SARCOIDOSIS VASCULITIS AND DIFFUSE LUNG DISEASES 2011; 28; 130-138

Optimal scoring of serial change on chest radiography in sarcoidosis

C.J. Zappala[1, 4], S.R. Desai[2], S.J. Copley[3], P. Spagnolo[1], D. Cramer[1], D. Sen[1], S.M. Alam[1], R.M. du Bois[1], D.M. Hansell[1], A.U. Wells[1]

[1] Royal Brompton Hospital and National Heart and Lung Institute, London, UK; [2] King's College Hospital, London UK; [3] Hammersmith Hospital, London, UK; [4] Department of Medicine, University of Melbourne, Melbourne, Australia

Abstract. *Background:* The optimal means of quantifying change on chest radiography in sarcoidosis is uncertain. In current guidelines, the role of serial measurement of carbon-monoxide diffusing capacity (DL_{CO}) remains undefined and the prevalence of discordance between serial chest radiographic change and pulmonary function tends is unknown. *Objective:* To identify and explore key uncertainties in the monitoring of sarcoidosis by serial pulmonary function tests and chest radiography. *Design:* 354 patients with sarcoidosis and concurrent tests (chest radiography and PFTs within three months at baseline, two years and/or four years) were studied. Chest radiographs were assessed by two radiologists for changes in stage and disease extent. Radiographic change and pulmonary function trends were quantified and compared. *Results:* Change in radiographic extent of lung disease was always more frequent than change in stage ($p<0.0001$) and there was poor agreement between change in stage and change in radiographic extent (Kw = 0.21 at two years; Kw = 0.23 at four years). Change in disease extent on chest radiography was linked to PFT trends on analysis of variance ($p<0.0005$ for FEV_1, FVC, DL_{CO}), whereas change in radiographic stage was not. Changes in gas transfer were often isolated or discordant with other serial data. Discordance between pulmonary function data and chest radiographic data was observed in 50% of cases. *Conclusions:* Change in radiographic extent is more applicable to routine monitoring in sarcoidosis than change in radiographic stage. In future guidelines, the role of serial gas transfer estimation and reconciliation of divergent chest radiographic and functional trends might usefully be addressed. *(Sarcoidosis Vasc Diffuse Lung Dis 2011; 28: 130-138)*

Key words: sarcoidosis, chest radiograph, pulmonary function tests, Scadding stage

Introduction

The most recent international consensus document on the management of sarcoidosis states serial chest radiography and spirometry should be used to identify changes in disease severity (1). However, the method of evaluating serial chest radiographs was not specified and no recommendation was made on the need to monitor changes in gas transfer. The medical literature provides little guidance on either issue.

In large studies, change on chest radiography is quantified as change in Scadding stage (2-5) or as change in the extent of individual radiographic patterns (6-9), but neither approach captures the rapid side by side evaluation of chest radiographs, usual in clinical practice. The use of Scadding stage in serial

Received: 4 July 2010
Accepted after Revision: 10 March 2011
Correspondence: Athol U. Wells
Interstitial Lung Disease Unit,
Royal Brompton Hospital and NHLI,
Imperial College, Emmanuel Kaye Building,
1B Manresa Road, London SW3 6LP, UK
Tel. +44 207 351 8327
Fax: +44 207 351 8336.
E-mail: Athol.Wells@rbht.nhs.uk

monitoring is hampered by lack of consensus on what constitutes Stage IV appearances (5, 8, 10), as "evidence of fibrosis on chest radiography" might, in principle, range from end-stage appearances to indirect evidence of less advanced fibrosis. Until recently, analyses of overall global change in chest radiographic severity had been reported only in relatively small groups of patients (3, 6, 7, 8). In one recent larger series, change in global disease extent correlated more strongly than change in radiographic pattern with forced vital capacity (FVC) trends (11). Similarly, although the potential utility of serial DL_{CO} is recognised (3, 6, 8, 12, 13), the relative sensitivity of serial DL_{CO} and spirometric variables has not been examined and the significance of isolated change in DL_{CO} has not been explored. To date, serial pulmonary function and chest radiographic trends have not been satisfactorily integrated in clinical studies.

We address these issues by evaluating chest radiographic and pulmonary function trends in a large cohort of patients with pulmonary sarcoidosis. This study explores potential difficulties in integrating serial data in clinical practice by examining: 1) variations in the prevalence of stage IV disease with alterations in radiographic criteria; 2) methods of serial chest radiographic scoring in relation to serial pulmonary function trends; 3) the prevalence of isolated and disproportionate changes in gas transfer; and 4) the prevalence of discordance between serial variables.

Methods

We reviewed the data of 354 patients, collected between January 1980 and December 2005, who met current clinical and histopathologic diagnostic criteria for pulmonary sarcoidosis (1) and had agreed to participate in previously reported genetic studies (14, 15). Retrospective review of our clinical and radiographic data-base was approved by our institutional ethics committee. Clinical data (including smoking status, drug treatment, clinical findings) was extracted from case records. Patients were categorised as never smokers or ex- and current smokers. Treatment regimens consisted of prednisolone, alone or in combination with other immunosuppressive agents. As the aim of the study was to examine relationships between radiographic change and PFT trends, treatment effects were not specifically evaluated.

PFTs at baseline, two years (± six months) and four years (± six months) were tabulated. The two year time-point was chosen to reflect current guideline recommendations (1) that intense surveillance should be continued for up to two years to monitor either spontaneous regression or need for treatment. The four year time-point was chosen in order to evaluate longer-term trends. Analysed lung function measurements consisted of FVC and FEV_1 (PKM spirometer, P. K. Morgan, Kent, UK, or the Jaeger Compact system, Viasys Healthcare) and DL_{CO} and K_{CO} (using a single breath technique or a re-breathing technique with adjustment to single breath values on a PK Morgan respirometer). Results were expressed as percentages of predicted values (16). Significant serial trends, expressed as percentages of baseline values, were defined using current ATS criteria (FEV_1 > 10%; FVC > 10%; DL_{CO} > 15%; K_{CO} > 15%) and were condensed to a three point scale (1 = improvement; 2 = stability; 3 = deterioration).

Paired chest radiographs were assessed independently by two experienced chest radiologists. To simulate routine clinical practice, blinding to chronological order of the films was not undertaken and images were evaluated side-by-side. Radiographic stage was scored, using a liberal interpretation of Scadding's criteria of 'irreversible fibrosis' (5) such that any radiographic evidence of fibrosis was taken to denote stage IV disease. Changes in radiographic stage were condensed into a three point scale: 1 = improvement (a reduction in stage); 2 = stability; 3 = deterioration (an increase in stage). Change in radiographic extent was quantified using a simple three point scale: 1 = improvement (a reduction in extent); 2 = stability; 3 = deterioration (an increase in extent). Divergent observations were resolved by consensus.

Radiographic staging was repeated in a randomly selected sub-group of 172 patients at baseline and two years, with the use of more stringent evidence of 'irreversible fibrosis' (major volume loss and/or lung destruction) in the definition of Stage IV appearances. This was done in order to quantify the change in prevalence of stage IV disease with variation in staging criteria.

Relationships between radiographic change and PFT trends were evaluated in 218 patients with con-

current data (within one month), either at two years (n=192) or, when two year data was unavailable, at four years (n=26).

Data analysis

Inter-observer variation for radiographic data was quantified using the weighted kappa coefficient of agreement (Kw), with quadratic weighting. Weighted kappa was also used to examine concordance between three point scales for change in radiographic extent, change in radiographic stage and PFT change. The prevalence of change in radiographic extent and change in radiographic change was compared using the McNemar chi-squared test. The relationships between numerical changes in individual PFT (i.e. analysed as continuous variables) and the three point scales for change in radiographic extent and change in radiographic stage were evaluated by analysis of variance.

Results

Baseline demographics (Table 1)

354 patients met the diagnostic inclusion criteria (mean age 40.3y; SD=10.9y; range=16-75y; no gender predilection). Most patients had received treatment and were never smokers.

Table 1. Baseline characteristics (n=354)

Gender	Male-184 (52.0%)/Female-170 (48.0%)
Age at diagnosis	40.30±10.9 years
Treatment at diagnosis	Treated-209 (78.6%)/ Not treated-57 (21.4%)
Smoking	Never-223 (71.9%)/ Ex or Current-87 (28.1%)
Chest radiography stage	0-44 (12.4%) I - 53 (15.0%) II - 74 (20.9%) III - 20 (5.7%) IV - 163 (46.1%)
DL_{CO} %	74.9±19.6
K_{CO} %	90.5±16.8
FVC%	93.6±19.9
FEV_1%	84.2±23.1

Serial change in chest X-ray

There was good inter-observer agreement for chest radiography stage at baseline (Kw = 0.81), at two years (Kw = 0.82), and at four years (Kw = 0.80). Similarly, there was good inter-observer agreement for change in disease extent at two years (Kw = 0.76) and at four years (Kw = 0.78). As shown in Table 2, changes in radiographic extent [234/343 (68.2%) at two years; 185/248 (74.6%) at four years] were more frequent than changes in radiographic stage [66/343 (19.2%) at two years; 60/248 (24.2%) at four years], p<0.0001 for both time intervals (McNemar chi square test). Agreement between change in stage and change in extent was poor, both at two years (Kw = 0.21), and at four years (Kw = 0.23).

With the use of stringent criteria for stage IV radiographic appearances in a randomly-selected sub-group of 172 patients, the proportion of patients with stage IV disease fell from 80/172 (47%) to 41/172 (24%), with a corresponding increase in stage II and III appearances (Table 3). Agreement between radiologists was good in assessing chest X-ray

Table 2. Prevalence of serial chest radiography change (stage and overall extent of disease) at two and four years. Changes in radiographic extent [234/343 (68.2%) at two years; 185/248 (74.6%) at four years] were more frequent than changes in radiographic stage [66/343 (19.2%) at two years; 60/248 (24.2%) at four years], p<0.0001 for both time intervals (McNemar chi square test)

	Change in chest radiography stage	Change in chest radiography overall extent of disease
Baseline to two years (n = 343)	Improvement 31 (9.0%)	Improvement 144 (42.0%)
	Stable 277 (80.8%)	Stable 109 (31.8%)
	Deterioration 35 (10.2%)	Deterioration 90 (26.2%)
Baseline to four years (n = 248)	Improvement 30 (12.1%)	Improvement 104 (41.9%)
	Stable 188 (75.8%)	Stable 63 (25.4%)
	Deterioration 30 (12.1%)	Deterioration 81 (32.7%)

Table 3. Difference in chest radiograph stage before and after the application of end-stage criteria for stage IV chest radiographic appearances in a randomly-selected sub-group of 172 patients. The proportion of patients with stage IV disease fell from 80/172 (47%) to 41/172 (24%), with a corresponding increase in the proportion with stage

Baseline stage	Initial scoring (liberal definition of fibrosis in) stage IV	Post-hoc analysis scoring (strict criteria of end-stage lung for fibrosis in stage IV)
0	22 (13%)	25 (15%)
I	25 (15%)	28 (16%)
II	35 (20%)	53 (31%)
III	10 (6%)	25 (15%)
IV	80 (47%)	41 (24%)

stage (Kw = 0.89 at baseline; Kw = 0.85 at two years). Change in stage was less prevalent 29/172 (16.9%) than change in disease extent 90/172 (52.3%), p<0.0001. Agreement between change in stage and change in overall extent of disease on chest X-ray remained poor (Kw = 0.20).

Serial change in PFT (FEV1, FVC, DL_{CO})

There was significant change in serial PFT in 156/343 (45.5%) at two years and 144/248 (58.1%) at four years (Table 4). At both time intervals, observed change consisted of significant change in a single PFT variable in approximately 50% of cases, with change in all three variables seen in only approximately 20% of cases. At both time points, there were seven patients with discordant change in PFT parameters, including five at each time point with discordance between DL_{CO} and one or both spirometric parameters.

Amongst patients with a significant (>15%) change in DL_{CO} (Table 5), there was concordant significant change in FEV_1 and/or FVC in two thirds of cases. However, at both time intervals, change in DL_{CO} was associated with neither significant nor marginal concordant spirometric changes in approximately 20%, including some cases in whom change in DL_{CO} was associated with discordant spirometric change. Evaluation of serial changes in K_{CO} (DL_{CO}/VA) provided further evidence of uncoupling of gas transfer and spirometric trends. In over half of cases with significant changes in K_{CO} at both time intervals, spirometric trends were stable or discordant.

Chest radiographic versus PFT change at two to four years (n=218)

There was fair to moderate agreement between significant PFT trends and change in radiographic extent (Kw = 0.38). By contrast, agreement between PFT trends and change in radiographic stage was poor (Kw = 0.09). On analysis of variance, numerical changes in DL_{CO}, FVC and DL_{CO} (but not K_{CO}) were

Table 4. Prevalence of significant serial PFT change in DL_{CO}, FEV_1 and FVC. There was significant change in serial PFT in 156/343 (45.5%) at two years and 144/248 (58.1%) at four years. At both time intervals, observed change consisted of significant change in a single PFT variable in approximately 50% of cases, with change in all three variables seen in only approximately 20% of cases

	Improvement	Deterioration	Total Change	Stable	Discordant	Total
Two years	n=97	n=59	n=156	n=180	n=7	n=343
All three variables	19	8	27 (17%)			27 (8%)
Two variables	38	19	57 (37%)			57 (17%)
DL_{CO} only	11	12	23 (15%)			23 (7%)
FVC only	15	5	20 (13%)			20 (6%)
FEV_1 only	14	15	29 (19%)			29 (8%)
Stable				180		180 (52%)
Discordant					7	7 (2%)
Four years	n=76	n=68	n=144	n=97	n=7	n=248
All three variables	13	17	30 (21%)			30 (12%)
Two variables	34	13	47 (33%)			47 (19%)
DL_{CO} only	5	21	26 (18%)			26 (10%)
FVC only	14	4	18 (13%)			18 (7%)
FEV_1 only	10	13	23 (16%)			23 (9%)
Stable				97		97 (39%)
Discordant					7	7 (3%)

Table 5. Prevalence of significant change in gas transfer with either significant (>10%) or marginal (5-10%) change in spirometric values. Amongst patients with a significant (>15%) change in DL_{CO}, there was concordant significant change in FEV_1 and/or FVC in two thirds of cases. In over half of cases with significant changes in K_{CO} at both time intervals, spirometric trends were stable or discordant

	Two years	Four years
Significant (>15%) Change in DL_{CO} with:	n=75	n=82
Concordant (>10%) change in spirometry	47 (63%)	53 (65%)
Concordant (5-10%) change in spirometry	15 (20%)	11 (13%)
Stable (<5% change) spirometry	6 (8%)	9 (11%)
Discordant >10% change in spirometry	5 (7%)	4 (5%)
Discordant 5-10% change in spirometry	2 (3%)	5 (6%)
Significant (>15%) Change in K_{CO} with:	n=41	n=54
Concordant (>10%) change in spirometry	14 (34%)	19 (35%)
Concordant (5-10%) change in spirometry	6 (15%)	9 (17%)
Stable (<5% change) spirometry	4 (10%)	4 (7%)
Discordant (>10%) change in spirometry	12 (29%)	13 (24%)
Discordant (5-10%) change in spirometry	5 (12%)	9 (17%)

linked to change in radiographic extent (all p<0.005). By contrast, numerical PFT change was not linked to change in radiographic stage. This negative finding persisted with the use of stringent criteria for stage IV appearances.

Categorical PFT trends were concordant with changes in radiographic extent (i.e. either concordant change or concordant stability) in only approximately 50% of cases (Table 6). In the remaining cases, isolated change in radiographic extent was more frequent than isolated change in PFT, but in 6% of cases, radiographic extent and PFT exhibited change in opposing directions.

Table 6. Prevalence of patients achieving concordance/discordance between PFT and serial chest radiograph extent of disease. Concordance was considered parallel change in PFT indices and chest radiography score or concomitant stability of PFT trends and chest radiography score. Discordant serial change was considered significant change in PFT indices with opposite trends in overall chest radiography appearance. Categorical PFT trends were concordant with changes in radiographic extent (i.e. either concordant change or concordant stability) in only approximately 50% of cases

	Frequency (n=218)
Concordant change	59 (27%)
Concordant stability	48 (22%)
Isolated change in PFT	21 (10%)
Isolated change in chest radiography extent	71 (33%)
Discordant change	12 (6%)

Excluded 7 patients who had divergent significant change in one or more PFT indices

Discussion

The aim of our study was to identify key uncertainties in the monitoring of sarcoidosis by serial pulmonary function tests and chest radiography, and to highlight areas in which current guidelines provide insufficient guidance to clinicians. Five broad conclusions can be drawn from our results:

1. Although the prognostic value of chest radiographic staging is not questioned and is not re-examined in the current study, major variability exists in the identification of stage IV disease, reflecting the lack of precision in the radiographic definition of 'irreversible fibrosis'.
2. In routine monitoring, this problem is not crucial because our findings show that change in radiographic stage does not correlate with pulmonary function trends, and is clearly inferior in this regard to a simple statement of change in chest radiographic extent, on side by side evaluation of images.
6. Isolated change in DL_{CO} or change disproportionate to lung volumes is frequent in pulmonary sarcoidosis and consideration should be given to a recommendation on the need for serial DL_{CO} measurement in guidelines on serial monitoring.
4. Discordance between serial chest radiography appearance and PFT trends, not addressed in current guidelines, is sufficiently prevalent to

justify the construction of a guideline algorithm as an aid to routine management.

5. The degree of discordance between tests makes it clear that the accurate definition of change in pulmonary sarcoidosis will need to be multidisciplinary, integrating also changes in pulmonary symptoms.

The definition of stage IV chest radiographic appearances

Because no precise, widely agreed definition exists of stage IV disease on chest radiography, especially with regard to the distribution and extent of fibrosis (5, 8, 10), the relative prevalence of stage IV disease, and stage II and III disease, necessarily varies from institution to institution. There appear to be two broad considerations in the definition of stage IV appearances. Firstly, the increasing use of high resolution CT scanning in diffuse lung disease in general has provided insights into chest radiographic findings. Many chest radiologists have become increasingly comfortable with the detection of subtle signs of fibrosis on chest radiography and are ready to classify appearances as indicative of 'irreversible fibrosis', even when loss of lung volume is focal and lung destruction is only minor. However, other clinicians and radiologists retain a threshold for the designation of stage IV appearances when there is gross volume loss and/or fibrotic destruction, although the definition of such a threshold is necessarily nebulous. This disparity may account for the low inter-observer agreement on chest radiographic stage in one recent series (11). In the present study, we defined stage IV appearances using contrasting liberal and stringent criteria for the presence of significant fibrosis, and this resulted in a two-fold variation in the prevalence of stage IV disease, with major corresponding changes in the frequency of stage II and stage III disease. Although, as discussed below, this may not have important implications for routine monitoring, our findings suggest strongly that a tighter definition of stage IV disease is needed, in view of the continuing prognostic importance of the Scadding staging system.

Serial chest radiography in routine monitoring

In most large studies of serial chest radiography in sarcoidosis (2, 4, 8), the methodology of scoring change consisted of the assessment of change in the Scadding change of stage and a detailed assessment of the change in individual chest radiographic patterns. Neither method captures the rapid side by side assessment of overall change in chest radiographic extent of disease that is usual in routine practice, and it should also be acknowledged that the Scadding system was developed for the evaluation of prognosis at baseline and has never been validated as a means of quantifying serial change. Thus, guideline statements on serial chest radiographic monitoring are necessarily based on anecdotal clinical experience.

Our findings indicate that change in Scadding stage is misleading as a measure of change in disease severity. Although inter-observer agreement was good, change in stage bore no relationship to PFT trends, whether using liberal criteria for Stage IV disease, or using strict criteria of end-stage lung in a post hoc analysis. Furthermore, in a number of patients, major improvement in pulmonary function indices was associated with striking regression of disease on chest radiography with the disclosure of limited residual fibrotic abnormalities. In other patients with stage IV disease at baseline, major radiographic deterioration could not be identified based on change in stage. Thus, the Scadding system should not be used for routine monitoring, based on findings in the present study.

By contrast, change in overall radiographic extent was much more prevalent than change in stage and, more importantly, was more closely linked with change in serial DL_{CO}, FVC and FEV_1. These observations strongly support the broad approach taken by Muers and colleagues in quantifying changes in the extent of individual patterns, rather than change in radiographic stage in the British Thoracic Society cohort (8, 17). This approach was associated with greater inter-observer agreement than was seen with the full application of the International Labour Organization pneumoconiosis reading system in another recent series (9). However, the system proposed by Muers did not allow for side by side comparison of chest radiographs and is necessarily more complex than the rapid evaluation of global change in extent on paired radiographs. It can be argued that HRCT has now supplanted chest radiography for the delineation of individual parenchymal patterns, but radiation considerations and the lack of evidence for the

cost-effectiveness of serial HRCT preclude its routine use. The serial functional-morphological linkages observed in the present study, using a rapid, easily applicable and reproducible system that simulates routine practice, were as strong as those reported in the BTS cohort. Our findings confirm and extend the recent observation that change in global disease extent correlates more strongly than change in the reticulonodular component of the Muers score with FVC trends (11). Thus, more sophisticated approaches to the assessment of chest radiography do not necessarily enhance its usefulness in clinical practice.

Serial PFT in routine monitoring

In keeping with current guidelines, significant change in serial PFT was most commonly observed in spirometric volumes (1, 18). However, isolated change in DL_{CO} was seen in 15-20% of cases. This surprisingly high prevalence mirrors that found in the BTS study (17) and justifies reappraisal of the current recommendation that pulmonary function monitoring should consist of spirometric evaluation alone (1). A finding of isolated change in gas transfer may be especially useful when spirometric stability occurs in individuals associated with symptomatic or radiographic change. However, the use of gas transfer requires that measurements be repeated routinely, in order to ensure that any observed change is recent. Previous investigators have found that gas transfer levels correlate better than other functional variables with histologic features and gas transfer (9, 19, 20) and this may in part underlie our findings.

However, a more difficult question, not resolved by our data, is whether isolated or disproportionate gas transfer trends capture progression of interstitial lung disease or reflect a pulmonary vascular process. Pulmonary hypertension is well recognised in sarcoidosis and is present in up to 50% of patients with chronic exercise intolerance. (21, 22) Although the identification of a pulmonary vascular component, based on trends in gas transfer, is presumptive, the high prevalence of isolated changes in DL_{CO} and K_{CO} in our study, and the discordance between change in K_{CO} and changes in spirometric variables and chest radiographic findings, highlight the need for guidance on the need for the serial estimation of gas transfer in clinical practice.

The integration of PFT and chest radiographic trends: multidisciplinary diagnosis

Because the evaluation of global chest radiographic change has not been reported in large series, there has been no detailed assessment of the difficulties in reconciling monitoring data in routine practice. The quantification of potential difficulties in routine monitoring was one aim of our study and we identified a strikingly high prevalence of chest radiographic change that was isolated or discordant with PFT trends. EJ Potchen once observed that "the only utility of a diagnostic test is to reduce uncertainty". If this principle applies equally to monitoring, it can be argued that in many cases, serial chest radiography increases uncertainty and, thus, violates the "Potchen precept". The problem of isolated change in chest radiography in functionally stable disease has long been recognised as an occasional clinical difficulty. However, in the current cohort, change on chest radiography was isolated or discordant with pulmonary function trends in approximately 40% of cases at two years and at four years. Isolated changes on chest radiography are often disregarded by experienced clinicians, but recommendations on the optimal approach in this situation and on the more problematic issues of a) change in pulmonary function trends without chest radiographic change; and b) discordance between serial chest radiographic and PFT trends might usefully be addressed in future guideline statements.

The prevalence of discordance between pulmonary function and radiographic trends makes it clear that the optimal definition of change will need to be multidisciplinary. In order to reconcile these divergences, changes in pulmonary symptoms will need to be integrated but that process will need to be carefully thought through as in less advanced disease, progression will not necessarily be associated with symptomatic change. This problem aside, we were unable to integrate clinical data because of the retrospective nature of our study: review of case records did not provide a consistently reliable indication of changes in exertional dyspnoea or other symptoms over the exact time intervals between tests. Further work to develop a multidisciplinary algorithm to define change will need to be prospective and in order to ensure robust statements on symptomatic change over exactly designated time intervals,

a shorter time interval than two to four years will need to be evaluated.

Study limitations

The study cohort consisted of patients in a specialised clinic who had agreed to participate in genetic studies and this resulted in two possible selection biases. As patients were routinely asked to participate in genetic studies and the acceptance rate was very high, exceeding 95%, it appears unlikely that this resulted in undue bias. However, the specialised nature of the clinic and the continuation of follow-up for four years did undoubtedly select for a particular sub-set of patients with pulmonary sarcoidosis with more difficult disease requiring prolonged follow-up. We suggest that this resulted in "appropriate" selection bias, given that the purpose of our study was to highlight potential shortfalls in current guidelines, with respect to more taxing clinical problems. However, it should be stressed that the prevalence of monitoring problems that we identified applies to more difficult pulmonary sarcoidosis and not to unselected patients with pulmonary sarcoidosis.

A second problem was the inability to adjust for treatment, as the choice, timing and duration of individual therapies varied widely during follow-up. However, there is no logical reason to suppose that linkage between chest radiographic and functional change is likely to differ according to therapeutic status. The underlying assumption in our study is that the problem of detecting disease progression and reconciling conflicting chest radiographic and functional information remains the same problem whether suspected progression occurs on or off therapy.

In conclusion, we identify striking variability in the prevalence of stage IV appearances with variations in the radiographic definition of irreversible fibrosis and this has important implications for future population studies. A simple, user-friendly method for scoring change in radiographic extent provided linkage to pulmonary function trends, unlike change in radiographic stage, which is clearly unsuited to routine monitoring. Our findings indicate that in future guidelines for the management of pulmonary sarcoidosis, the role of serial gas transfer estimation in routine monitoring should be addressed and guidance is needed on how best to reconcile divergent chest radiographic and functional trends.

References

1. ATS/ERS/WASOG. Statement on Sarcoidosis. American Journal of Respiratory and Critical Care Medicine 1999; 160: 736-55.
2. Judson MA, Baughman RP, Thompson BW, et al. Two year prognosis of sarcoidosis: the ACCESS experience. Sarcoidosis Vasculitis Diffuse Lung Disease 2003; 20: 201-11.
3. Winterbauer RH, Hutchinson JF. Clinical significance of pulmonary function tests. Chest 1980; 78 (4): 640-7.
4. Sones M, Israel HL. Course and prognosis of sarcoidosis. American Journal of Medicine July 1960: 84-93.
5. Scadding JG. Prognosis of intrathoracic sarcoidosis in England: a review of 136 cases after five years' observation. British Medical Journal 1961; 2: 1267-78.
6. Lieberman J, Schleissner LA, Nosal , Sastre A, Mishkin FS. Clinical correlations of serum angiotensin-converting enzyme (ACE) in sarcoidosis. Chest 1983; 84 (5): 522-8.
7. Baumann MH, Strange C, Sahn SA. Do chest radiographic findings reflect the clinical course of patients with sarcoidosis during corticosteroid withdrawal? American Journal of Radiology 1990; 154: 481-5.
8. Muers MF, Middleton WG, Gibson GJ, et al. A simple radiographic scoring method for monitoring pulmonary sarcoidosis: relations between radiographic scores, dyspnoea grade and respiratory function in the British Thoracic Society study of long-term corticosteroid treatment. Sarcoidosis Vasculitis Diffuse Lung Disease 1997; 14: 46-56.
9. Judson MA, Gilbert GE, Rodgers JK, et al. The utility of the chest radiograph in diagnosing exacerbations of pulmonary sarcoidosis. Respirology 2008; 13: 97-102.
10. Emirgil C, Sobol BJ, Williams MH. Long-term study of pulmonary sarcoidosis. Journal of Chronic Disease 1969; 22: 69-86.
11. Baughman RP, Desai S, Drent M, et al. Changes in chest roentgenogram of sarcoidosis patients during a clinical trial of infliximab therapy: comparion of different methods of evaluation. Chest 2009; 136: 526-35.
12. Mihailovic-Vucinic V, Zugic V, Videnovic-Ivanov J. Hew observations on pulmonary function changes in sarcoidosis. Current opinion in pulmonary medicine 2003; 9: 436-41.
13. Sharma OP, Colp C, Williams MH. Pulmonary function studies in patients with bilateral sarcoidosis of hilar lymph nodes. Archives of Internal Medicine 1966; 117: 436-40.
14. Spagnolo P, Renzoni EA, Wells AU, et al. C-C chemokine receptor 2 and sarcoidosis: association with Lofgren's syndrome. American Journal of Respiratory and Critical Care Medicine 2003; 168 (10): 1162-6.
15. Spagnolo P, Renzoni EA, Wells AU, et al. C-C chemokine receptor 5 gene variants in relation to lung disease in sarcoidosis. American Journal of Respiratory and Critical Care Medicine Sep 2005; 172 (6): 721-8.
16. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Report of working party. Standardized lung function testing. Official statement of the European Respiratory Society. European community of steel & coal. European Respiratory Journal 1993; 16: 1-100.
17. Gibson GJ, Prescott RJ, Muers MF, et al. British Thoracic Society Sarcoidosis study: effects of long term corticosteroid treatment. Thorax 1996; 51: 238-47.
18. Baughman RP, Winget DB, Bowen EH, Lower EE. Predicting res-

piratory failure in sarcoidosis patients. Sarcoidosis Vasculitis Diffuse Lung Disease 1997; 14 (2): 154-8.
19. Huang CT, Heurich AE, Rosen Y, Moon S, Lyons HA. Pulmonary sarcoidosis. Roentgenographic, functional and pathologic correlations. Respiration 1979; 37: 337-45.
20. Colp C. Sarcoidosis. Course and treatment. Medical Clinics of North America 1977; 61 (6): 1267-78.
21. Baughman RP, Engel PJ, Meyer CA, Barrett AB, Lower EE. Pulmonary hypertension in sarcoidosis. Sarcoidosis, Vasculitis, Diffuse Lung Disease 2006; 23 (2): 108-16.
22. Bourbonnais JM, Samavati L. Clinical predictors of pulmonary hypertension in sarcoidosis. European Respiratory Journal Aug 2008; 32 (2): 296-302.