
heart_rate INTEGER
respirate INTEGER
o2sat INTEGER
sbp INTEGER
dbp INTEGER
pain INTEGER
acuity VARCHAR
chief_complaint VARCHAR
```
Data Processing – Pipeline
<table>
<thead>
<tr>
<th>Finding</th>
<th>Train Dataset</th>
<th>Validation Dataset</th>
<th>Test Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>train_set_chexpert_4_findings_single_label_unbalanced.json</td>
<td>validation_set_chexpert_4_findings_single_label_unbalanced.json</td>
<td>test_set_chexpert_4_findings_single_label_unbalanced.json</td>
</tr>
<tr>
<td>no_finding</td>
<td>2,228</td>
<td>557</td>
<td>557</td>
</tr>
<tr>
<td>lung_opacity</td>
<td>6,865</td>
<td>1,716</td>
<td>1,715</td>
</tr>
<tr>
<td>cardiomegaly</td>
<td>7,242</td>
<td>1,810</td>
<td>1,809</td>
</tr>
<tr>
<td>atelectasis</td>
<td>7,292</td>
<td>1,822</td>
<td>1,822</td>
</tr>
<tr>
<td>pleural_effusion</td>
<td>7,405</td>
<td>1,851</td>
<td>1,849</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Finding</th>
<th>Train Dataset - Balanced</th>
<th>Validation Dataset - Balanced</th>
<th>Test Dataset - Balanced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>train_set_chexpert_4_findings_single_label_balanced.json</td>
<td>validation_set_chexpert_4_findings_single_label_balanced.json</td>
<td>test_set_chexpert_4_findings_single_label_balanced.json</td>
</tr>
<tr>
<td>no_finding</td>
<td>1,412</td>
<td>1,412</td>
<td>1,412</td>
</tr>
<tr>
<td>lung_opacity</td>
<td>1,765</td>
<td>1,765</td>
<td>1,765</td>
</tr>
<tr>
<td>cardiomegaly</td>
<td>1,765</td>
<td>1,765</td>
<td>1,765</td>
</tr>
<tr>
<td>atelectasis</td>
<td>1,765</td>
<td>1,765</td>
<td>1,765</td>
</tr>
<tr>
<td>pleural_effusion</td>
<td>1,765</td>
<td>1,765</td>
<td>1,765</td>
</tr>
</tbody>
</table>

Data Processing – Split
Data Processing – Pathology Selection

Multiple combinations (multilabel)

Positive mentions for the top 4 findings/pathologies after cleaning the data (they total 64 combinations)

Look for the top 4 findings/pathologies

<table>
<thead>
<tr>
<th>pathology</th>
<th>not_mention</th>
<th>positive_mention</th>
</tr>
</thead>
<tbody>
<tr>
<td>no_finding</td>
<td>152372</td>
<td>75455</td>
</tr>
<tr>
<td>support_devices</td>
<td>157783</td>
<td>66558</td>
</tr>
<tr>
<td>pleural_effusion</td>
<td>146369</td>
<td>54300</td>
</tr>
<tr>
<td>lung_opacity</td>
<td>173233</td>
<td>51525</td>
</tr>
<tr>
<td>atelectasis</td>
<td>180488</td>
<td>48908</td>
</tr>
<tr>
<td>cardiomegaly</td>
<td>187071</td>
<td>44845</td>
</tr>
<tr>
<td>edema</td>
<td>175168</td>
<td>27018</td>
</tr>
<tr>
<td>pneumonia</td>
<td>186933</td>
<td>16556</td>
</tr>
<tr>
<td>consolidation</td>
<td>239082</td>
<td>10778</td>
</tr>
<tr>
<td>pneumothorax</td>
<td>157113</td>
<td>10358</td>
</tr>
<tr>
<td>enlarged_cardiomediastinum</td>
<td>215365</td>
<td>7179</td>
</tr>
<tr>
<td>lung_lesion</td>
<td>220681</td>
<td>6294</td>
</tr>
<tr>
<td>fracture</td>
<td>222551</td>
<td>4390</td>
</tr>
<tr>
<td>pleural_other</td>
<td>225690</td>
<td>2011</td>
</tr>
</tbody>
</table>

1726 rows x 3 columns
Problem: Data leakage (some notes contain the pathology/finding associated with the x-ray)

Solution: Extraction of paragraphs that don’t contain explanation of results from the ED (via algorithm)
Data Processing – Images

Key challenges with imaging data:

- Patients can have more than 1 X-ray
- X-rays can be taken from multiple angles and positions
- Images are of varying sizes

Progress on imaging data:

- Found positioning metadata connected with each image to identify the orientation of the image systematically
- Developed a logic to select which image we will keep from each patient
Data Processing – Image Padding

Originals

Padded & Reduced
Logic for reducing/selecting which image to keep from each patient:

1. For each patient, if # of images per study > 1: [Main criteria for reducing all images per patient down to 1 single image]
2. Non-lateral views are preferred* (from SME conversations)
3. Exclude 'Recumbent' orientation wherever possible
4. Prefer images with larger 'Rows' pixels if orientations vary
5. Latest 'StudyTime' if times vary
6. Remove record with NaN in meta data for two images with similar other meta data
7. Remove record with lower 'Columns' if column pixels is the only difference in the meta data between 2 images
8. Preference to 'antero-posterior' view over 'posterior-antero' if this is the only difference in the meta data
## NLP Models – Top Performers

<table>
<thead>
<tr>
<th>Model</th>
<th>Train Data</th>
<th>Class Weights</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio_ClinicalBERT</td>
<td>Unbalanced</td>
<td>No</td>
<td>0.66531</td>
</tr>
<tr>
<td>Bio_ClinicalBERT</td>
<td>Unbalanced</td>
<td>Yes</td>
<td>0.65781</td>
</tr>
<tr>
<td>Bio_ClinicalBERT</td>
<td>Balanced</td>
<td>No</td>
<td>0.66177</td>
</tr>
<tr>
<td>Bio_Discharge_Summary_BERT</td>
<td>Unbalanced</td>
<td>No</td>
<td>0.61101</td>
</tr>
<tr>
<td>Bio_Discharge_Summary_BERT</td>
<td>Unbalanced</td>
<td>Yes</td>
<td>0.51912</td>
</tr>
<tr>
<td>Bio_Discharge_Summary_BERT</td>
<td>Balanced</td>
<td>No</td>
<td>0.65772</td>
</tr>
<tr>
<td>BioBERT</td>
<td>Unbalanced</td>
<td>No</td>
<td>0.60079</td>
</tr>
<tr>
<td>BioBERT</td>
<td>Unbalanced</td>
<td>Yes</td>
<td>0.52576</td>
</tr>
<tr>
<td>BioBERT</td>
<td>Balanced</td>
<td>No</td>
<td>0.67084</td>
</tr>
</tbody>
</table>
Top NLP Model (Bio_ClinicalBERT)

<table>
<thead>
<tr>
<th></th>
<th>Unbalanced</th>
<th>Unbalanced with class weights</th>
<th>Balanced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>precision</td>
<td>recall</td>
<td>f1-score</td>
</tr>
<tr>
<td>no_finding</td>
<td>0.77</td>
<td>0.99</td>
<td>0.82</td>
</tr>
<tr>
<td>atelectasis</td>
<td>0.15</td>
<td>0.94</td>
<td>0.87</td>
</tr>
<tr>
<td>cardiomegaly</td>
<td>0.27</td>
<td>0.70</td>
<td>0.23</td>
</tr>
<tr>
<td>lung opacity</td>
<td>0.35</td>
<td>0.27</td>
<td>0.30</td>
</tr>
<tr>
<td>pleural effusion</td>
<td>0.35</td>
<td>0.19</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>accuracy</th>
<th>macro avg</th>
<th>weighted avg</th>
<th>accuracy</th>
<th>macro avg</th>
<th>weighted avg</th>
<th>accuracy</th>
<th>macro avg</th>
<th>weighted avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>no_finding</td>
<td>0.69</td>
<td>0.38</td>
<td>0.63</td>
<td>0.30</td>
<td>0.37</td>
<td>0.63</td>
<td>0.29</td>
<td>0.36</td>
<td>0.63</td>
</tr>
<tr>
<td>atelectasis</td>
<td>0.42</td>
<td>0.30</td>
<td>0.63</td>
<td>0.24</td>
<td>0.30</td>
<td>0.63</td>
<td>0.24</td>
<td>0.30</td>
<td>0.63</td>
</tr>
<tr>
<td>cardiomegaly</td>
<td>0.34</td>
<td>0.20</td>
<td>0.34</td>
<td>0.21</td>
<td>0.14</td>
<td>0.34</td>
<td>0.21</td>
<td>0.14</td>
<td>0.34</td>
</tr>
<tr>
<td>lung opacity</td>
<td>0.55</td>
<td>0.03</td>
<td>0.55</td>
<td>0.43</td>
<td>0.15</td>
<td>0.43</td>
<td>0.43</td>
<td>0.15</td>
<td>0.43</td>
</tr>
<tr>
<td>pleural effusion</td>
<td>0.55</td>
<td>0.03</td>
<td>0.55</td>
<td>0.43</td>
<td>0.15</td>
<td>0.43</td>
<td>0.43</td>
<td>0.15</td>
<td>0.43</td>
</tr>
</tbody>
</table>
### Top Imaging Model: EfficientNet-B3 Results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atelectasis</td>
<td>0.79</td>
<td>0.69</td>
<td>0.74</td>
<td><strong>0.83</strong></td>
</tr>
<tr>
<td>Cardiomegaly</td>
<td>0.89</td>
<td>0.78</td>
<td>0.83</td>
<td><strong>0.88</strong></td>
</tr>
<tr>
<td>Lung Opacity</td>
<td>0.76</td>
<td>0.70</td>
<td>0.73</td>
<td><strong>0.82</strong></td>
</tr>
<tr>
<td>Pleural Effusion</td>
<td>0.92</td>
<td>0.71</td>
<td>0.80</td>
<td><strong>0.85</strong></td>
</tr>
<tr>
<td>No Findings</td>
<td>0.86</td>
<td>0.76</td>
<td>0.80</td>
<td><strong>0.83</strong></td>
</tr>
</tbody>
</table>

Other computer vision models assessed:
- RestNet18
- DenseNet121
- Various custom CNNs
Individual Model Performance

**Tabular Model**
- XGBoost

**Notes Model**
- Bio_ClinicalBERT

**Images Model**
- EfficientNet-B3
Fusion Model Comparisons – Additional Figures
Average AUC

- **Individual Models vs MedFusion Analytics**

- **Research Summary**
  - **Case** | **Model** | **estimator** | **basis** | **data** | **Average** | **no findings** | **atelectasis** | **cardiomegaly** | **lung opacity** | **pleural effusion**
  - 1 | Tabular | XGBoost | train | trainset | 0.6617 | 0.6655 | 0.6497 | 0.6777 | 0.6641 | 0.6469
  - 2 | Notes1 | Bert | train | trainset | 0.8344 | 0.8040 | 0.7226 | 0.8652 | 0.8556 | 0.9257
  - 3 | Image2 | EfficientNet | train | large modified trainset | 0.9964 | 0.9962 | 0.9918 | 0.9920 | 0.9964 | 0.9993
  - 13 | EF1a: Tabular + Notes1 + Image2 | XGBoost | train | trainset + valset | 0.9635 | 0.9526 | 0.9323 | 0.9758 | 0.9580 | 0.9782

**Notes:**
1. All data are for 4-balanced (4 pathologies, balanced data, and single pathology per image).
2. All models (except Image model) used the same following datasets:
   - trainset: 2118
   - valset: 1940
   - testset: 1990
### Research Summary

<table>
<thead>
<tr>
<th>Case</th>
<th>Model</th>
<th>assertion</th>
<th>bias</th>
<th>data</th>
<th>Average</th>
<th>No. Patients</th>
<th>sensitivity</th>
<th>specificity</th>
<th>accuracy</th>
<th>E[call]</th>
<th>ROC AUC Score</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tabular</td>
<td>KBtarget</td>
<td>train</td>
<td>male</td>
<td>0.667</td>
<td>0.637</td>
<td>0.926</td>
<td>0.677</td>
<td>0.151</td>
<td>0.068</td>
<td>0.865</td>
<td>0.595</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.673</td>
<td>0.647</td>
<td>0.925</td>
<td>0.682</td>
<td>0.155</td>
<td>0.069</td>
<td>0.863</td>
<td>0.598</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Notes 1</td>
<td>GNN</td>
<td>train</td>
<td>male</td>
<td>0.684</td>
<td>0.654</td>
<td>0.930</td>
<td>0.685</td>
<td>0.157</td>
<td>0.071</td>
<td>0.864</td>
<td>0.599</td>
<td>0.107</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.688</td>
<td>0.658</td>
<td>0.931</td>
<td>0.688</td>
<td>0.159</td>
<td>0.072</td>
<td>0.865</td>
<td>0.601</td>
<td>0.107</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Notes 2</td>
<td>GNN</td>
<td>train</td>
<td>male</td>
<td>0.691</td>
<td>0.663</td>
<td>0.936</td>
<td>0.696</td>
<td>0.161</td>
<td>0.074</td>
<td>0.869</td>
<td>0.605</td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.695</td>
<td>0.667</td>
<td>0.938</td>
<td>0.698</td>
<td>0.164</td>
<td>0.076</td>
<td>0.871</td>
<td>0.610</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Image 1</td>
<td>CNN</td>
<td>train</td>
<td>male</td>
<td>0.462</td>
<td>0.433</td>
<td>0.916</td>
<td>0.393</td>
<td>0.159</td>
<td>0.072</td>
<td>0.788</td>
<td>0.585</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.502</td>
<td>0.473</td>
<td>0.925</td>
<td>0.403</td>
<td>0.164</td>
<td>0.078</td>
<td>0.800</td>
<td>0.595</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Image 2</td>
<td>CNN</td>
<td>train</td>
<td>male</td>
<td>0.508</td>
<td>0.482</td>
<td>0.928</td>
<td>0.523</td>
<td>0.170</td>
<td>0.082</td>
<td>0.800</td>
<td>0.595</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.512</td>
<td>0.486</td>
<td>0.931</td>
<td>0.526</td>
<td>0.172</td>
<td>0.083</td>
<td>0.803</td>
<td>0.598</td>
<td>0.114</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Tabular + Image (CNN)</td>
<td>KBtarget</td>
<td>train</td>
<td>male</td>
<td>0.456</td>
<td>0.431</td>
<td>0.916</td>
<td>0.406</td>
<td>0.160</td>
<td>0.073</td>
<td>0.790</td>
<td>0.583</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.484</td>
<td>0.456</td>
<td>0.925</td>
<td>0.420</td>
<td>0.164</td>
<td>0.077</td>
<td>0.801</td>
<td>0.595</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Tabular + Notes 1</td>
<td>KBtarget</td>
<td>train</td>
<td>male</td>
<td>0.670</td>
<td>0.641</td>
<td>0.928</td>
<td>0.680</td>
<td>0.157</td>
<td>0.070</td>
<td>0.863</td>
<td>0.600</td>
<td>0.108</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.674</td>
<td>0.645</td>
<td>0.929</td>
<td>0.684</td>
<td>0.160</td>
<td>0.072</td>
<td>0.866</td>
<td>0.604</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Tabular + Notes 2</td>
<td>KBtarget</td>
<td>train</td>
<td>male</td>
<td>0.688</td>
<td>0.658</td>
<td>0.936</td>
<td>0.695</td>
<td>0.161</td>
<td>0.074</td>
<td>0.870</td>
<td>0.610</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.692</td>
<td>0.662</td>
<td>0.938</td>
<td>0.698</td>
<td>0.164</td>
<td>0.076</td>
<td>0.873</td>
<td>0.610</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Late Fusion</td>
<td>-</td>
<td>train</td>
<td>male</td>
<td>0.624</td>
<td>0.591</td>
<td>0.916</td>
<td>0.518</td>
<td>0.154</td>
<td>0.072</td>
<td>0.794</td>
<td>0.587</td>
<td>0.108</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.628</td>
<td>0.595</td>
<td>0.917</td>
<td>0.522</td>
<td>0.157</td>
<td>0.073</td>
<td>0.800</td>
<td>0.595</td>
<td>0.112</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>EF to Tabular + Notes 1</td>
<td>Image 2</td>
<td>train</td>
<td>male</td>
<td>0.803</td>
<td>0.764</td>
<td>0.935</td>
<td>0.756</td>
<td>0.165</td>
<td>0.079</td>
<td>0.900</td>
<td>0.637</td>
<td>0.118</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.807</td>
<td>0.768</td>
<td>0.937</td>
<td>0.760</td>
<td>0.166</td>
<td>0.080</td>
<td>0.902</td>
<td>0.640</td>
<td>0.120</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>EF to Tabular + Notes 2</td>
<td>Image 2</td>
<td>train</td>
<td>male</td>
<td>0.925</td>
<td>0.885</td>
<td>0.948</td>
<td>0.905</td>
<td>0.175</td>
<td>0.081</td>
<td>0.956</td>
<td>0.693</td>
<td>0.128</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.929</td>
<td>0.890</td>
<td>0.951</td>
<td>0.907</td>
<td>0.177</td>
<td>0.083</td>
<td>0.959</td>
<td>0.695</td>
<td>0.130</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>EF to Tabular + Notes 1</td>
<td>Image 3</td>
<td>train</td>
<td>male</td>
<td>0.780</td>
<td>0.741</td>
<td>0.926</td>
<td>0.736</td>
<td>0.158</td>
<td>0.074</td>
<td>0.887</td>
<td>0.597</td>
<td>0.107</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.784</td>
<td>0.745</td>
<td>0.928</td>
<td>0.740</td>
<td>0.160</td>
<td>0.076</td>
<td>0.890</td>
<td>0.601</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>EF to Tabular + Notes 2</td>
<td>Image 3</td>
<td>train</td>
<td>male</td>
<td>0.805</td>
<td>0.766</td>
<td>0.931</td>
<td>0.759</td>
<td>0.164</td>
<td>0.079</td>
<td>0.900</td>
<td>0.637</td>
<td>0.118</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.809</td>
<td>0.770</td>
<td>0.933</td>
<td>0.763</td>
<td>0.166</td>
<td>0.080</td>
<td>0.903</td>
<td>0.640</td>
<td>0.120</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Image 3</td>
<td>-</td>
<td>train</td>
<td>male</td>
<td>0.813</td>
<td>0.775</td>
<td>0.935</td>
<td>0.780</td>
<td>0.167</td>
<td>0.081</td>
<td>0.905</td>
<td>0.642</td>
<td>0.121</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>test</td>
<td>male</td>
<td>0.817</td>
<td>0.780</td>
<td>0.936</td>
<td>0.785</td>
<td>0.169</td>
<td>0.082</td>
<td>0.908</td>
<td>0.645</td>
<td>0.123</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
1. All data sets contain the LIDC data, but single pathology data sets were used.
2. All models except Image 3 used the same following datasets:
   - training: 2700
   - validation: 1040
   - test: 960
3. Each hyperparameter set is different for different cases.