European Cystic Fibrosis Society

Journal of Cystic Fibrosis
The Official Journal of the European Cystic Fibrosis Society

Cystic Fibrosis Research News

Title:
Deep Learning to Automate Brasfield Chest Radiographic Scoring for Cystic Fibrosis

Authors:
Evan J. Zucker[a], Zachary A. Barnes[b], Matthew P. Lungren[a], Yekaterina Shpanskaya[a], Jayne M. Seekins[a], Safwan S. Halabi[a], David B. Larson[a]

Affiliations:
[a]Department of Radiology, Stanford University School of Medicine, 725 Welch Road, Stanford, CA 94305, USA
[b]Department of Computer Science, Stanford University, 353 Serra Mall, Stanford, CA 94305, USA

What was your research question?
The goal of this study was to determine if a deep learning computer model could be developed to measure the severity of CF lung disease on chest x-rays (CXRs) using the traditional Brasfield scoring method and whether this model can perform similarly to a paediatric radiologist (physician expert in medical imaging).

Why is this important?
Scoring systems such as the Brasfield method help to define the severity of CF lung disease in a reproducible, quantitative fashion for follow-up and treatment purposes. However, accurate scoring requires the time and expertise of radiologists in short supply. Recently, deep learning, a technique that allows computers to "learn" based on prior examples, has been successfully used to perform a variety of image recognition and classification tasks, including medical applications (e.g., pneumonia detection on CXRs). We hypothesized that the Brasfield method could similarly be automated with deep learning, potentially leading to increased patient access to fast and reliable scoring.

What did you do?
2058 CF CXRs performed over 10 years at our institution were scored by a paediatric radiologist using the Brasfield method, then divided into a "training/validation" set of 1858 exams and a "test set" of 200 exams. A deep learning model was iteratively optimized to predict Brasfield scores using only the "training/validation" set as "learning" inputs (i.e., CXRs annotated with scores). Then, the "test set" was scored by both the model and 5 total paediatric radiologists and their predictions compared. "Heat maps" were also generated to

show the portions of the CXRs that the model considered "important" in making its predictions.

What did you find?

The model, taking approximately 5 seconds to score the 200 test set exams, performed nearly as well as the radiologists in predicting Brasfield scores. The average difference between the model's total score predictions and those of the radiologists (averaged) was close to 0. In addition, model vs. radiologist total score correlation was almost as high as radiologist vs. radiologist total score correlation. Based on several metrics, the model also matched or exceeded radiologist performance for 3/5 Brasfield features (subscores). Finally, heat maps showed visual correlation between salient features identified by the model and CXR disease severity.

What does this mean and reasons for caution?

Deep learning techniques are promising for predicting Brasfield total and subscores at a level approaching that of paediatric radiologists, with rapid results. The main study limitations relate to the generalizability of the CXRs used to train and test the model. The model's performance would likely be improved with access to both a greater quantity and diversity of training data from multiple centres. In addition, although the test set was chosen to encompass a broad range of Brasfield feature severities, the model might perform differently on another test set (e.g., using CXRs from a different institution).

What's next?

Next steps would involve further refinement and multicentre validation of the model, which could then be evaluated in a clinical setting to determine its potential benefits (e.g., faster score reporting). Ultimately, the model could be widely distributed (e.g., via a website) to enable broad access to reliable CF CXR scoring.

Original manuscript citation in PubMed

https://www.ncbi.nlm.nih.gov/pubmed/?term=Deep+Learning+to+Automate+Brasfield+Chest+Radiographic+Scoring+for+Cystic+Fibrosis