Observer agreement in the diagnosis of interstitial lung diseases based on HRCT scans*, **

Concordância entre observadores no diagnóstico das doenças pulmonares intersticiais por imagens de TCAR

Viviane Baptista Antunes, Gustavo de Souza Portes Meirelles,
Dany Jasinowodolinski, Carlos Alberto de Castro Pereira,
Carlos Gustavo Yuji Verrastro, Fabíola Goda Torlai, Giuseppe D'Ippolito

Abstract

Objective: To determine the interobserver and intraobserver agreement in the diagnosis of interstitial lung diseases (ILDs) based on HRCT scans and the impact of observer expertise, clinical data and confidence level on such agreement. **Methods:** Two thoracic radiologists and two general radiologists independently reviewed the HRCT images of 58 patients with ILDs on two distinct occasions: prior to and after the clinical anamnesis. The radiologists selected up to three diagnostic hypotheses for each patient and defined the confidence level for these hypotheses. One of the thoracic and one of the general radiologists re-evaluated the same images up to three months after the first readings. In the coefficient analyses, the kappa statistic was used. **Results:** The thoracic and general radiologists, respectively, agreed on at least one diagnosis for each patient in 91.4% and 82.8% of the patients. The thoracic radiologists agreed on the most likely diagnosis in 48.3% ($\kappa = 0.42$) and 62.1% ($\kappa = 0.58$) of the cases, respectively, prior to and after the clinical anamnesis; likewise, the general radiologists agreed on the most likely diagnosis in 37.9% ($\kappa = 0.32$) and 36.2% ($\kappa = 0.30$) of the cases. For the thoracic radiologist, the intraobserver agreement on the most likely diagnosis was 0.73 and 0.63 prior to and after the clinical anamnesis, respectively. That for the general radiologist was 0.38 and 0.42. The thoracic radiologists presented almost perfect agreement for the diagnostic hypotheses defined with the high confidence level. **Conclusions:** Interobserver and intraobserver agreement in the diagnosis of ILDs based on HRCT scans ranged from fair to almost perfect and was influenced by radiologist expertise, clinical history and confidence level.

Keywords: Lung diseases, interstitial; Tomography, X-ray computed; Observer variation.

Resumo

Objetivo: Determinar a concordância interobservador e intraobservador no diagnóstico de doenças pulmonares intersticiais (DPIs) por TCAR e o impacto da experiência dos observadores, dos dados clínicos e do grau de confiança nessas concordâncias. **Métodos:** Dois radiologistas torácicos e dois gerais independentemente avaliaram imagens de TCAR de 58 pacientes com DPIs em dois momentos: antes e após da anamnese clínica. Os observadores selecionaram até três hipóteses diagnósticas para cada paciente e definiram o grau de confiança dessas hipóteses. Um dos radiologistas torácicos e um dos gerais reavaliaram as mesmas imagens até três meses após a primeira leitura. As análises estatísticas foram feitas utilizando o coeficiente kappa. **Resultados:** Os radiologistas torácicos e os gerais, respectivamente, concordaram com uma ou mais hipóteses diagnósticas em 91,4% e 82,8% dos pacientes. Os radiologistas torácicos concordaram com o diagnóstico mais provável em 48,3% ($\kappa = 0,42$) e 62,1% ($\kappa = 0,58$) dos casos, respectivamente, antes e após a anamnese clínica; de forma semelhante; os radiologistas gerais concordaram com o diagnóstico mais provável em 37,9% ($\kappa = 0,32$) e 36,2% ($\kappa = 0,30$). A concordância intraobservador do radiologista torácico no diagnóstico mais provável foi de 0,73 e 0,63, antes e após da anamnese clínica, respectivamente; para o radiologista geral, essa foi de 0,38 e 0,42. Os radiologistas torácicos apresentaram graus de concordância quase perfeitos nas hipóteses diagnósticas definidas com o grau de confiança alto. **Conclusões:** A concordância interobservador e intraobservador no diagnóstico das DPIs por TCAR variaram de regular a quase perfeita, tendo sido influenciadas pela experiência do radiologista, pela história clínica e pelo grau de confiança.

Descritores: Doenças pulmonares intersticiais; Tomografia computadorizada por raios X; Variações dependentes do observador.

Introduction

The use of HRCT has an established role in the detection and in the differential diagnosis of interstitial lung diseases (ILDs).[1-6] In selected cases, a specific HRCT pattern suffices for the presumptive diagnosis, even in the absence of histological confirmation.[7,8]

The evaluation of HRCT scans for ILDs relies on the subjective interpretation of images and the recognition of abnormal patterns, which are associated with interobserver and intraobserver variations. The lack of adequate training contributes to the variability of image interpretation and confidence in diagnoses made on the basis of HRCT findings.[7-9] The impact of clinical information, the observer expertise and the confidence level for the diagnostic hypotheses on the interobserver and intraobserver agreement in the diagnosis of ILDs based on HRCT scans are important issues that have not been fully evaluated.[7,10]

The aim of this study was to determine the interobserver and intraobserver agreement on the diagnosis of ILDs based on HRCT scans, as well as to evaluate the influence of the observer expertise, the confidence level and the clinical history of patients on the observer agreement, comparing two experienced thoracic radiologists with two young general radiologists who have received basic training in HRCT.

Methods

In order to select the study population, all 189 of the HRCT scans available between March of 2004 and June of 2006 in the archives at the Pulmonology Department of São Paulo Hospital, a tertiary referral center, were retrospectively reviewed. Only patients with technically adequate examinations who had been definitely diagnosed with ILDs by means of appropriate standards (clinical, laboratory and radiological data, as well as histological confirmation when necessary) were included. Those with postsurgical changes, active infections, malignant disease, predominantly airway disease or unavailable clinical records were excluded from the study. All of the scans were selected by the same radiologist, who did not take part in the evaluation of the scans. The study sample comprised the HRCT scans of 58 patients (30 females and 28 males). The mean age of females and males was, respec-

tively, 49.6 and 58.1 years (range: 29-81 years for the sample as a whole).

The diagnostic spectrum of the 58 cases encompassed a representative sample of the ILDs seen in our population. The final diagnoses and methods used to establish them are shown in Tables 1 and 2.

The institutional review board approval was issued by the Research Ethics Committee of the center where the study was performed on September 30th, 2005. The informed consent was not necessary, since it was a retrospective study.

We acquired HRCT images under the following conditions: 1.0 to 2.0 mm collimation, 10 to 15 mm interspaces, breath-holding at full inspiration and supine position. None of the patients received i.v. contrast material. The images were reconstructed with a high spatial frequency algorithm and photographed at appropriate window settings for viewing the lung parenchyma (window center, –600 to –800 HU; window width, 1,200 to 1,500 HU) and the mediastinum (window center, 30 to 50 HU; window width, 350 to 400 HU).

Two thoracic radiologists, each with more than 5 years of experience, and two radiologists, each with 2 years of experience in general radiology, independently reviewed the HRCT images. The general radiologists had received standard training in HRCT during their radiology academic studies, including the analysis of approximately 100 scans as a part of the residence program.

Table 1 – Definitive diagnosis of the patients included in the study.

Diagnosis	n	%
Sarcoidosis	11	19,0
Collagen/vascular diseases	11	19,0
Hypersensitivity pneumonitis	11	19,0
Idiopathic pulmonary fibrosis	9	15,5
Silicosis	5	8,6
Asbestosis	3	5,2
Cryptogenic organizing pneumonia	2	3,4
Langerhans cell histiocytosis	2	3,4
Coal worker's pneumoconiosis	1	1,7
Lymphangioleiomyomatosis	1	1,7
Alveolar proteinosis	1	1,7
Pulmonary ossification	1	1,7
Total	58	100

Table 2 – Definitive diagnosis, methods used to establish the diagnosis and histological diagnosis of each patient included in the study.

P	Definitive diagnosis	Methods	Histological diagnosis
1	Silicosis	SLB + TBLB	Silicosis
2	IPF	SLB	UIP
3	HP	TBLB	HP
4	HP	SLB	HP
5	Collagenosis	SLB	NSIP
6	Silicosis	TBLB + SLB	Silicosis
7	Silicosis	Clin + HRCT	-
8	Asbestosis	TBLB	Asbestosis
9	HP	SLB	HP
10	Sarcoidosis	TBLB	Sarcoidosis
11	Sarcoidosis	TBLB	Sarcoidosis
12	Sarcoidosis	Clin + HRCT	-
13	Collagenosis	TBLB + SLB	NSIP
14	Histiocytosis	SLB	Histiocytosis
15	Sarcoidosis	SLB	Sarcoidosis
16	Pulm Ossif	TBLB + SLB	Pulm Ossif
17	Sarcoidosis	TBLB	Sarcoidosis
18	Sarcoidosis	Clin + HRCT	-
19	IPF	Clin + HRCT	-
20	Asbestosis	Clin + HRCT	-
21	HP	SLB	HP
22	IPF	Clin + HRCT	-
23	IPF	TBLB	UIP
24	Silicosis	Clin + HRCT	-
25	Sarcoidosis	TBLB + SLB	Sarcoidosis
26	Proteinosis	Clin + HRCT + BAL	-
27	IPF	Clin + HRCT	-
28	CMP	TBLB + SLB	CMP
29	HP	SLB	HP
30	Silicosis	TBLB	Silicosis
31	HP	Clin + HRCT	-
32	Sarcoidosis	TBLB	Sarcoidosis
33	Collagenosis	SLB	LIP
34	Collagenosis	Clin + HRCT	-
35	Collagenosis	Clin + HRCT	-
36	HP	SLB	HP
37	Collagenosis	SLB	NSIP
38	Collagenosis	SLB	NSIP
39	LLM	SLB	LLM
40	Collagenosis	TBLB	LIP
41	Collagenosis	Clin + HRCT	-
42	Sarcoidosis	TBLB + SLB	Sarcoidosis
43	HP	TBLB + SLB	HP
44	IPF	TBLB + SLB	UIP
45	Sarcoidosis	SLB	Sarcoidosis
46	COP	TBLB	COP
47	IPF	TBLB + SLB	UIP
48	Asbestosis	Clin + HRCT	-
49	Collagenosis	SLB	NSIP
50	HP	TBLB	HP
51	HP	SLB	HP
52	HP	SLB	HP
53	IPF	SLB	UIP
54	HP	SLB	HP
55	Sarcoidosis	TBLB + SLB	Sarcoidosis
56	Histiocytosis	SLB	Histiocytosis
57	COP	SLB	COP
58	IPF	Clin + HRCT	-

P: patient; SLB: surgical lung biopsy; TBLB: transbronchial lung biopsy; IPF: idiopathic pulmonary fibrosis; HP: hypersensitivity pneumonitis; UIP: usual interstitial pneumonia; NSIP: nonspecific interstitial pneumonia; Pulm Ossif: pulmonary ossification; CMP: coal worker's pneumoconiosis; LIP: lymphoid interstitial pneumonia; LLM: lymphangioleiomyomatosis; COP: cryptogenic organizing pneumonia;; and Clin: clinical data.

None of the observers were aware of the specific diagnoses or had prior knowledge of any of the cases.

Images were initially reviewed without any knowledge of the clinical findings. All observers could list one to three diagnostic hypotheses for each patient and then select a confidence level of 1, 2 or 3 (low, medium or high, respectively) for each hypothesis. For the diagnostic hypotheses, radiologists had been asked to use the histopathological patterns rather than the clinical diagnoses. They had also been asked to use the latest American Thoracic Society/European Respiratory Society (ATS/ERS) classification of idiopathic interstitial pneumonias.[11] Immediately thereafter, all of the observers listed again one to three diagnostic hypotheses for each patient and selected a level of confidence for these hypotheses after the clinical information was provided. Sex; age; respiratory and systemic symptoms; smoking history; presence of other diseases, such as collagenoses and gastroesophageal reflux; environmental and occupational exposures; and physical examination were the items included in the clinical information list provided.

In order to evaluate intraobserver agreement, all images were reviewed by one of the thoracic radiologists and one of the general radiologists 1-3 months after the first analysis.

Statistical analyses were performed using the Minitab software program, version 14.2 (Minitab Inc., State College, MA, USA). Interobserver and intraobserver agreement on the most likely diagnoses and on at least one of the diagnoses for each patient, prior to and after the assessment of the clinical data, was quantified using the unadjusted kappa coefficient, categorized as "poor" for $\kappa \leq 0.20$; "fair" for $0.21 \leq \kappa \leq 0.40$; "moderate" for $0.41 \leq \kappa \leq 0.60$; "substantial" for $0.61 \leq \kappa \leq 0.80$; and "almost perfect" for $0.81 \leq \kappa \leq 1.00$.[12]

To a more accurate demonstration of differences, the chi-square test and Fisher's exact test were used. The level of significance was set at $p < 0.05$.

Results

The two thoracic radiologists agreed on at least one diagnosis for each patient in 53 (91.4%) and 54 (93.1%) of the cases, respectively, prior to and after the clinical information was provided,
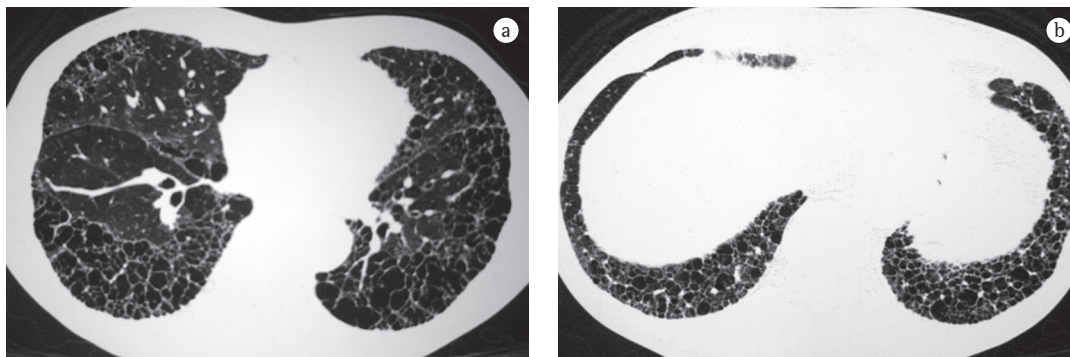
Figure 1a and b – HRCT scans of a patient with usual idiopathic pulmonary fibrosis (patient # 27 in Table 2). All observers agreed on usual interstitial pneumonia as the most likely diagnosis.

as did the two general radiologists in 48 (82.8%) and 47 (81.1%) of the cases.

The two thoracic radiologists agreed on the most likely diagnosis for each patient in 48.3% (κ = 0.42) and 62.1% (κ = 0.58) of the cases, respectively, prior to and after the clinical data were provided, as did the two general radiologists in 37.9% (κ = 0.32) and in 36.2% (κ = 0.30) of the cases (Figures 1-3).

The interobserver agreement between the thoracic and the general radiologists was statistically significant (p = 0.005) only for the most likely diagnostic hypothesis after the assessment of the clinical data.

Prior to receiving the clinical data of the patients in study, the two thoracic radiologists, respectively, selected the high level of confidence for the diagnosis for 15 (25.8%) and 36 (62.0%) of the patients, 9 (15.5%) of whom were selected by both. After reviewing

the clinical information, the two thoracic radiologists, respectively, selected the high level of confidence for the diagnosis for 36 (62.0%) and 47 (81.0%) of the patients, 30 (51.7%) of whom were selected by both. Subsequently, only the cases in which the level of confidence was defined as high by both thoracic radiologists were taken into consideration. There was interobserver agreement on the most likely diagnosis for 6 (66.7%) and 19 (63.3%) of these patients, respectively, prior to and after the review of the clinical data. The interobserver agreement was considered moderate (κ = 0.57) and almost perfect (κ = 0.85), respectively, prior to and after the clinical data were available.

One of the general radiologists selected the high level of confidence for the diagnosis for, respectively, 13 (22.4%) and 14 (24.1%) of the patients prior to and after the clinical information was provided. The other general radiologist
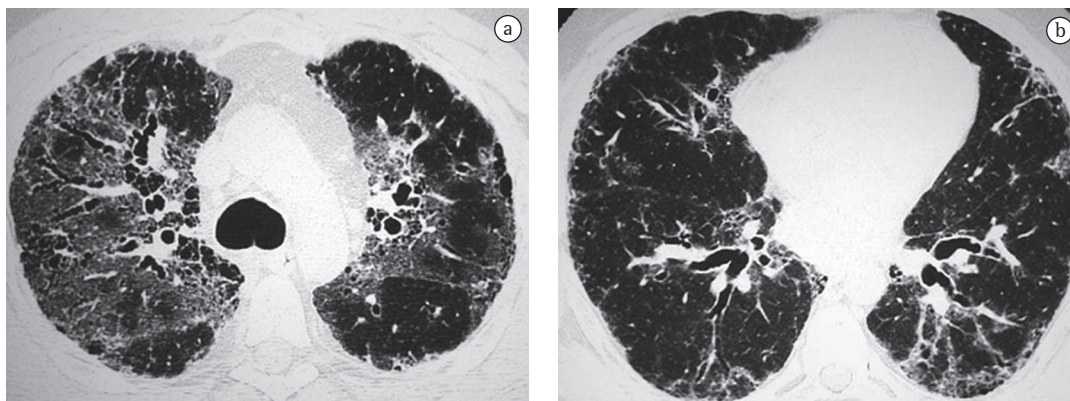

Figure 2a and b – HRCT scans of a patient with sarcoidosis, diagnosed by surgical lung biopsy (patient # 25 in Table 2). The observers selected chronic hypersensitivity pneumonitis or sarcoidosis as the most likely diagnosis. After the clinical data were provided, with no history of exposure, they agreed on sarcoidosis as the first-choice diagnosis.
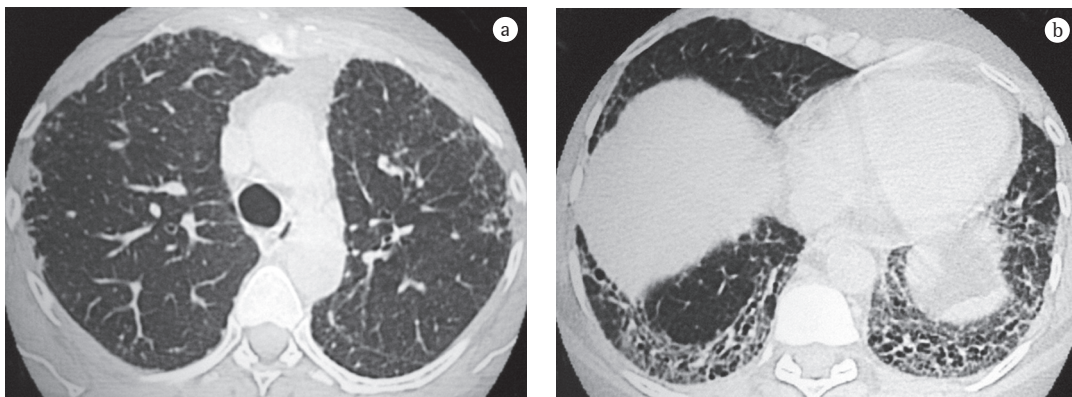
Figure 3a and b – HRCT scans of a patient with chronic hypersensitivity pneumonitis, diagnosed by surgical lung biopsy (patient # 29 in Table 2). The observers selected chronic hypersensitivity pneumonitis, usual interstitial pneumonia or non-specific pneumonia as the first-choice diagnosis. After the information of clinical history of bird exposure, the thoracic radiologists agreed on chronic hypersensitivity pneumonitis as the most likely hypothesis; the general radiologists agreed on usual interstitial pneumonia.

selected the high level of confidence for the diagnosis for none of the patients; therefore, we could not estimate the interobserver agreement on the most likely diagnosis for the general radiologists.

The thoracic radiologist who reviewed the HRCT scans on two distinct occasions agreed on at least one diagnostic hypothesis for each patient in 56 (96.6%) and 57 (98.3%) of the cases, respectively, prior to and after the clinical information was provided. The general radiologist who reviewed the HRCT scans on two distinct occasions agreed on at least one diagnosis for each patient in 51 (87.9%) and 50 (86.2%) of the cases, respectively, prior to and after the clinical information was provided.

The intraobserver agreement on the most likely diagnosis for each patient for the thoracic radiologist was, respectively, 74.1% ($\kappa = 0.73$) and 65.5% ($\kappa = 0.63$) prior to and after the clinical data were provided. The intraobserver agreement on the most likely diagnosis for each patient for the general radiologist was, respectively, 43.1% ($\kappa = 0.38$) and 46.6% ($\kappa = 0.42$) prior to and after the clinical data were provided.

The difference between the intraobserver agreement on the most likely diagnosis for the thoracic radiologist and for the general radiologist was statistically significant both prior to and after the clinical data were provided (p = 0.001 and p = 0.040, respectively).

The thoracic radiologist selected the high level of confidence for the diagnostic hypoth-

esis, on both occasions, for 10 (17.2%) and 31 (53.4%) of the patients, respectively, prior to and after the clinical information was provided. There was intraobserver agreement on the most likely diagnosis (high level of confidence) for 9 (90.0%) and 26 (86.7%) of the patients, respectively, prior to and after the clinical data were provided. This agreement was considered almost perfect on both occasions ($\kappa = 0.87$ and $\kappa = 0.85$ respectively).

Because the general radiologist did not select the high level of confidence for any of the cases, we were unable to estimate the agreement on the most likely diagnosis.

Discussion

In previous studies, interobserver agreement on the diagnosis of ILDs based on HRCT scans was estimated. Two groups of authors[3,6] found κ-values of 0.78 and 0.75, respectively, for the interobserver agreement on the most likely diagnosis between experienced radiologists. These studies, however, predate the latest ATS/ERS classification of idiopathic interstitial pneumonias[11]; therefore, the definitions of these diseases were unclear.

One group of authors,[13] evaluating only HRCT images from patients with idiopathic interstitial pneumonia, found a κ-value of 0.55 for the interobserver agreement regardless of the level of confidence, and a κ-value of 0.65 for the interobserver agreement on the diagnosis

with a high level of confidence. Another group of authors[9] studied the interobserver variation between 11 thoracic radiologists in the diagnosis of ILDs, comparing images from secondary and tertiary centers, and found an overall κ-value of 0.48. In that study, the κ-values ranged from 0.60, in secondary centers, to 0.34, in tertiary centers. These data well demonstrate the difference between observer agreement for complex cases at tertiary centers and for simple cases at secondary centers. In another study,[7] interobserver agreement on the diagnosis of idiopathic interstitial pneumonia was determined. A significant increase in agreement was found after the clinical information was provided, since κ-values increased from 0.72 to 0.80.

Observer experience and expertise are peculiar issues that were not considered in the previous studies, because in most of them only highly specialized observers were included.[3,6,7,9,13] In many centers, HRCT images are usually reviewed by general radiologists, who might not be familiar with the imaging aspects and the classification of ILDs. For this reason, we decided to compare interobserver and intraobserver agreement between thoracic and general radiologists. We found greater agreement between the thoracic radiologists than between the general radiologists. On the most likely diagnosis, the interobserver agreement for the thoracic radiologists was moderate (κ = 0.42); for the general radiologists, it was fair (κ = 0.32).

As the diagnostic hypotheses for ILDs based on HRCT scans are seldom confined to one possibility, it is also important to evaluate observer agreement taking into account other differential hypotheses. In our study, interobserver agreement was much higher when we took into account at least one of the three differential diagnoses listed by the radiologists for each patient. There was agreement between the two thoracic radiologists and the two general radiologists in 91.4% and 82.8% of the cases, respectively.

Clinical data had a major impact on the interobserver agreement for the thoracic radiologists but only when the most likely diagnosis was taken into account. The interobserver agreement on the most likely diagnostic hypothesis for the thoracic radiologists increased from 48.3% (κ = 0.42) to 62.1% (κ = 0.58) of the cases when the clinical information was provided.

As expected, a high level of confidence was much more frequently selected by the thoracic radiologists than by the general radiologists. The high level of confidence increased the κ-values from 0.42 to 0.57 for the interobserver agreement on the most likely diagnosis for the thoracic radiologists. The clinical information also increased the agreement on the selection of the high level of confidence. The interobserver agreement on the most likely diagnosis for the thoracic radiologists improved from moderate (κ = 0.58) to almost perfect (κ = 0.85). This improvement in agreement, especially when the clinical information was provided, confirms the importance of a confident hypothesis based on HRCT scans.

The intraobserver agreement was also higher for the thoracic radiologist than for the general radiologist (κ = 0.73 vs. κ = 0.38). Again, as expected, the agreement on the diagnostic hypotheses was higher for the most experienced observer. Intraobserver agreement was not statistically influenced by the clinical information. Similarly to the interobserver agreement, a diagnosis for which the level of confidence was classified as high was made more frequently by the thoracic radiologist than by the general radiologist. This significantly increased the degree of intraobserver agreement. For the thoracic radiologist, the intraobserver agreement on the diagnostic hypothesis with the high level of confidence was almost perfect prior to and after the clinical information was provided (κ = 0.87 and κ = 0.85, respectively).

In recent decades, the widely accepted gold standard for the diagnosis of ILDs was the histological diagnosis based on surgical biopsy. Concerns about the morbidity of the procedure, interobserver variations and various nonrepresentative specimens obtained from the lung biopsies have led to a reappraisal of this "gold standard" for the diagnosis of ILDs in clinical practice. It is increasingly accepted that the diagnosis of ILDs requires a multidisciplinary approach with the reconciliation of clinical, radiological and histological findings. Therefore, it is extremely important to have the most accurate observer agreement on the diagnoses based on HRCT scans.[14,15]

Our study has some limitations. Although ILDs are relatively uncommon, the number of cases we analyzed was smaller than was that

of other series. Future studies should involve larger patient samples, so that more significant conclusions can be made. Another limitation is the fact that the patients were selected from a highly specialized outpatient clinic at a tertiary center, which could have led to the inclusion of subjects with complex ILDs, which would have an impact on the observer agreement. The outpatient clinic is also a referral center for some entities, including hypersensitivity pneumonitis (HP). This explains the unusual distribution of cases, in which HP is one of the most common diseases and idiopathic pulmonary fibrosis is less common than in other series. However, the patients included encompassed a representative sample of diseases seen in our population.

In conclusion, interobserver and intraobserver agreement on the diagnosis of ILDs based on HRCT scans in our population ranged from fair to almost perfect and was highly influenced by the expertise of the radiologists, the clinical information provided and the level of confidence for the diagnostic hypothesis. The best agreement was achieved with the review of HRCT scans by experienced thoracic radiologists with the assessment of the clinical data and a high level of confidence for the diagnostic hypothesis. Basic training in HRCT during the residence program can provide good reproducibility of the method; however, experience seems to be crucial in order to increase confidence for the definition of a specific diagnosis and to enhance the value of the clinical data.

References

1. Padley SP, Hansell DM, Flower CD, Jennings P. Comparative accuracy of high resolution computed tomography and chest radiography in the diagnosis of chronic diffuse infiltrative lung disease. Clin Radiol. 1991;44(4):222-6.
2. Mathieson JR, Mayo JR, Staples CA, Müller NL. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography. Radiology. 1989;171(1):111-6.
3. Grenier P, Valeyre D, Cluzel P, Brauner MW, Lenoir S, Chastang C. Chronic diffuse interstitial lung disease: diagnostic value of chest radiography and high-resolution CT. Radiology. 1991;179(1):123-32.
4. Grenier P, Chevret S, Beigelman C, Brauner MW, Chastang C, Valeyre D. Chronic diffuse infiltrative lung disease: determination of the diagnostic value of clinical data, chest radiography, and CT and Bayesian analysis. Radiology. 1994;191(2):383-90.
5. Nishimura K, Izumi T, Kitaichi M, Nagai S, Itoh H. The diagnostic accuracy of high-resolution computed tomography in diffuse infiltrative lung diseases. Chest. 1993;104(4):1149-55.
6. Lee KS, Primack SL, Staples CA, Mayo JR, Aldrich JE, Müller NL. Chronic infiltrative lung disease: comparison of diagnostic accuracies of radiography and low- and conventional-dose thin-section CT. Radiology. 1994;191(3):669-73.
7. Flaherty KR, King TE Jr, Raghu G, Lynch JP 3rd, Colby TV, Travis WD, et al. Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? Am J Respir Crit Care Med. 2004;170(8):904-10.
8. Wells AU. Histopathologic diagnosis in diffuse lung disease: an ailing gold standard. Am J Respir Crit Care Med. 2004;170(8):828-9.
9. Aziz ZA, Wells AU, Hansell DM, Bain GA, Copley SJ, Desai SR, et al. HRCT diagnosis of diffuse parenchymal lung disease: inter-observer variation. Thorax. 2004;59(6):506-11.
10. Wells AU. High-resolution computed tomography in the diagnosis of diffuse lung disease: a clinical perspective. Semin Respir Crit Care Med. 2003;24(4):347-56.
11. American Thoracic Society; European Respiratory Society. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. Am J Respir Crit Care Med. 2002;165(2):277-304. Erratum in: Am J Respir Crit Care Med. 2002;166(3):426.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
13. Johkoh T, Müller NL, Cartier Y, Kavanagh PV, Hartman TE, Akira M, et al. Idiopathic interstitial pneumonias: diagnostic accuracy of thin-section CT in 129 patients. Radiology. 1999;211(2):555-60.
14. Thomeer M, Demedts M, Behr J, Buhl R, Costabel U, Flower CD, et al. Multidisciplinary interobserver agreement in the diagnosis of idiopathic pulmonary fibrosis. Eur Respir J. 2008;31(3):585-91.
15. Bradley B, Branley HM, Egan JJ, Greaves MS, Hansell DM, Harrison NK, et al. Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. Thorax. 2008;63 Suppl 5:v1-58.

About the authors

Viviane Baptista Antunes
Collaborating Physician. Department of Diagnostic Imaging, São Paulo Hospital, Federal University of São Paulo, São Paulo (SP) Brazil.

Gustavo de Souza Portes Meirelles
Collaborating Physician. Department of Diagnostic Imaging, São Paulo Hospital, Federal University of São Paulo, São Paulo (SP) Brazil.

Dany Jasinowodolinski
Collaborating Physician. Department of Diagnostic Imaging, São Paulo Hospital, Federal University of São Paulo, São Paulo (SP) Brazil.

Carlos Alberto de Castro Pereira
Attending Physician. Department of Internal Medicine, São Paulo Hospital, Federal University of São Paulo, São Paulo (SP) Brazil.

Carlos Gustavo Yuji Verrastro
Masters Student. Department of Diagnostic Imaging, São Paulo Hospital, Federal University of São Paulo, São Paulo (SP) Brazil.

Fabíola Goda Torlai
Collaborating Physician. Department of Diagnostic Imaging, São Paulo Hospital, Federal University of São Paulo, São Paulo (SP) Brazil.

Giuseppe D'Ippolito
Adjunct Professor. Federal University of São Paulo, São Paulo (SP) Brazil.