SHAREABLE PDF

Collaboration between explainable artificial intelligence and pulmonologists improves the accuracy of pulmonary function test interpretation

Nilakash Das[1], Sofie Happaerts [2], Iwein Gyselinck [1,2], Michael Staes[1,2], Eric Derom [3], Guy Brusselle [3], Felip Burgos [4], Marco Contoli[5], Anh Tuan Dinh-Xuan [6], Frits M.E. Franssen [7], Sherif Gonem[8], Neil Greening[9], Christel Haenebalcke[10], William D-C. Man [11,12], Jorge Moisés [13], Rudi Peché[14], Vitalii Poberezhets [15], Jennifer K. Quint [11,12], Michael C. Steiner[9], Eef Vanderhelst [16], Mustafa Abdo[17], Marko Topalovic[18] and Wim Janssens [1,2]

[1]Laboratory of Respiratory Diseases and Thoracic Surgery, Department of Chronic Diseases Metabolism and Ageing, KU Leuven, Leuven, Belgium. [2]Clinical Department of Respiratory Diseases, University Hospitals Leuven, Leuven, Belgium. [3]UZ Gent, University of Ghent, Ghent, Belgium. [4]Department of Pulmonary Medicine, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona, Spain. [5]Department of Translational Medicine, University of Ferrara, Ferrara, Italy. [6]Service de Physiologie-Explorations Fonctionnelles, AP-HP, Hôpital Cochin, Université Paris Cité, Paris, France. [7]Department of Respiratory Medicine and School of Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University Medical Center, Maastricht, The Netherlands. [8]Nottingham University Hospitals NHS Trust, Nottingham, UK. [9]Leicester NIHR Biomedical Research Centre – Respiratory, Department of Respiratory Sciences, University of Leicester, Leicester, UK. [10]AZ Sint-Jan Brugge-Oostende, Bruges, Belgium. [11]National Heart and Lung Institute, Imperial College London, London, UK. [12]Royal Brompton and Harefield Clinical Group, Guy's and St Thomas' NHS Foundation Trust, London, UK. [13]Biomedical Research Networking Center on Respiratory Diseases (CIBERES), Madrid, Spain. [14]CHU Charleroi, Charleroi, Belgium. [15]Department of Propedeutics of Internal Medicine, National Pirogov Memorial Medical University, Vinnytsya, Ukraine. [16]University Hospital of Brussels, Vrije Universiteit Brussel, Brussels, Belgium. [17]LungenClinic Grosshansdorf, Grosshansdorf, Germany. [18]ArtiQ NV, Leuven, Belgium.

Corresponding author: Wim Janssens (wim.janssens@uzleuven.be)

Check for updates

Abstract

Background Few studies have investigated the collaborative potential between artificial intelligence (AI) and pulmonologists for diagnosing pulmonary disease. We hypothesised that the collaboration between a pulmonologist and AI with explanations (explainable AI (XAI)) is superior in diagnostic interpretation of pulmonary function tests (PFTs) than the pulmonologist without support.

Methods The study was conducted in two phases, a monocentre study (phase 1) and a multicentre intervention study (phase 2). Each phase utilised two different sets of 24 PFT reports of patients with a clinically validated gold standard diagnosis. Each PFT was interpreted without (control) and with XAI's suggestions (intervention). Pulmonologists provided a differential diagnosis consisting of a preferential diagnosis and optionally up to three additional diagnoses. The primary end-point compared accuracy of preferential and additional diagnoses between control and intervention. Secondary end-points were the number of diagnoses in differential diagnosis, diagnostic confidence and inter-rater agreement. We also analysed how XAI influenced pulmonologists' decisions.

Results In phase 1 (n=16 pulmonologists), mean preferential and differential diagnostic accuracy significantly increased by 10.4% and 9.4%, respectively, between control and intervention (p<0.001). Improvements were somewhat lower but highly significant (p<0.0001) in phase 2 (5.4% and 8.7%, respectively; n=62 pulmonologists). In both phases, the number of diagnoses in the differential diagnosis did not reduce, but diagnostic confidence and inter-rater agreement significantly increased during

intervention. Pulmonologists updated their decisions with XAI's feedback and consistently improved their baseline performance if AI provided correct predictions.

Conclusion A collaboration between a pulmonologist and XAI is better at interpreting PFTs than individual pulmonologists reading without XAI support or XAI alone.