

INTERSTITIAL LUNG DISEASE

HRCT diagnosis of diffuse parenchymal lung disease: inter-observer variation

Z A Aziz, A U Wells, D M Hansell, G A Bain, S J Copley, S R Desai, S M Ellis, F V Gleeson, S Grubnic, A G Nicholson, S P G Padley, K S Pointon, J H Reynolds, R J H Robertson, M B Rubens

Thorax 2004;59:506–511. doi: 10.1136/thx.2003.020396

See end of article for authors' affiliations

Correspondence to: Professor D M Hansell, Department of Radiology, Royal Brompton Hospital, London SW3 6NP, UK; d.hansell@rbh.nthames.nhs.uk

Received 16 December 2003
Accepted 25 March 2004

Background: This study was designed to measure inter-observer variation between thoracic radiologists in the diagnosis of diffuse parenchymal lung disease (DPLD) using high resolution computed tomography (HRCT) and to identify areas of difficulty where expertise, in the form of national panels, would be of particular value.

Methods: HRCT images of 131 patients with DPLD (from a tertiary referral hospital (n = 66) and regional teaching centres (n = 65)) were reviewed by 11 thoracic radiologists. Inter-observer variation for the first choice diagnosis was quantified using the unadjusted kappa coefficient of agreement. Observers stated differential diagnoses and assigned a percentage likelihood to each. A weighted kappa was calculated for the likelihood of each of the six most frequently diagnosed disease entities.

Results: Observer agreement on the first choice diagnosis was moderate for the entire cohort ($\kappa = 0.48$) and was higher for cases from regional centres ($\kappa = 0.60$) than for cases from the tertiary referral centre ($\kappa = 0.34$). 62% of cases from regional teaching centres were diagnosed with high confidence and good observer agreement ($\kappa = 0.77$). Non-specific interstitial pneumonia (NSIP) was in the differential diagnosis in most disagreements (55%). Weighted kappa values quantifying the likelihood of specific diseases were moderate to good (mean 0.57, range 0.49–0.70).

Conclusion: There is good agreement between thoracic radiologists for the HRCT diagnosis of DPLD encountered in regional teaching centres. However, cases diagnosed with low confidence, particularly where NSIP is considered as a differential diagnosis, may benefit from the expertise of a reference panel.

The evaluation of high resolution computed tomography (HRCT) as a diagnostic test has centred on its diagnostic accuracy. One aspect of HRCT that has not been comprehensively evaluated is observer variation in the context of diffuse parenchymal lung disease (DPLD), particularly among radiologists at regional teaching centres. Observer variation is an important aspect of the reproducibility of a diagnostic test, and is relevant to both radiological and histological evaluations which rely on subjective interpretation. In both disciplines the essential skill is pattern recognition and the classification of abnormal morphological patterns. Currently, the majority of patients with non-sarcoid interstitial lung disease are managed on the basis of HRCT observations, without histological data, making knowledge of observer variation pivotal to the routine clinical use of HRCT.

The review of HRCT by expert consensus opinion has been advocated in the recent American Thoracic Society/European Respiratory Society statement on the idiopathic interstitial pneumonias.¹ British Thoracic Society guidelines for the investigation and management of diffuse parenchymal lung disease have also recommended that national panels should be formed in order to standardise both HRCT and histopathological evaluation.² Although a national panel of UK histopathologists has existed for several years, a similar initiative has yet to be undertaken by radiologists with expertise in HRCT.

The aim of our study was to determine the level of observer variation for the HRCT diagnosis of diffuse lung disease and to identify areas of difficulty where expertise, in the form of a national panel, would be of particular use. This question was addressed by quantifying the extent to which experienced radiologists agree with each other in the HRCT diagnosis of diffuse lung disease in general, and in separate subgroups of

(a) consecutive unselected cases and (b) cases posing greater diagnostic difficulties in which surgical biopsy was undertaken.

METHODS

Patient population

The patient population (n = 131) consisted of two cohorts: 66 consecutive patients undergoing HRCT at a single tertiary referral hospital between January 1996 and December 1998 in whom a surgical lung biopsy was performed within 1 month and a histological diagnosis of diffuse lung disease was made (group A), and 65 consecutive patients undergoing HRCT at regional teaching centres in whom appearances were considered compatible with diffuse lung disease by the radiologist providing the case (group B). Ten of the 11 participating radiologists were asked to provide either six or seven consecutive cases to match the number of cases that were evaluated in the two groups. Cases of predominantly airway disease (such as bronchiectasis or constrictive bronchiolitis) or infection were excluded.

HRCT scanning protocol

All HRCT scans at the tertiary centre (group A, n = 66) were obtained on a CT scanner (Imatron Inc, San Francisco, CA, USA) with 1.5 mm collimation at full inspiration. Scans were obtained at 10 mm intervals in the supine position and images were reconstructed with a high spatial frequency

Abbreviations: AIP, acute interstitial pneumonia; COP, cryptogenic organising pneumonia; DPLD, diffuse parenchymal lung disease; EAA, extrinsic allergic alveolitis; IPF, idiopathic pulmonary fibrosis; LAM, lymphangioleiomyomatosis; LCH, Langerhans' cell histiocytosis; NSIP, non-specific interstitial pneumonia; SRILD, smoking related interstitial lung disease

algorithm and photographed at window settings appropriate for viewing the lung parenchyma (window centre -550 HU; window width 1500 HU). Currently accepted protocols³ for the acquisition of the HRCT scans in group B (n = 65) were used by the participating centres. All the images were evaluated on hard copy.

Observer characteristics and evaluation of images

The participating radiologists, all working at teaching hospitals, had completed their general radiological training 5–26 years previously and all had a declared interest in the HRCT diagnosis of diffuse lung disease. Details of each observer's experience are summarised in table 1. The HRCT images were reviewed without the provision of any clinical information. Differential diagnoses were specified with percentage likelihoods (censored at 5%, summing to 100% in each case). Observers were free to diagnose any disease entity that they considered classifiable as a diffuse lung disease, the only stipulation being that the recent ATS/ERS classification and terminology for the idiopathic interstitial pneumonias was used when applicable.¹ In addition, the likelihood that the disease was reversible as judged by the HRCT appearances^{4–6} was graded on a scale of 1–5 (<5%, 5–25%, 30–65%, 70–90%, and 95–100%, respectively).

Data analysis

For the purposes of analysis, diagnostic statements were categorised into 17 diagnostic subgroups (box 1). All statistical analyses were performed using STATA data analysis software (Computing Resource Centre, Santa Monica, CA, USA).

In each case the diagnosis of first choice was assigned a confidence rating of 1 (diagnostic likelihood <70% = low confidence), 2 (diagnostic likelihood 70–95% = high confidence), or 3 (100% = pathognomonic). The categories chosen were based on those used to assess the clinical probability of pulmonary embolism in the PIOPED study.⁷

Unadjusted kappa coefficients of agreement (κ) were computed in (a) the entire cohort, (b) in separate subgroups with summed confidence scores above and below the median value (that is, cases diagnosed with high and low confidence), and (c) in groups A and B.

The weighted kappa coefficient of agreement (κ_w) was then used to calculate the observer variation for the estimation of the probability of each of the six most frequently diagnosed conditions. In order to do this, the percentage likelihood given to each diagnosis was assigned a grade of 0–4 representing clinically useful probabilities: grade 0 = condition not included in the differential diagnosis, grade 1 = low probability (5–25%), grade 2 = intermediate probability (30–65%), grade 3 = high probability (70–95%), and grade 4 = pathognomonic (100%). Weighted kappa

Box 1 Categories of disease used for statistical analyses

- Idiopathic pulmonary fibrosis;
- Non-specific interstitial pneumonia;
- Smoking related interstitial lung disease (respiratory-bronchiolitis interstitial lung disease and desquamative interstitial pneumonia);
- Cryptogenic organising pneumonia;
- Lymphoid interstitial pneumonia;
- Acute interstitial pneumonia;
- Sarcoidosis;
- Extrinsic allergic alveolitis;
- Asbestosis;
- Drug induced lung disease;
- Langerhans' cell histiocytosis;
- Lymphangitis carcinomatosa;
- Eosinophilic pneumonia;
- Lymphangiomyomatosis;
- Bronchoalveolar cell carcinoma;
- Other (amyloidosis, silicosis, follicular bronchiolitis, idiopathic pulmonary haemorrhage, lipoid pneumonia, pulmonary oedema, Churg-Strauss syndrome and Wegener's granulomatosis);
- Unclassifiable.

values were calculated for idiopathic pulmonary fibrosis (IPF), non-specific interstitial pneumonia (NSIP), smoking related interstitial lung disease (SRILD), cryptogenic organising pneumonia (COP), sarcoidosis, and extrinsic allergic alveolitis (EAA). Weighted kappa values were calculated between paired observers and hence κ_w is expressed as median values with ranges for the 55 possible combinations of 11 observers (observer 1 v observer 2, observer 1 v observer 3, etc). κ_w values for the prediction of reversibility of disease were calculated for the entire cohort.

Data were interrogated to identify the sources of inter-observer variation and cases in which divergent diagnoses were made by at least two observers were tabulated—for example, IPF (n = 6), NSIP (n = 2) and EAA (n = 3) was categorised as diagnostic discordance between IPF and NSIP, IPF and EAA, and EAA and NSIP.

Observer agreement was categorised as poor, fair moderate, good, or excellent according to κ values of <0.20, 0.20–0.39, 0.40–0.59, 0.60–0.79, and >0.80, respectively.⁸

Table 1 Details of the 11 observers

Name	Age	Sex	No of years since completing general radiology training	Chest radiology fellowship undertaken	% of workload dedicated to thoracic work
Observer 1	53	M	17	Yes	70%
Observer 2	43	M	12	Yes	60%
Observer 3	37	F	6	Yes	30%
Observer 4	43	M	13	Yes	30%
Observer 5	40	F	4	No	40%
Observer 6	37	F	7	No	40%
Observer 7	55	M	26	Yes	75%
Observer 8	38	F	7	Yes	60%
Observer 9	41	M	8	Yes	95%
Observer 10	37	M	3	Yes	90%
Observer 11	42	M	12	Yes	55%

Table 2 Prevalence of individual diseases in groups A and B based on the diagnosis offered most frequently by the radiologists for each case

Disease	Tertiary referral cases (group A, n = 66)	Teaching hospital cases (group B, n = 65)
NSIP	25	11
IPF	10	14
Sarcoidosis	5	11
EAA	5	9
COP	8	4
SRILD	6	5
LCH	2	2
Asbestosis	1	3
LAM	1	1
Lymphangitis	0	2
carcinomatosis		
Amyloidosis	1	1
AIP	1	1
LIP	0	1
Alveolar proteinosis	1	0

NSIP, non-specific interstitial pneumonia; IPF, idiopathic pulmonary fibrosis; SRILD, smoking related interstitial lung disease; COP, cryptogenic organising pneumonia; EAA, extrinsic allergic alveolitis; LCH, Langerhans' cell histiocytosis; LAM, lymphangioleiomyomatosis; AIP, acute interstitial pneumonia; LIP, lymphoid interstitial pneumonia.

RESULTS

The prevalence of the conditions in groups A and B, as judged by the diagnosis offered most frequently by the radiologists for each case, is shown in table 2.

Prevalence of pathognomonic, high confidence, and low confidence observations

The median prevalence of first choice diagnoses thought to be pathognomonic or made with high confidence (>70% likelihood) was 69% (range 41–79%); appearances were considered by the 11 radiologists to be pathognomonic in a median of 25% of cases (range 0–49%). First choice diagnoses were made with low confidence (<70% likelihood) in a median of 31% of cases (range 21–59%). When the cohort was subdivided, first choice diagnoses were made with high confidence in a median of 60% (range 27–80%) and 77% (range 55–82%) for groups A and B, respectively.

Variation in first choice diagnosis

There was moderate agreement ($\kappa = 0.48$) on the first choice diagnosis for the entire cohort. Agreement for first choice diagnoses of the six most frequently offered diagnoses (IPF, NSIP, sarcoidosis, EAA, COP, and SRILD) was moderate to good, with agreement greatest for an HRCT diagnosis for sarcoidosis (table 3).

Observer agreement was substantially higher in unselected consecutive cases from regional teaching centres (group B, $\kappa = 0.60$) than in biopsied tertiary referral cases (group A,

Table 3 Unweighted kappa (κ) coefficients of agreement for the first choice diagnosis for the six most frequently diagnosed conditions

Diagnostic category	κ coefficient of first choice diagnosis
Idiopathic pulmonary fibrosis	0.50
Non-specific interstitial pneumonia	0.38
Sarcoidosis	0.62
Extrinsic allergic alveolitis	0.59
Cryptogenic organising pneumonia	0.37
Smoking related interstitial lung disease	0.30

$\kappa = 0.34$). Similarly, observer agreement was substantially higher in cases diagnosed with high confidence ($\kappa = 0.68$) than in those diagnosed with low confidence ($\kappa = 0.28$). For group B cases diagnosed with high confidence (40/66 = 62%), agreement was good to excellent ($\kappa = 0.77$).

Variation in diagnostic probabilities and in the probability of reversible disease

There was moderate to good agreement on the probability of the six most prevalent diagnoses as shown in (table 4). Weighted kappa values were moderate for the likelihood of COP, NSIP, and SRILD, and good for EAA, IPF, and sarcoidosis.

Agreement on the likelihood that disease was reversible was good (median $\kappa_w = 0.61$; 25th to 75th percentile 0.56–0.67).

Sources of inter-observer variation

The sources of inter-observer variation are shown in table 5. 138 individual disagreements were made in the entire cohort with 31 (22%) related to IPF/NSIP discordance. Overall, the diagnosis of NSIP was a frequent source of observer variation and was involved in 76 of the 138 disagreements (55%).

DISCUSSION

HRCT is the major diagnostic advance of the past two decades in diffuse lung disease,³ yet the inter-observer agreement in HRCT reporting has not been fully evaluated. Quantifying the observer agreement of a diagnostic test should form part of its formal evaluation;⁹ it is an important insight into a test's usefulness and may disclose strengths and expose weaknesses of the test that are not readily apparent from more conventional diagnostic accuracy studies. Hence, the aim of this study was to quantify the level of observer agreement among practising thoracic radiologists in the diagnosis of diffuse lung disease in order to determine the need for a reference panel.

For most cases from regional centres the first choice diagnosis was made with high confidence and good observer agreement ($\kappa = 0.77$). However, in a minority of regional cases where the diagnosis is made with low confidence (approximately one third), reference panel review is likely to be beneficial, especially if NSIP is suspected.

An important facet of this study is the use of weighted kappa to evaluate agreement in the estimation of diagnostic probabilities. In certain cases the diagnosis may be a close call—for example, observer 1 may state IPF 45%, NSIP 55% and observer 2 IPF 55%, NSIP 45%. The use of the unweighted kappa in this scenario would give the impression of spurious disagreement between the two observers, despite the fact that the percentage probabilities assigned to each condition were very similar. By converting percentage probabilities into five categories and then applying the weighted kappa, we were able to assess agreement across a range of clinically useful probabilities in specific diffuse lung

Table 4 Weighted kappa coefficients (κ_w) for the six most frequently diagnosed diffuse lung diseases

Diagnostic category	Median (range) κ_w coefficient of agreement
Idiopathic pulmonary fibrosis	0.63 (0.48–0.78)
Non-specific interstitial pneumonia	0.51 (0.27–0.78)
Sarcoidosis	0.70 (0.58–0.84)
Extrinsic allergic alveolitis	0.60 (0.36–0.78)
Cryptogenic organising pneumonia	0.49 (0.06–0.76)
Smoking related interstitial lung disease	0.51 (0.20–0.73)

Median (range) values for combinations of paired observations (n = 55).

Table 5 Noise analysis where divergent observations were made on four or more occasions

Combination of diseases where discordance was observed	No of disagreements
NSIP-IPF	31
NSIP-EAA	13
NSIP-SRILD	12
SRILD-EAA	9
NSIP-COP	7
NSIP-sarcoidosis	7
Sarcoidosis-IPF	6
Sarcoidosis-COP	6
IPF-COP	5
EAA-IPF	4
EAA-sarcoidosis	4

IPF, idiopathic pulmonary fibrosis; NSIP, non-specific interstitial pneumonia; SRILD, smoking related interstitial lung disease; COP, cryptogenic organising pneumonia; EAA, extrinsic allergic alveolitis.

diseases. The good weighted kappa value for IPF ($\kappa_w = 0.63$) therefore reflects agreement among radiologists for the exclusion of the disease. This is of particular relevance in view of the poor prognosis of patients with a typical HRCT appearance of IPF.¹⁰ Weighted kappa values were good or moderate for the other five most frequently offered diagnoses (range 0.49–0.70).

We also analysed the sources of inter-observer variation by identifying the frequency with which specific diseases were offered as a differential diagnosis. Our data indicate that 55% of observer noise was related to the diagnosis of NSIP. The greatest area of disagreement was in the distinction between IPF and NSIP (22% of overall noise), but there were also problems distinguishing NSIP from other diseases, particularly EAA, SRILD, COP, and sarcoidosis. The difficulties in making a diagnosis of NSIP most likely stems from differences in the HRCT descriptions of NSIP in the published literature.^{11–14} In addition, several studies have emphasised the significant overlap between NSIP and IPF, NSIP and EAA, and NSIP and COP.^{12 15 16}

Agreement in identifying reversible disease on HRCT scans was good ($\kappa = 0.61$). Studies have established that parenchymal consolidation,¹⁷ nodules,^{18–20} and ground glass opacity not associated with traction bronchiectasis or bronchiolactasis²¹ are signs of reversible lung disease. The good kappa value indicates that radiologists are aware of, and agreed on, the features of disease reversibility on the HRCT scan. Arguably, a statement on disease reversibility in some cases is as useful as the diagnosis itself.

To the best of our knowledge, there are no similar studies that have evaluated observer agreement for the HRCT diagnosis of consecutive cases of diffuse lung disease. A study by Collins *et al* assessed observer variation in pattern type and disease extent in fibrosing alveolitis on HRCT scans,²² but agreement on individual CT patterns does not necessarily translate into overall diagnostic agreement. The requisites of an observer agreement study include a large number of observers, cases that are representative of those encountered in everyday clinical practice, and observers who are not all academic radiologists but who provide a substantial proportion of the total number of opinions on cases of DPLD for the patient population.

Some of the early studies that evaluated the diagnostic accuracy of HRCT also included statements on observer variation, but they are limited because they predate the recent classification of the idiopathic interstitial pneumonias, unusual diagnoses were over-represented,^{23 24} and the number of observers was small.^{24–26} Kappa values were 0.78 and

0.75 for the studies by Grenier *et al*²³ and Lee *et al*,²⁵ respectively; apparently much higher than in the present study, but comparison of κ values between studies in which disease prevalence varies considerably is fraught as the κ value is highly dependent on disease prevalence.²⁷ More recently, a study by Johkoh *et al* achieved a κ value of 0.55 for a correct HRCT diagnosis,¹⁶ although in this study all the observers were aware that the differential diagnosis was limited to just the five types of idiopathic interstitial pneumonia which probably increased the value of κ . In a study assessing the need for a lung biopsy in patients with suspected IPF,²⁸ agreement for the presence or absence of IPF was similar to that found in our study ($\kappa = 0.54$ and 0.50, respectively).

One of the strengths of this study was the comparison between cases from a tertiary centre and those from secondary practices. Reliance on tertiary referral practice cases alone would have produced biased results; indeed, this has been a criticism of early studies that have quantified observer variation in cases that have not been representative of those encountered in routine clinical practice. The inclusion of cases from both secondary and tertiary practices provides a more representative picture of the observer variation that actually exists. The difference in κ values for the tertiary cases (group A) compared with regional centre cases (group B) was striking (0.34 and 0.60, respectively). A possible explanation is that cases at a referral centre are more likely to represent those at the unusual end of the spectrum and, by virtue of referral patterns, comprise the more challenging cases. Additionally, all these cases (group A) had a surgical biopsy implying that the HRCT appearances were not characteristic, although it is possible that referrals to a tertiary centre trigger a biopsy response more readily than at regional centres. Nevertheless, the difference between the two groups is clear.

Our results also show, not surprisingly, that a greater proportion of group A cases than group B cases were made with low confidence and, in addition, that observer agreement was highest in cases diagnosed with high confidence. The study by Mathieson *et al*²⁴ in the early 1990s established that, when a confident HRCT diagnosis was made, it was correct in 93% of cases. The link between confidence and accuracy suggests that cases where the diagnosis is made with low confidence may benefit from interpretation by a panel of radiologists with particular expertise in HRCT.

A recent study published in a companion paper has evaluated the observer variation between pathologists in diffuse lung disease.²⁹ The basic design of the studies was similar although, in the study of inter-observer variation between pathologists, observers chose a diagnosis from a specified list of 15 categories. In the present HRCT study observers could state any disease entity that was classified as a diffuse lung disease. This difference in methodology may have artificially increased inter-observer agreement for the pathologists. Nevertheless, κ values for tertiary referral cases were similar for radiologists and pathologists (0.34 and 0.38, respectively). These results highlight the fact that, in difficult cases of DPLD, reliance on either imaging or pathology in isolation is inadvisable. As suggested by the ATS/ERS guidelines,¹ a concerted effort should be made to integrate clinical information, HRCT findings, and the pathology (if this is available) before a final diagnosis is formulated.

Several issues surrounding this study require clarification. Firstly, we included cases without a histological diagnosis to allow for a comparison of observer variation between cases from a tertiary centre that come to biopsy with non-biopsied secondary practice cases. In addition, with increasing reliance on HRCT, biopsy cases are no longer representative of the larger population of patients with interstitial lung disease.

Indeed, our results show that the quantification of observer variation exclusively in biopsied cases of interstitial lung disease produces biased results. Secondly, all observer variation studies are inherently artificial because the test under evaluation is assessed in isolation, without the assimilation of clinical information that contributes to the diagnostic process. This is necessary as all aspects of a diagnostic test—observer variation being no exception—need to be evaluated without clinical information so that the results reflect the true properties of the test under scrutiny. The integration of clinical information with radiological assessment would be inappropriate in this study where the aim was specifically to quantify the observer noise for HRCT.

Non-thoracic radiologists were not included in the study because we felt it would be more appropriate to assess observer variation among thoracic radiologists who report HRCT scans on a regular basis and provide the opinions on which decisions are made. A reference panel is only warranted when experts disagree, so we specifically included thoracic radiologists with a designated interest in interstitial lung disease. Finally, the diagnoses used throughout the study represent the radiologists' diagnoses based on HRCT appearances. There was no independent "gold standard" as this concept is irrelevant to a study assessing observer agreement. This study has not attempted to measure the accuracy of HRCT in diffuse lung disease; importantly, a high level of agreement is not equivalent to high accuracy. There are mathematical models that may be used to estimate accuracy from agreement³⁰ but, as a general rule, agreement (as measured by kappa) should not be used as a surrogate for accuracy.

This type of study is subject to differences in behaviour that individuals demonstrated when asked to express confidence using numerical probabilities—for example, in our group of radiologists one observer never used 100% probability. A further related factor is that qualitative expressions of probability have different numerical meanings to different individuals, even those in the medical profession.³¹ A radiologist interpreting an HRCT scan may come to the conclusion that the appearances are "likely" to represent sarcoidosis. He states the likelihood of this diagnosis to be 80%, whereas another radiologist who also thinks that sarcoidosis is "likely" may record 60%. However, it is hoped that, by placing the percentages into clinically useful categories (for the weighted kappa analyses), some of the noise introduced by the different perceptions of the observers will have been minimised.

In conclusion, we have shown that thoracic radiologists are within the clinically acceptable range of observer variation for cases of diffuse lung disease encountered in regional teaching centres. However, the low agreement observed for cases diagnosed with low confidence justifies a reference panel that would parallel the existing UK pathology group. The purpose of the reference panel would not be to improve the accuracy of HRCT against biopsy, but to provide an opportunity to reach consensus and standardise diagnoses in areas of contention.

Authors' affiliations

Z A Aziz, D M Hansell, M B Rubens, Department of Radiology, Royal Brompton Hospital, London, UK
A U Wells, Interstitial Lung Unit, Royal Brompton Hospital, London, UK
G A Bain, Department of Radiology, Central Middlesex Hospital, London, UK
S J Copley, Department of Radiology, Hammersmith Hospital, London, UK
S R Desai, Department of Radiology, King's College Hospital, London, UK
S M Ellis, Department of Radiology, London Chest Hospital, London, UK

F V Gleeson, Department of Radiology, Churchill Hospital, Oxford, UK
S Grubnic, Department of Radiology, St George's Hospital, London, UK
A G Nicholson, Department of Histopathology, Royal Brompton Hospital, London, UK
S P G Padley, Department of Radiology, Chelsea and Westminster Hospital, London, UK
K S Pointon, Department of Radiology, Nottingham City Hospital, Nottingham, UK
J H Reynolds, Department of Radiology, Birmingham Heartlands Hospital, Birmingham, UK
R J H Robertson, Department of Radiology, Leeds General Infirmary, Leeds, UK

REFERENCES

- 1 **Travis WD**, King TE. American Thoracic Society/European Respiratory Society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2002;**165**:277–304.
- 2 **British Thoracic Society**. The diagnosis, assessment and treatment of diffuse parenchymal lung disease in adults. Introduction. *Thorax* 1999;**54**(Suppl 1):S1–14.
- 3 **Kazerooni EA**. High-resolution CT of the lungs. *AJR Am J Roentgenol* 2001;**177**:501–19.
- 4 **Muller NL**, Staples CA, Miller RR, et al. Disease activity in idiopathic pulmonary fibrosis: CT and pathologic correlation. *Radiology* 1987;**165**:731–4.
- 5 **Wells AU**, Hansell DM, Rubens MB, et al. The predictive value of thin-section computed tomography in fibrosing alveolitis. *Am Rev Respir Dis* 1993;**148**:1076–82.
- 6 **MacDonald SL**, Muller NL. Insights from HRCT: how they affect the management of diffuse parenchymal lung disease. *Semin Respir Crit Care Med* 2003;**24**:357–63.
- 7 **PIOPED investigators**. Value of ventilation/perfusion scan in acute pulmonary embolism. *JAMA* 1990;**263**:2753–9.
- 8 **Coblentz CL**, Babcock CJ, Alton D, et al. Observer variation in detecting the radiologic features associated with bronchiolitis. *Invest Radiol* 1991;**26**:115–8.
- 9 **Hansell DM**, Wells AU. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Radiol* 2003;**58**:573–4.
- 10 **Faherty KR**, Thwaite EL, Kazerooni EA, et al. Radiological versus histological diagnosis in UIP and NSIP: survival implications. *Thorax* 2003;**58**:143–8.
- 11 **Park JS**, Lee KS, Kim JS, et al. Nonspecific interstitial pneumonia with fibrosis: radiographic and CT findings in seven patients. *Radiology* 1995;**195**:645–8.
- 12 **MacDonald SL**, Rubens MB, Hansell DM, et al. Nonspecific interstitial pneumonia and usual interstitial pneumonia: comparative appearances at and diagnostic accuracy of thin-section CT. *Radiology* 2001;**221**:600–5.
- 13 **Cottin V**, Donsbeck AV, Revel D, et al. Nonspecific interstitial pneumonia. Individualization of a clinicopathologic entity in a series of 12 patients. *Am J Respir Crit Care Med* 1998;**158**:1286–93.
- 14 **Nagai S**, Kitaichi M, Itoh H, et al. Idiopathic nonspecific interstitial pneumonia/fibrosis: comparison with idiopathic pulmonary fibrosis and bronchiolitis obliterans organizing pneumonia. *Eur Respir J* 1998;**12**:1010–19.
- 15 **Hartman TE**, Swensen SJ, Hansell DM, et al. Nonspecific interstitial pneumonia: variable appearance at high-resolution chest CT. *Radiology* 2000;**217**:701–5.
- 16 **Johkoh T**, Muller NL, Cartier Y, et al. Idiopathic interstitial pneumonias: diagnostic accuracy of thin-section CT in 129 patients. *Radiology* 1999;**211**:555–60.
- 17 **Hommeyer SH**, Godwin JD, Takasugi JE. Computed tomography of air-space disease. *Radiol Clin North Am* 1991;**29**:1065–84.
- 18 **Brauner MW**, Lenoir S, Grenier P, et al. D. Pulmonary sarcoidosis: CT assessment of lesion reversibility. *Radiology* 1992;**182**:349–54.
- 19 **Remy-Jardin M**, Remy J, Wallaert B, et al. Subacute and chronic bird breeder hypersensitivity pneumonitis: sequential evaluation with CT and correlation with lung function tests and bronchoalveolar lavage. *Radiology* 1993;**189**:111–8.
- 20 **Soler P**, Bergeron A, Kambouchner M, et al. Is high-resolution computed tomography a reliable tool to predict the histopathological activity of pulmonary Langerhans cell histiocytosis? *Am J Respir Crit Care Med* 2000;**162**:264–70.
- 21 **Remy-Jardin M**, Giraud F, Remy J, et al. Importance of ground-glass attenuation in chronic diffuse infiltrative lung disease: pathologic-CT correlation. *Radiology* 1993;**189**:693–8.
- 22 **Collins CD**, Wells AU, Hansell DM, et al. Observer variation in pattern type and extent of disease in fibrosing alveolitis on thin section computed tomography and chest radiography. *Clin Radiol* 1994;**49**:236–40.
- 23 **Grenier P**, Valeyre D, Cluzel P, et al. Chronic diffuse interstitial lung disease: diagnostic value of chest radiography and high-resolution CT. *Radiology* 1991;**179**:123–32.
- 24 **Mathieson JR**, Mayo JR, Staples CA, et al. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography. *Radiology* 1989;**171**:111–6.
- 25 **Lee KS**, Primack SL, Staples CA, et al. Chronic infiltrative lung disease: comparison of diagnostic accuracies of radiography and low- and conventional-dose thin section CT. *Radiology* 1994;**191**:669–73.

- 26 **Padley SPG**, Hansell DM, Flower CDR, *et al.* Comparative accuracy of high resolution computed tomography and chest radiography in the diagnosis of chronic diffuse infiltrative lung disease. *Clin Radiol* 1991;**44**:227–31.
- 27 **Kundel HL**, Polansky M. Measurement of observer agreement. *Radiology* 2003;**228**:303–8.
- 28 **Hunninghake GW**, Zimmerman MB, Schwartz DA, *et al.* Utility of a lung biopsy for the diagnosis of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2001;**164**:193–6.
- 29 **Nicholson AG**, Addis BJ, Bharucha H, *et al.* Inter-observer variation between pathologists in diffuse parenchymal lung disease. *Thorax* 2004;**59**:000–0.
- 30 **Kundel HL**, Polansky M. Mixture distribution and receiver operating characteristic analysis of bedside chest imaging with screen-film and computed radiography. *Acad Radiol* 1997;**4**:1–7.
- 31 **Kong A**, Barnett GO, Mosteller F, *et al.* How medical professionals evaluate expressions of probability. *N Engl J Med* 1986;**315**:740–4.

LUNG ALERT

Evidence to support RSV vaccination of the elderly?

▲ Walsh EE, Peterson DR, Falsey AR. Risk factors for severe respiratory syncytial virus infection in elderly persons. *J Infect Dis* 2004;**189**:233–8.

RSV is an important respiratory pathogen in infants. There is increasing evidence that it also plays a significant role in seasonal morbidity of the elderly and those with underlying cardiopulmonary disease. This study followed two “at risk” populations for up to two winter seasons: a cohort of healthy elderly subjects (>65 years: n = 216 (year 1), n = 289 (year 2)) and adults with cardiopulmonary disease (n = 204 (year 1), n = 265 (year 2)). A third group of 625 patients admitted acutely into hospital with respiratory symptoms and who were also in either of these “at risk” groups was also studied. A PCR positive swab, detection of RSV in cell culture, or a fourfold rise in serum IgG to RSV envelope glycoproteins was taken as evidence of RSV infection. Overall, 130 RSV infections were identified over 2 years: 61 in the hospitalised group, 32 in the healthy elderly subjects, and 37 in those with cardiopulmonary disease. A multivariate analysis revealed that age <65 years, poor baseline functional status, and low initial titres of protective RSV neutralising antibody were independently associated with RSV related hospitalisation. There was a non-significant trend for an effect of chronic pulmonary disease. Neutralising antibody levels to RSV were lower in the hospitalised group.

These data suggest that inadequate humoral immunity is an independent risk factor for hospitalisation during RSV infection in at risk populations, and that these groups might benefit most from protective vaccination.

T M A Wilkinson

Clinical Research Fellow, St Bartholomew's Hospital, London, UK;
tomw1970@hotmail.com

LUNG ALERT

Hypermethylation of FHIT predicts poor prognosis in NSCLC

▲ Maruyama R, Sugio K, Yoshino I, *et al.* Hypermethylation of FHIT as a prognostic marker in non-small cell lung carcinoma. *Cancer* 2004;**100**:1472–7.

DNA methylation of CpG sites in the promoter regions of tumour suppressor genes is a frequently acquired epigenetic event in the pathogenesis of many human cancers. The fragile histidine triad (FHIT) gene, located at chromosome 3p14.2, is a tumour suppressor gene. Methylation results in inactivation of these genes.

Maruyama and colleagues looked at the relationship between the methylation status of various genes and survival in patients with NSCLC. Surgically resected specimens from 124 patients (79 men) of median (range) age 69 (25–82) years with NSCLC were studied using PCR. The majority had early stage disease (59 stage I, 24 stage II). The frequency of methylation was highest for E-cadherin (52%) followed by RAS association domain family protein (41%), FHIT and adenomatous polyposis coli (38%), retinoic acid receptor beta and H-cadherin (27%), p16^{INK4A} (20%), and O⁶-methylguanine-DNA-methyltransferase (0.8%). The 5 year survival rates of the patients with FHIT methylation-positive tumours was found to be significantly shorter (31.9% v 51.4%, p = 0.03), even in those patients with early stage disease (p = 0.007). However, the methylation status of other genes was not associated with any survival difference. In a multivariate analysis, FHIT methylation-positive status was found to be independently associated with poor survival (p = 0.046).

This study shows that methylation of FHIT is an independent factor associated with poor prognosis in patients with NSCLC.

K Pavithran

Consultant, Department of Medical Oncology, Rajiv Gandhi Cancer Institute & Research Centre,
New Delhi, India;
drkpavithran@hotmail.com