A multimedia library of interstitial lung diseases at the University Hospitals of Geneva

Alejandro Vargas, Adrien Depeursinge, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, Henning Müller

Geneva University Hospitals & University of Geneva, Switzerland

Contact address: Alejandro Vargas, Service of Medical Informatics, Geneva University Hospitals & University of Geneva, Rue Gabrielle-Perret-Gentil 4, 1211 Geneva 14, Switzerland. Email: alejandro.vargas@sim.hcuge.ch

Purpose

Challenges

• **Interstitial lung diseases (ILD)** include around 150 diseases that affect the lung parenchyma and are classified together because of similar **clinical**, **roentgenographic**, **physiologic**, or **pathologic** manifestations.

• The initial evaluation is based on the complete history, physical examination, laboratory tests, pulmonary function testing as well as a chest radiograph.

• In most of the cases, **high resolution computer tomography (HRCT)** of the chest is required to allow a rapid and accurate visual assessment of the lung tissue.

• Interpreting HRCT is often challenging with numerous differential diagnoses and requires experience for a correct interpretation compared to x-ray images of the chest.

Goals

• To retrospectively collect at least **150 cases** representative of the 15 most frequent ILDs at the University Hospitals of Geneva (HUG).

• To store the cases in a structured database containing selected clinical parameters and annotated HRCT image series.

• The database constitutes a basis for developing **image-based diagnosis aid computer tools** to assist the radiologist to the diagnosis workup of ILDs in clinical routine emergency.

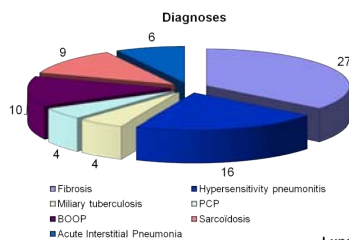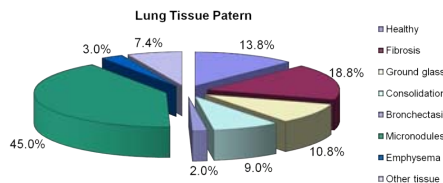• The database also creates opportunities for specialized studies and teaching.

Figure 1. Distribution of the diagnoses.

Figure 2. Distribution of the volumes of lung tissue patterns.

Results

Data acquisition

• A raw list of 1266 potential cases was extracted from the EHR data warehouse.

• So far, among 527 studied cases, 60 cases with annotated HRCT image series were entered into the database (see Table 1).

Diagnostic	Cases	%	VOIs (cm³)	%
Fibrosis	27	35.5	7305	21
Hypersensitivity pneumonitis	16	21.1	5433	15.6
Miliary tuberculosis	4	5.3	13654	39.2
PCP	4	5.3	2665	7.7
BOOP	10	13.2	1460	4.2
Sarcoidosis	9	11.9	2693	7.7
Acute Interstitial Pneumonia	6	7.9	1545	4.4
TOTAL	76		34775	

Table 1. Distribution of the cases, images series and VOIs according to diagnoses.

• 7 of the 15 diagnosis were retained as occurring sufficiently (see Figure 1).

• 34775 cm³ of lung tissue showing 16 types of healthy and pathological lung tissue were delineated in the 76 images series. Among the 16 types of lung tissue, the repartition of the 7 most represented is shown in Figure 2.

Statistics

• The mean age over the 60 cases is 65 years with a standard deviation of 19.

• 58.3% (35 cases of 60) are men.

• 76.7% (46 cases) underwent a confirming biopsy, 73.3% (44 cases) present a BAL. Among the 60 cases, 16.7% (10 cases) have neither a biopsy nor a BAL.

Computer tools

• A web-based interface was created to:
  • Browse the database.
  • Retrieve similar cases (see Figure 4).
  • Analyze a whole HRCT image series to carry out an automatic categorization of the pulmonary tissue (see Figure 5).

Methods & Materials

Database structure

• A set of 99 clinical parameters were defined corresponding to the 15 most frequent ILDs, describing a precise panorama of the patient's clinical state at the time of the stay when the HRCT image series was acquired.

Case selection process & data entry

• A raw list of patients that underwent a thorax CT within a stay in the pneumology service between 2003 and 2009 was extracted from the central data repository of HUG.

• Only cases with HRCT (without contrast agent, 1mm slice thickness) were kept.

• The diagnosis of the remaining cases was retraced in the electronic health record (EHR) based on the clinical history, reports and specific tests. Cases that underwent a pathological exam (biopsy, bronchoalveolar lavage (BAL)) that confirmed the diagnosis were kept for annotation sessions.
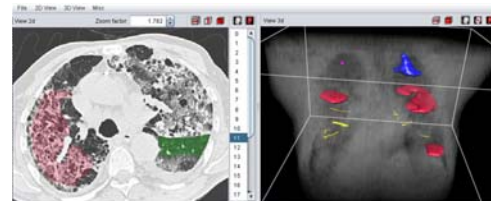

Figure 3. A screenshot of the graphical tool for the annotation of image regions.

• During regular annotation sessions with experienced radiologists, **3D volumes of interest (VOI)** showing lung tissue patterns that were consistent with the associated diagnosis were delineated in the HRCT images using an in-house annotation tool. (see Figure 3).

• Each retained case was added to the database by filling an HTML form containing the 99 defined clinical parameters associated with HRCT image series. Clinical data were taken within a margin of 1 month around the date of the image.

Figure 4. A screenshot of the graphical tool for similar cases retrieval.

Conclusions

Now available

• The database contains 40% (60 cases with at least one annotated HRCT image series) of the 150 projected cases. Each of the 60 cases has either a pathology study (biopsy, BAL) or a laboratory/specific test confirming the diagnosis.

• Computer tools have been implemented enabling image-based diagnosis aid with automatic detection of pathological lung tissue in HRCT images, retrieval of similar cases and database browsing.

Future work

• Beyond continuing the collection of cases.

• Add healthy cases in order to have a rich representation of healthy lung tissue and associated clinical parameters.
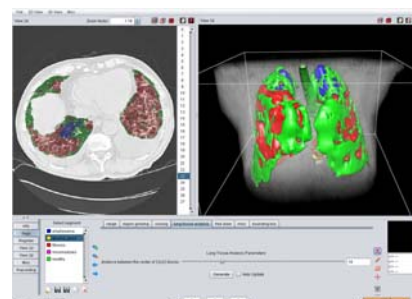

Figure 5. A screenshot of the graphical tool for the visualization of automatically segmented lung tissue.