

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

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

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

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.

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.

* Trabalho realizado no Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.
Endereço para correspondência: Viviane Baptista Antunes. Rua Machado Bitencourt, 379, apto. 13, Vila Clementino, CEP 04044-001, São Paulo, SP, Brasil.
Tel 55 11 5014-6819. E-mail: antunes.viviane@gmail.com
Apoio financeiro: Nenhum.
Recebido para publicação em 16/7/2009. Aprovado, após revisão, em 21/8/2009.

Introdução

O uso da TCAR tem um papel estabelecido na detecção e no diagnóstico diferencial das doenças pulmonares intersticiais (DPIs).⁽¹⁻⁶⁾ Em casos selecionados, um padrão específico na TCAR é suficiente para o diagnóstico presuntivo, mesmo na ausência de confirmação histológica.^(7,8)

A avaliação de imagens de TCAR para DPIs baseia-se na interpretação subjetiva das imagens e no reconhecimento de padrões de anormalidade, associados a variações interobservador e intraobservador. A falta de treinamento adequado contribui para que haja variações na interpretação de imagens e na confiança nos diagnósticos feitos com base em achados na TCAR.⁽⁷⁻⁹⁾ O impacto das informações clínicas, da experiência dos observadores e do grau de confiança para as hipóteses diagnósticas sobre as concordâncias interobservador e intraobservador no diagnóstico de DPIs por TCAR são importantes aspectos que ainda não foram completamente avaliados.^(7,10)

O objetivo do presente estudo foi determinar as concordâncias interobservador e intraobservador no diagnóstico de DPIs por TCAR, bem como avaliar a influência da experiência dos observadores, do grau de confiança e da história clínica dos pacientes na concordância entre observadores, comparando dois experientes radiologistas torácicos e dois radiologistas gerais com treinamento básico em TCAR.

Métodos

Para selecionar a população do estudo, todas as 189 imagens de TCAR disponíveis nos arquivos do Departamento de Pneumologia do Hospital São Paulo, um centro de referência terciário, entre março de 2004 e junho de 2006, foram revisadas retrospectivamente. Apenas pacientes com exames tecnicamente adequados e com diagnóstico confirmado de DPIs por meio de padrões apropriados (dados clínicos, laboratoriais e radiológicos, bem como confirmação histológica se necessário) foram incluídos. Pacientes com alterações pós-cirúrgicas, infecções ativas, doenças malignas, doença predominante das vias aéreas ou registros clínicos indisponíveis foram excluídos do estudo. Todas as imagens foram selecionadas pelo mesmo radiologista, que não participou da avaliação dos exames. A amostra

do estudo foi composta por imagens de TCAR de 58 pacientes (30 mulheres e 28 homens). A idade média dos pacientes femininos e masculinos foi, respectivamente, 49,6 e 58,1 anos (variação: 29-81 anos para toda a amostra).

O espectro diagnóstico dos 58 casos compreendeu uma amostra representativa de DPIs observada em nossa população. Os diagnósticos finais e os métodos usados para estabelecê-los são apresentados nas Tabelas 1 e 2.

A aprovação do conselho de revisão da instituição foi emitida pelo Comitê de Ética em Pesquisa do centro onde o estudo foi realizado em 30 de setembro de 2005. O consentimento livre e esclarecido não foi necessário, uma vez que o presente foi um estudo retrospectivo.

Obtivemos imagens de TCAR nas seguintes condições: 1.0 a 2.0 mm de colimação a intervalos de 10 a 15 mm, apneia voluntária sob inspiração máxima e decúbito dorsal. Nenhum dos pacientes recebeu meio de contraste i.v. As imagens foram reconstruídas com algoritmo de alta frequência espacial e fotografadas com configuração adequada da janela para a visualização do parênquima pulmonar (centro, -600 a -800 HU; largura, 1.200 a 1.500 UH) e do mediastino (centro, 30 a 50 HU; largura, 350 a 400 UH).

Dois radiologistas torácicos, cada um com mais de 5 anos de experiência, e dois radiologistas, cada um com 2 anos de experiência em radiologia geral, revisaram as imagens de TCAR de forma independente. Os radiologistas gerais haviam recebido treinamento padrão para TCAR durante seus estudos acadêmicos em radio-

Tabela 1 - Diagnóstico definitivo dos pacientes incluídos no estudo.

Diagnóstico	n	%
Sarcoidose	11	19,0
Doenças do colágeno/vasculares	11	19,0
Pneumonite de hipersensibilidade	11	19,0
Fibrose pulmonar idiopática	9	15,5
Silicose	5	8,6
Asbestose	3	5,2
Pneumonia em organização criptogênica	2	3,4
Histiocitose de células de Langerhans	2	3,4
Pneumoconiose dos mineiros de carvão	1	1,7
Linfangioliomatose	1	1,7
Proteinose alveolar	1	1,7
Ossificação pulmonar	1	1,7
Total	58	100

Tabela 2 – Diagnóstico definitivo, métodos usados para estabelecer o diagnóstico e diagnóstico histológico de cada paciente incluído no estudo.

P	Diagnóstico definitivo	Métodos	Diagnóstico histológico
1	Silicose	BPC + BPTB	Silicose
2	FPI	BPC	PIU
3	PH	BPTB	PH
4	PH	BPC	PH
5	Colagenose	BPC	PINE
6	Silicose	BPTB + BPC	Silicose
7	Silicose	Clin + HRCT	-
8	Asbestose	BPTB	Asbestose
9	PH	BPC	PH
10	Sarcoidose	BPTB	Sarcoidose
11	Sarcoidose	BPTB	Sarcoidose
12	Sarcoidose	Clin + HRCT	-
13	Colagenose	BPTB + BPC	PINE
14	Histiocitose	BPC	Histiocitose
15	Sarcoidose	BPC	Sarcoidose
16	Ossif Pulm	BPTB + BPC	Ossif Pulm
17	Sarcoidose	BPTB	Sarcoidose
18	Sarcoidose	Clin + HRCT	-
19	FPI	Clin + HRCT	-
20	Asbestose	Clin + HRCT	-
21	PH	BPC	PH
22	FPI	Clin + HRCT	-
23	FPI	BPTB	PIU
24	Silicose	Clin + HRCT	-
25	Sarcoidose	BPTB + BPC	Sarcoidose
26	Proteinose	Clin + HRCT + BAL	-
27	FPI	Clin + HRCT	-
28	PMC	BPTB + BPC	PMC
29	PH	BPC	PH
30	Silicose	BPTB	Silicose
31	PH	Clin + HRCT	-
32	Sarcoidose	BPTB	Sarcoidose
33	Colagenose	BPC	PIL
34	Colagenose	Clin + HRCT	-
35	Colagenose	Clin + HRCT	-
36	PH	BPC	PH
37	Colagenose	BPC	PINE
38	Colagenose	BPC	PINE
39	LLM	BPC	LLM
40	Colagenose	BPTB	PIL
41	Colagenose	Clin + HRCT	-
42	Sarcoidose	BPTB + BPC	Sarcoidose
43	PH	BPTB + BPC	PH
44	FPI	BPTB + BPC	PIU
45	Sarcoidose	BPC	Sarcoidose
46	POC	BPTB	POC
47	FPI	BPTB + BPC	PIU
48	Asbestose	Clin + HRCT	-
49	Colagenose	BPC	PINE
50	PH	BPTB	PH
51	PH	BPC	PH
52	PH	BPC	PH
53	FPI	BPC	PIU
54	PH	BPC	PH
55	Sarcoidose	BPTB + BPC	Sarcoidose
56	Histiocitose	BPC	Histiocitose
57	POC	BPC	POC
58	FPI	Clin + HRCT	-

P: paciente; BPC: biópsia pulmonar cirúrgica; BPTB: biópsia pulmonar transbrônquica; FPI: fibrose pulmonar idiopática; PH: pneumonite de hipersensibilidade; PIU: pneumonia intersticial usual; PINE: pneumonia intersticial não específica; Ossif Pulm: Ossificação pulmonar; PMC: pneumoconiose dos mineiros de carvão; PIL: pneumonia intersticial linfocítica; LLM: linfangioliomatose; POC: pneumonia em organização criptogênica; e Clin: dados clínicos.

logia, incluindo a análise de aproximadamente 100 imagens de TCAR como parte do programa de residência. Nenhum dos observadores tinha conhecimento dos diagnósticos específicos ou tinha conhecimento prévio dos casos.

As imagens foram inicialmente revisadas sem o conhecimento dos achados clínicos. Todos os observadores poderiam listar até três hipóteses diagnósticas para cada paciente e então selecionar um grau de confiança de 1, 2 ou 3 (baixo, médio ou alto, respectivamente) para cada hipótese. Foi solicitado que os radiologistas usassem os padrões histopatológicos ao invés dos diagnósticos clínicos para selecionar as hipóteses diagnósticas. Também foi solicitado que usassem a mais recente classificação de pneumonias intersticiais idiopáticas da *American Thoracic Society/European Respiratory Society* (ATS/ERS).⁽¹¹⁾ Subsequentemente, todos os observadores receberam as informações clínicas, listaram novamente até três hipóteses diagnósticas para cada paciente e selecionaram um grau de confiança para tais hipóteses. Os itens incluídos na lista de informações clínicas fornecida foram os seguintes: sexo; idade; sintomas respiratórios; sintomas sistêmicos; história de tabagismo; presença de outras doenças, tais como colagenoses e refluxo gastroesofágico; exposição ambiental; exposição ocupacional e exame físico.

Para avaliar a concordância intraobservador, todas as imagens foram revisadas por um dos radiologistas torácicos e por um dos radiologistas gerais 1-3 meses após a primeira análise.

As análises estatísticas foram feitas usando o programa Minitab, versão 14.2 (Minitab Inc., State College, MA, EUA). As concordâncias interobservador e intraobservador quanto ao diagnóstico mais provável e quanto a pelo menos um dos diagnósticos para cada paciente, antes e após a avaliação dos dados clínicos, foram quantificadas utilizando o coeficiente kappa não ajustado e categorizadas como “fraca” para $\kappa \leq 0,20$; “regular” para $0,21 \leq \kappa \leq 0,40$; “moderada” para $0,41 \leq \kappa \leq 0,60$; “substancial” para $0,61 \leq \kappa \leq 0,80$; e “quase perfeita” para $0,81 \leq \kappa \leq 1,00$.⁽¹²⁾

Para uma demonstração de diferenças mais precisa, o teste do qui-quadrado e o teste exato de Fisher foram usados. O nível de significância estabelecido foi $p < 0,05$.

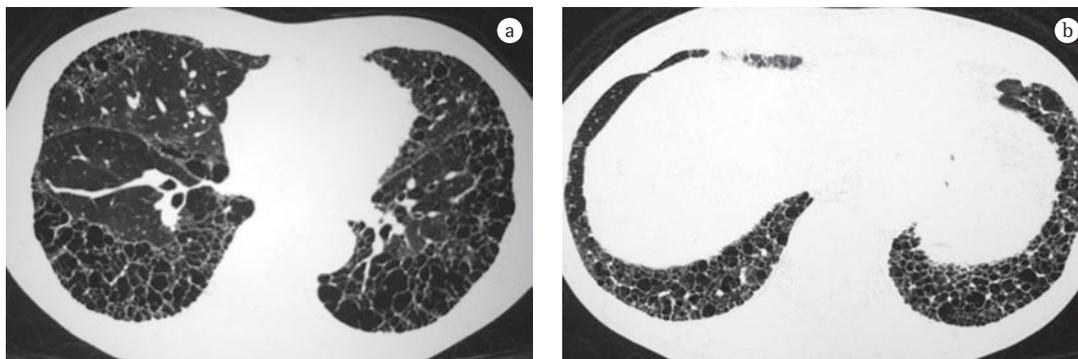


Figura 1a e 1b – TCAR de paciente com fibrose pulmonar idiopática usual (paciente nº 27 na Tabela 2). Todos os observadores concordaram que a pneumonia intersticial usual era o diagnóstico mais provável.

Resultados

Os dois radiologistas torácicos concordaram quanto a pelo menos um diagnóstico para cada paciente em 53 (91,4%) e 54 (93,1%) dos casos, respectivamente, antes e após o recebimento das informações clínicas. Os dois radiologistas gerais concordaram quanto a pelo menos um diagnóstico para cada paciente em 48 (82,8%) e 47 (81,1%) dos casos, respectivamente, antes e após o recebimento das informações clínicas.

Os dois radiologistas torácicos concordaram quanto ao diagnóstico mais provável para cada paciente em 48,3% ($\kappa = 0,42$) e 62,1% ($\kappa = 0,58$) dos casos, respectivamente, antes e após o recebimento das informações clínicas. Os dois radiologistas gerais concordaram quanto ao diagnóstico mais provável para cada paciente em 37,9% ($\kappa = 0,32$) e 36,2% ($\kappa = 0,30$) dos casos, respectivamente, antes e após o recebimento das informações clínicas (Figuras 1-3).

A concordância interobservador entre os radiologistas torácicos e os radiologistas gerais foi estatisticamente significativa ($p = 0,005$) apenas para a hipótese diagnóstica mais provável após a avaliação dos dados clínicos.

Antes do recebimento dos dados clínicos dos pacientes estudados, os dois radiologistas torácicos, respectivamente, selecionaram um alto grau de confiança no diagnóstico de 15 (25,8%) e 36 (62,0%) dos pacientes, 9 (15,5%) dos quais foram selecionados por ambos. Após receberem as informações clínicas, os dois radiologistas torácicos, respectivamente, selecionaram um alto grau de confiança no diagnóstico de 36 (62,0%) e 47 (81,0%) dos pacientes, 30 (51,7%) dos quais foram selecionados por ambos. Posteriormente, apenas os casos em que o grau de confiança foi definido como alto pelos dois radiologistas torácicos foram considerados. Houve concordância interobservador quanto ao diagnóstico

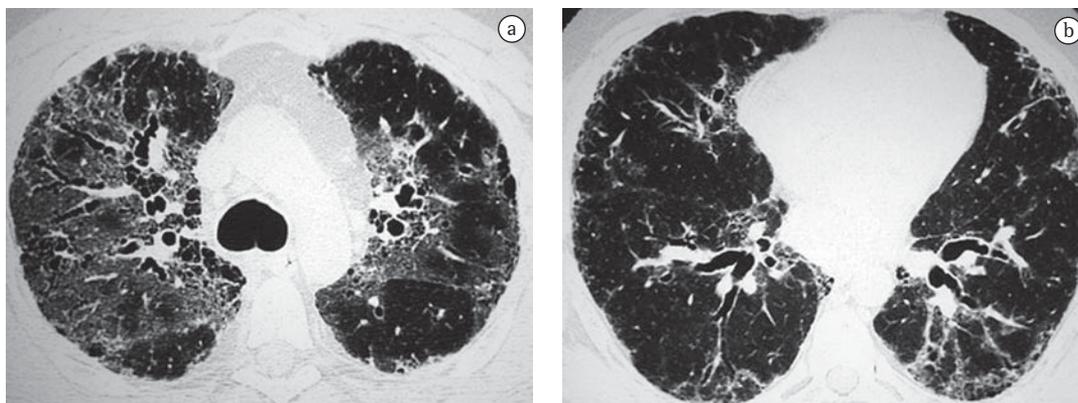


Figura 2a e 2b – TCAR de paciente com sarcoidose, diagnosticada através de biópsia pulmonar (paciente nº 25 na Tabela 2). Os observadores selecionaram pneumonia de hipersensibilidade ou sarcoidose como o diagnóstico mais provável. Após o recebimento dos dados clínicos, sem história de exposição, concordaram que sarcoidose era o diagnóstico mais provável.

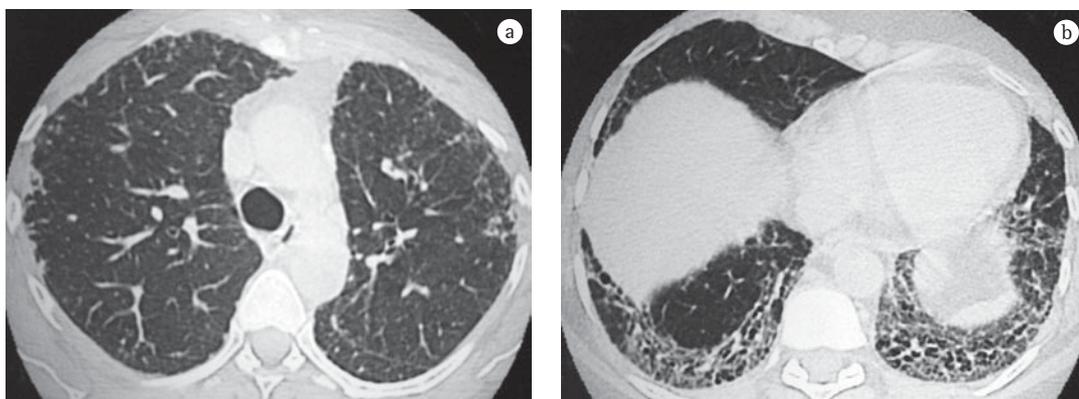


Figura 3a e 3b – TCAR de paciente com pneumonia de hipersensibilidade crônica, diagnosticada por biópsia pulmonar cirúrgica (paciente nº 29 na Tabela 2). Os observadores selecionaram pneumonia de hipersensibilidade crônica, pneumonia intersticial usual ou pneumonia não-específica como diagnóstico mais provável. Após o recebimento da história clínica de exposição a pássaros, os radiologistas torácicos concordaram que pneumonia de hipersensibilidade crônica era a hipótese mais provável e os radiologistas gerais concordaram que pneumonia intersticial usual era a hipótese mais provável.

mais provável em 6 (66,7%) e 19 (63,3%) desses pacientes, respectivamente, antes e após a revisão dos dados clínicos. A concordância inter-observador foi considerada moderada ($\kappa = 0,57$) e quase perfeita ($\kappa = 0,85$), respectivamente, antes e após o recebimento dos dados clínicos.

Um dos radiologistas gerais selecionou o grau de confiança alto no diagnóstico de 13 (22,4%) e 14 (24,1%) dos pacientes, respectivamente, antes e após o recebimento das informações clínicas. O outro radiologista geral não selecionou o grau de confiança alto para o diagnóstico de nenhum dos pacientes; portanto, não foi possível estimar a concordância interobservador dos radiologistas gerais quanto ao diagnóstico mais provável.

O radiologista torácico que revisou as imagens de TCAR em dois momentos distintos concordou quanto a pelo menos uma hipótese diagnóstica para cada paciente em 56 (96,6%) e 57 (98,3%) dos casos, respectivamente, antes e após o recebimento das informações clínicas. O radiologista geral que revisou as imagens de TCAR em dois momentos distintos concordou quanto a pelo menos um diagnóstico para cada paciente em 51 (87,9%) e 50 (86,2%) dos casos, respectivamente, antes e após o recebimento das informações clínicas.

A concordância intraobservador do radiologista torácico quanto ao diagnóstico mais provável para cada paciente foi 74,1% ($\kappa = 0,73$) e 65,5% ($\kappa = 0,63$), respectivamente, antes e após o recebimento das informações clínicas. A concordância intraobservador do radiologista

geral quanto ao diagnóstico mais provável para cada paciente foi 43,1% ($\kappa = 0,38$) e 46,6% ($\kappa = 0,42$), respectivamente, antes e após o recebimento das informações clínicas.

A diferença entre a concordância intraobservador do radiologista torácico e a do radiologista geral quanto ao diagnóstico mais provável foi estatisticamente significativa tanto antes quanto após o recebimento dos dados clínicos ($p = 0,001$ e $p = 0,040$, respectivamente).

O radiologista torácico selecionou o grau de confiança alto para a hipótese diagnóstica, nos dois momentos, de 10 (17,2%) e 31 (53,4%) dos pacientes, respectivamente, antes e após o recebimento das informações clínicas. Houve concordância intraobservador quanto ao diagnóstico mais provável (alto grau de confiança) em 9 (90,0%) e 26 (86,7%) dos pacientes, respectivamente, antes e após o recebimento dos dados clínicos. Essa concordância foi considerada quase perfeita nas duas ocasiões ($\kappa = 0,87$ e $\kappa = 0,85$, respectivamente).

Uma vez que o radiologista geral não selecionou o grau de confiança alto para nenhum dos casos, não foi possível estimar a concordância quanto ao diagnóstico mais provável.

Discussão

Em estudos prévios, a concordância interobservador para o diagnóstico de DPLs por TCAR foi estimada. Dois grupos de autores^(3,6) encontraram valores de kappa de 0,78 e 0,75, respectivamente, para a concordância interobservador de

radiologistas experientes quanto ao diagnóstico mais provável. Tais estudos, no entanto, são anteriores à classificação mais recente da ATS/ERS para pneumonias intersticiais idiopáticas⁽¹¹⁾; portanto, a definição dessas doenças ainda não estava clara.

Um grupo de autores,⁽¹³⁾ ao avaliar apenas imagens de TCAR de pacientes com pneumonia intersticial idiopática, encontrou um valor de kappa de 0,55 para a concordância interobservador independentemente do grau de confiança e um valor de kappa de 0,65 para a concordância interobservador quanto ao diagnóstico com alto grau de confiança. Outro grupo de autores⁽⁹⁾ estudou a variação interobservadores para o diagnóstico de DPLs em uma amostra de 11 radiologistas torácicos, comparando imagens de centros secundários e terciários, e encontrou um valor geral de kappa de 0,48. Nesse mesmo estudo, os valores de kappa variaram de 0,60, em centros secundários, a 0,34, em centros terciários. Esses dados demonstram bem a diferença de concordância entre observadores para casos complexos de centros terciários e para casos simples de centros secundários. Em outro estudo,⁽⁷⁾ determinou-se a concordância interobservador quanto ao diagnóstico de pneumonia intersticial idiopática. Um aumento significativo na concordância foi encontrado após o recebimento das informações clínicas, uma vez que os valores de kappa aumentaram de 0,72 para 0,80.

A experiência e a especialidade do observador são aspectos peculiares que não foram considerados em estudos prévios, uma vez que a maior parte deles incluiu apenas observadores altamente especializados.^(3,6,7,9,13) Em muitos centros, imagens de TCAR são geralmente revisadas por radiologistas gerais, que podem não estar familiarizados com os aspectos observados nas imagens e com a classificação das DPLs. Por isso, decidimos comparar as concordâncias interobservador e intraobservador entre radiologistas torácicos e radiologistas gerais. Encontramos melhor concordância entre os radiologistas torácicos do que entre os radiologistas gerais. A concordância interobservador dos radiologistas torácicos quanto ao diagnóstico mais provável foi moderada ($\kappa = 0,42$); a concordância interobservador dos radiologistas gerais quanto ao diagnóstico mais provável foi regular ($\kappa = 0,32$).

Uma vez que as hipóteses diagnósticas das DPLs por imagens de TCAR raramente estão restritas a uma possibilidade, é importante avaliar a concordância entre observadores considerando-se outras hipóteses diferenciais. No presente estudo, a concordância interobservador foi bem mais alta quando consideramos pelo menos um dos três diagnósticos diferenciais listados pelos radiologistas para cada paciente. Houve concordância entre os dois radiologistas torácicos e os dois radiologistas gerais em 91,4% e 82,8% dos casos, respectivamente.

Os dados clínicos tiveram um papel fundamental na concordância interobservador para os radiologistas torácicos, mas apenas quando o diagnóstico mais provável foi considerado. A concordância interobservador quanto à hipótese diagnóstica mais provável para os radiologistas torácicos aumentou de 48,3% ($\kappa = 0,42$) para 62,1% ($\kappa = 0,58$) dos casos após o recebimento das informações clínicas.

Como esperado, os radiologistas torácicos selecionaram um alto grau de confiança com muito mais frequência do que os radiologistas gerais. O alto grau de confiança aumentou os valores de kappa para a concordância interobservador dos radiologistas gerais quanto ao diagnóstico mais provável de 0,42 para 0,57. O recebimento de informações clínicas também aumentou a concordância quanto à seleção de grau de confiança alto. A concordância interobservador dos radiologistas torácicos quanto ao diagnóstico mais provável melhorou, passando de moderada ($\kappa = 0,58$) para quase perfeita ($\kappa = 0,85$). Essa melhora na concordância, especialmente após o recebimento de informações clínicas, confirma a importância de uma hipótese confiável baseada em imagens de TCAR.

A concordância intraobservador foi mais alta para o radiologista torácico do que para o radiologista geral ($\kappa = 0,73$ vs. $\kappa = 0,38$). Novamente, como esperado, a concordância quanto às hipóteses diagnósticas foi mais alta para o observador mais experiente. A concordância intraobservador não foi estatisticamente influenciada pelas informações clínicas. De forma semelhante à concordância interobservador, o diagnóstico com grau de confiança classificado como alto foi mais frequentemente estabelecido pelo radiologista torácico do que pelo radiologista geral. Isso aumentou significativamente o grau de concordância intraobservador. A concor-

dância intraobservador do radiologista torácico quanto à hipótese diagnóstica com alto grau de confiança foi quase perfeita antes e após o recebimento das informações clínicas ($\kappa = 0,87$ e $\kappa = 0,85$, respectivamente).

Nas últimas décadas, o diagnóstico histológico baseado em biópsia cirúrgica foi amplamente aceito como o padrão ouro para o diagnóstico de DPLs. Preocupações quanto à morbidade desse procedimento, a variações interobservadores e à obtenção de vários espécimes não representativos em biópsias pulmonares levaram à reavaliação desse “padrão ouro” para o diagnóstico de DPLs na prática clínica. Cada vez mais, reconhece-se que o diagnóstico de DPLs requer uma abordagem multidisciplinar, com a conciliação de achados clínicos, radiológicos e histológicos. Dessa forma, é extremamente importante que se tenha a mais precisa concordância entre observadores em diagnósticos por imagens de TCAR.^(14,15)

Nosso estudo teve algumas limitações. Embora as DPLs sejam relativamente incomuns, o número de casos analisados foi menor que o investigado em outras séries. Estudos futuros devem investigar amostras maiores para a obtenção de conclusões mais significantes. Outra limitação foi o fato de os pacientes terem sido selecionados em um ambulatório altamente especializado de um centro terciário, o que pode ter levado à inclusão de indivíduos com DPLs complexas, afetando a concordância entre observadores. O ambulatório é também um centro de referência para algumas afecções, inclusive pneumonite de hipersensibilidade (PH). Isso explica a distribuição não usual dos casos, sendo que a PH foi uma das doenças mais comuns, e a fibrose pulmonar idiopática foi menos frequente que em outras séries. No entanto, os pacientes incluídos constituíram uma amostra representativa das doenças observadas em nossa população.

Concluímos que a concordância interobservador e intraobservador para o diagnóstico de DPLs por imagens de TCAR na população estudada variou de regular a quase perfeita e foi altamente influenciada pela experiência dos radiologistas, pelas informações clínicas recebidas e pelo grau de confiança para a hipótese diagnóstica. A melhor concordância foi obtida através da revisão das imagens de TCAR por radiologistas torácicos experientes após a

avaliação dos dados clínicos e com alto grau de confiança para a hipótese diagnóstica. O treinamento básico em TCAR durante o programa de residência pode fornecer uma boa reprodutibilidade do método. No entanto, a experiência parece ser crucial para aumentar a confiança na definição de um diagnóstico específico e para aumentar o valor dos dados clínicos.

Referências

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.

Sobre os autores

Viviane Baptista Antunes

Médico Colaborador. Departamento de Diagnóstico por Imagem, Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.

Gustavo de Souza Portes Meirelles

Médico Colaborador. Departamento de Diagnóstico por Imagem, Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.

Dany Jasinowodolinski

Médico Colaborador. Departamento de Diagnóstico por Imagem, Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.

Carlos Alberto de Castro Pereira

Médico Contratado. Departamento de Medicina, Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.

Carlos Gustavo Yuji Verrastro

Mestrando. Departamento de Diagnóstico por Imagem, Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.

Fabiola Goda Torlai

Médico Colaborador. Departamento de Diagnóstico por Imagem, Hospital São Paulo, Universidade Federal de São Paulo, São Paulo (SP) Brasil.

Giuseppe D'Ippolito

Professor Adjunto. Universidade Federal de São Paulo, São Paulo (SP) Brasil.