Poster

Deep learning improves sensitivity of UIP classification on CT

American Thoracic Society (ATS) International Conference

May 13-18, 2022

SC-US-74373

P

Deep learning improves sensitivity of UIP classification on CT

Stephen M. Humphries PhD [1]
D. Thieke[1], A. Notary[1], A. Oh[1], J. H. Chung[2], M. E. Strek[3], A. L. Olson[4], N. M. Patel[4], D. A. Lynch[1]

1. Dept. of Radiology, National Jewish Health, Denver, CO
2. Dept. of Radiology, University of Chicago, Chicago, IL
3. Dept. of Medicine, University of Chicago, Chicago, IL
4. Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT

Introduction

- In the correct clinical context, confident identification of a usual interstitial pneumonia (UIP) pattern on computed tomography (CT) is sufficient to diagnose idiopathic pulmonary fibrosis (IPF) without surgical lung biopsy
- However, visual assessment of CT is subjective and its overall sensitivity for a histologic UIP diagnosis is limited
- We developed a computer algorithm for prediction of UIP from CT and tested its accuracy against histologic diagnosis, visual CT assessment, and presence of the MUC5B promoter variant

Methods

- Using a multiple instance learning (MIL) paradigm, we trained a deep learning (DL) model for classification of UIP from CT
- The MIL algorithm was trained using n=1,770 chest CT exams with binary UIP labels (+/-)
- The model was tested, using Receiver Operating Characteristic (ROC) and Decision Curve Analysis (DCA)[1], in a separate group of n=128 CTs with radiologist visual assessment, histologic diagnosis and MUC5b genotyping[2]

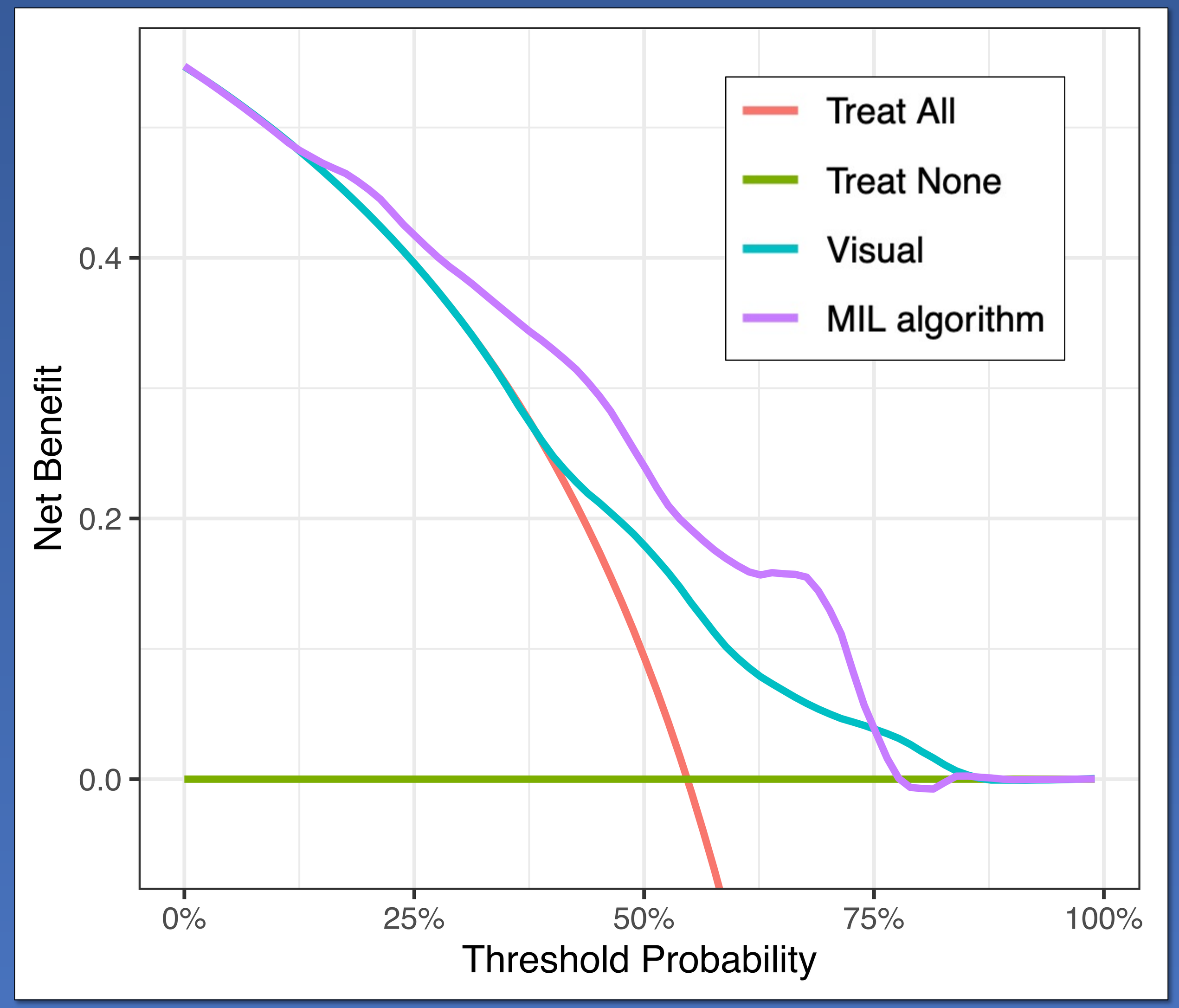
Results

- In the testing cohort, visual radiologic UIP diagnosis agreed moderately with histologic UIP (sensitivity 0.31, specificity 0.88) and MUC5b (sensitivity 0.29, specificity 0.82)
- The MIL algorithm showed improved sensitivity for histologic UIP (sensitivity 0.66, specificity 0.81) and MUC5b genotype (sensitivity 0.64, specificity 0.67) compared to visual CT assessment

	MIL Algorithm UIP classification		Visual radiologic UIP	
	Sensitivity	Specificity	Sensitivity	Specificity
Visual radiologic UIP	0.90	0.76	-	-
Histologic UIP	0.66	0.81	0.31	0.88
MUC5b	0.64	0.67	0.29	0.82

- Decision curve analysis showed that MIL classification provided greater net benefit than visual CT assessment for diagnosis of UIP

A **deep learning** algorithm trained to predict **UIP** from CT images can **increase** the **sensitivity** of UIP diagnosis



Decision curve analysis[1] *shows that the MIL algorithm provides greater "net benefit" for UIP classification than visual assessment of CT*

Scan QR score or visit URL for a webpage featuring all BI-supported publications at ATS 2022

https://www.usscicomms.com/respiratory/ATS2022

Scan QR score or visit URL for a device-friendly version of this poster

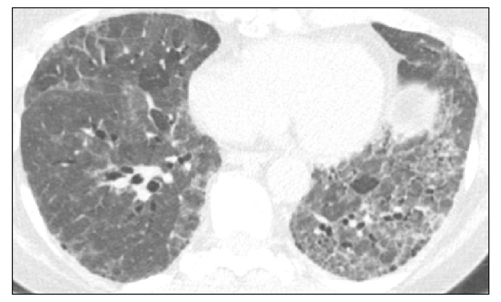
https://www.usscicomms.com/respiratory/ATS2022/Humphries

National Jewish Health®

Algorithm development

- MIL[3] is a deep learning approach that can produce subject-level predictions based on a collection of 2D image patches sampled from within lungs on CT
- Training is based on "bags", meaning collections of observations – in this application 2D 96x96 pixel patches sampled from a CT are a bag
  - A bag is labeled negative if all of the observations (patches) in it are negative
  - A bag is labeled positive if at least one observation (patch) in it is positive
- Training cases were drawn from the following cohorts:
  - Lung Tissue Research Consortium[4] (n=750), 10% UIP by central pathology
  - IPFNet[5] (n=370), 95% UIP positive by central radiology read
  - COPDGene[6] (n=500), 100% non-UIP (fibrotic ILD excluded)
  - RSNA RICORD[7] (n=150), 100% non-UIP (open database of acute COVID-19)

Example test images

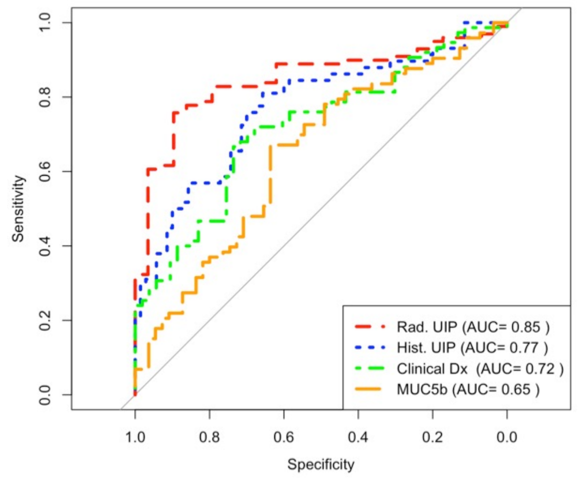


MIL score: 0.19
Radiologist assessment: Alternative diagnosis
Clinical Diagnosis: Hypersensitivity Pneumonitis

MIL score: 0.92
Radiologist assessment: definite UIP
Clinical Diagnosis: IPF

ROC analysis

- n=128 CT scans and diagnostic data from a prior study1, not overlapping with training data
- Probable + Definite categories for radiology and histology reads collapsed to obtain binary UIP classification
- ROC analysis of MIL algorithm outputs showed good concordance with radiologist visual classification (AUC=0.85) and histologic diagnosis (AUC=0.77)



Decision Curve Analysis

- DCA is a simple graphical method for evaluating prediction models and diagnostic tests
- It calculates a clinical "net benefit" for one or more prediction models in comparison to default strategies of treating all or no patients
- Net benefit differs from accuracy metrics such as discrimination and calibration because it incorporates the consequences of the decisions made on the basis of a model or test

net benefit = sensitivity × prevalence – (1 – specificity) × (1 – prevalence) × w

where w = the odds at the threshold probability

References:

1. Vickers AJ, et al. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26:565–574.
2. Chung JH, et al. CT scan findings of probable usual interstitial pneumonitis have a high predictive value for histologic usual interstitial pneumonitis. Chest. 2015;147:450–459.
3. Ilse M, et al. Attention-based deep multiple instance learning. PMLR 2018;80:2127–2136.
4. Lung Tissue Consortium (LTRC): https://www.nhlbi.nih.gov/science/lung-tissue-research-consortium-ltrc
5. Andrade J, et al. The Idiopathic Pulmonary Fibrosis Clinical Research Network (IPFnet): diagnostic and adjudication processes. Chest. 2015;148:1034–1042.
6. Regan EA, et al. Genetic epidemiology of COPD (COPDGene) study design. COPD. 2011;7:32–43.
7. Tsai EB, et al. The RSNA international COVID-19 open radiology database (RICORD). Radiology. 2021;299:E204–13.

Research support provided by:

Boehringer Ingelheim Pharmaceuticals, Inc.

The authors meet criteria for authorship as recommended by the International Committee of Medical Journal Editors (ICMJE). This was a collaborative research study where Boehringer Ingelheim Pharmaceuticals, Inc. (BIPI) was involved in the design, analysis or interpretation of the results but was not the regulatory sponsor. BI was given the opportunity to review the poster for medical and scientific accuracy as it relates to BI substances, as well as intellectual property considerations.

SC-US-74373