Adaptation and validation of a natural language processing algorithm to use in electronic health records to identify patients with progressive fibrosing-interstitial lung disease in Spain: a real-world, cross-sectional, retrospective, observational study. The REVEAL-PF study

RWD94

Balcells, E[1]; Castellví, I[2]; Caballero, P[3]; Salinas, MB[4], Tort, M[5]

[1]Respiratory Department Hospital del Mar. Barcelona, Spain; [2]Rheumatologist at Hospital de la Santa Creu i Sant Pau, Barcelona, [3]Radiologist at Hospital Universitario de la Princesa, Madrid; [4]Pulmonologist at Hospital Universitario Basurto, Bilbao; [5]Boehringer Ingelheim España. Sant Cugat del Vallès, Spain

INTRODUCTION

- Progressive fibrosing-interstitial lung disease (PF-ILD) is a recognized condition that is often not registered in a homogeneous way in electronic health records (EHR).
- This is a limitation when describing epidemiological, clinical management information, use of resources, and costs related to PF-ILD.
- The REVEAL-PF study has two phases: phase 1 (one-site pilot study) and phase 2 (on-going multicenter study). This communication shows the results of phase 1.

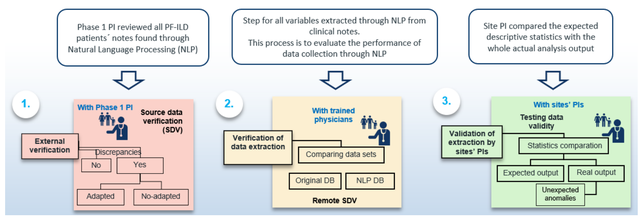
AIM

- The aim of phase 1 of the REVEAL-PF study was to adapt and validate a natural language processing (NLP) algorithm that identifies PF-ILD patients in the non-structured data of EHR in one Spanish hospital by using experts' verification as reference standard.

METHODS

- This cross-sectional, retrospective, and observational study included adults registered in the *Hospital del Mar* (Barcelona, Spain) between the 1st of January 2015 and the 31st of December 2019.
- The main study outcomes were precision, sensitivity, and accuracy metrics of the algorithm performance when identifying PF-ILD patients. PF-ILD cases were detected according to predefined progression criteria occurring within 24 months (decline in forced vital capacity, worsening of respiratory symptoms, and increased extent of fibrotic changes) and the consecutive identification of patients with fibrosing interstitial lung disease (**F-ILD) and** differentiation between Idiopathic Pulmonary Fibrosis (**IPF**) and other **F-ILD.**
- The algorithm validation process consisted in 3 steps (Figure 1).

Figure 1. Phase 1 - Algorithm validation (3 steps)



CONCLUSIONS

- The adapted algorithm showed to be precise, sensitive, and accurate in identifying PF-ILD diagnosis that were not explicitly coded in EHR. Considering the algorithm performance -which will likely improve as the algorithm learns- and its high applicability (e.g., standardized and economic), its implementation in health care settings to refine epidemiological indicators and guide resources allocation should be considered. In addition, the application of this methodology to other diseases showing differences in their EHR registration could be assessed.
- The phase 2 of the study is ongoing and aims to estimate the prevalence and incidence of PF-ILD according to underlying pathology in several Spanish hospitals.

RESULTS

Algorithm validation - step 1

The verification process was conducted by the PI. Each medical concept found by the algorithm was validated by the principal investigator (PI) who labeled it as correct (true positive) or incorrect (false positive) (Table 1).

Table 1. "Patient-level" precision and recall per each key variable, resulting from the PI's verification

Disease	Precision	Recall
IPF	84.51%	98.36%
Interstitial lung disease	98.53%	95.71%
Fibrosing lung disease	73.33%	91.67%
PF-ILD	100.00%	100.00%

Precision: Proportion of patients identified by the algorithm which were corroborated by the PI.

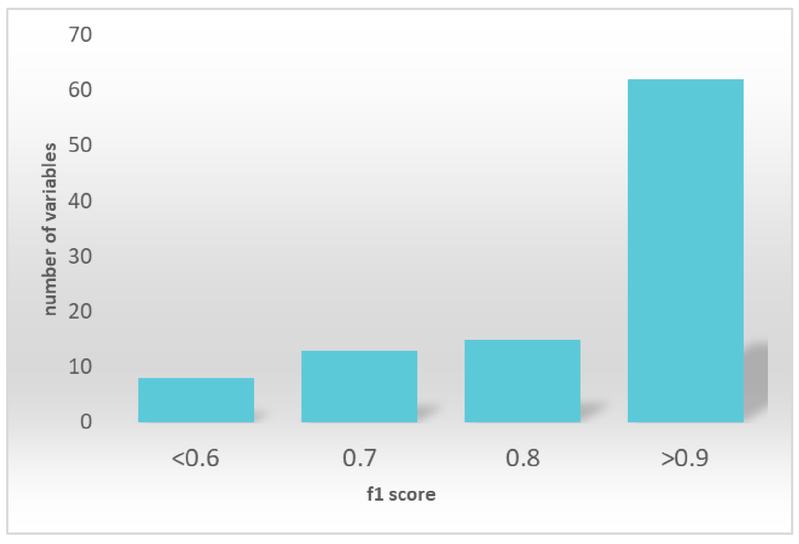
Recall: Proportion of patients identified by the PI which were also identified by the algorithm.

In **all cases**, **precision and recall** were **over 80%** and for PF-ILD, there were no false positives or false negatives, which translates into 100% precision and recall.

The fibrosing lung disease's low precision was not considered a constraint of the study since it is commonly used to describe general imaging fibrosing features not necessarily related with an interstitial lung disease and was used as the first filter for the inclusion criteria.

Algorithm validation - step 2

Figure 2. Number of variables according to different ranges of f1



F1 score: aggregated indicator of verification metrics (accuracy, recall and precision) which indicates the algorithm success/performance on capturing each variable found through NLP.

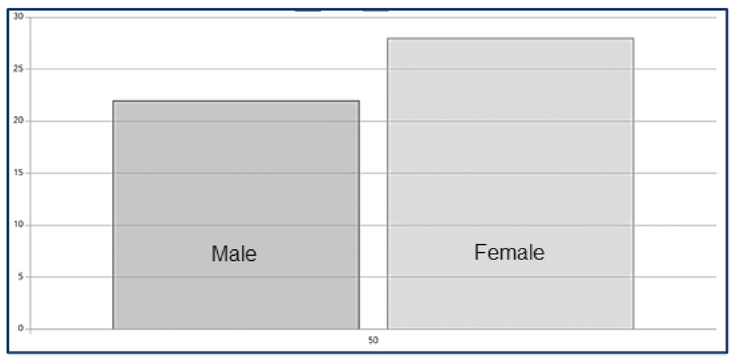
A 63% of the total number of variables analyzed (94 variables), had a f1 score above 90% (including the most important ones for this pathology, i.e., PF-ILD, ILD, IPF).

The performance of the variables will be improved in the phase 2 of the study due to the Machine Learning process.

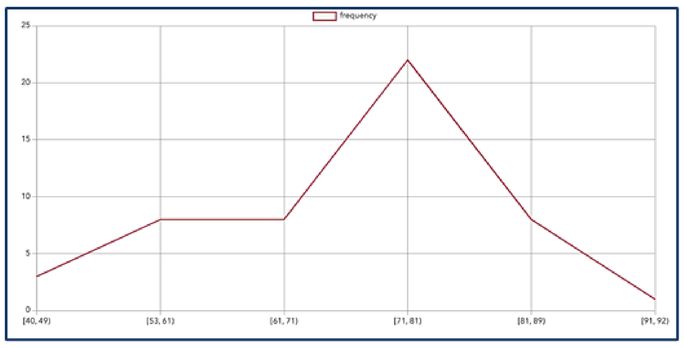
Algorithm validation - step 3

Figure 3. Distribution of observed values in study's population

Male vs. Female



Study's population distribution according to age



Few patients had an explicit free text of PF-ILD in the clinical notes indicating that "PF-ILD" may not be a term often used.

A total of **43 patients met the PF-ILD selection criteria**, which was **consistent with the PI's estimation** that, during study period, the total number of patients with PF-ILD were approximately between 40 to 50.

ABBREVIATIONS

EHR = Electronic Health Records; F-ILD = Fibrosing Interstitial Lung Disease; ILD = Interstitial Lung Disease; IPF = Idiopathic Pulmonary Fibrosis; NLP = Natural Language Processing; PF-ILD = Progressive-fibrosing Interstitial Lung Disease; PI = Principal Investigator

DISCLOSURES

The author(s) meet criteria for authorship as recommended by the International Committee of Medical Journal Editors (ICMJE). BI was given the opportunity to review the poster for medical and scientific accuracy as well as intellectual property considerations.
The authors did not receive payment related to the development of the poster.
The study was supported and funded by BI.

Boehringer Ingelheim

Poster presented at ISPOR Europe 2023, 12-15 November, Copenhagen, Denmark.