The chest radiograph in cystic fibrosis: a new scoring system compared with the Chrispin–Norman and Brasfield scores

S P Conway, M N Pond, I Bowler, D L Smith, E J Simmonds, D N Joanes, G Hambleton, EJ Hiller, D E Stableforth, P Weller, J M Littlewood

Abstract

Background – Scoring systems for the chest radiograph in cystic fibrosis are used to compare patients and different treatment regimens, and to monitor respiratory disease in individual patients. The Northern chest radiograph score was designed to allow one person to assess the radiological features of lung involvement in cystic fibrosis in as simple, rapid and equally reproducible manner as the established Chrispin and Norman, and Brasfield scoring systems.

Methods – Forty five chest radiographs were scored by 10 physicians with a special interest in cystic fibrosis according to the Brasfield and Northern methods, and by five pairs of physicians according to the Chrispin–Norman criteria. Three individuals and two pairs rescored the radiographs after an interval of 3–5 months. The Northern score was performed with and without a lateral view, using the original posteroanterior radiograph.

Results – The Northern score showed a better agreement between observers for the ranking of the radiographs. It was equally well related to respiratory function tests, the Shwachman–Kulczycki score of overall clinical status, and in its discrimination between different radiographs as the other two scoring systems. The Northern score performed equally well with or without a lateral film.

Conclusions – The Northern system fulfils the requirements of a chest radiograph score more successfully than the Chrispin–Norman or Brasfield systems, and does not require a lateral film.

(*Thorax* 1994;49:860–862)

North of England Cystic Fibrosis Units at Leeds, Birmingham, Manchester and Nottingham
S P Conway
M N Pond
I Bowler
D L Smith
E J Simmonds
G Hambleton
E J Hiller
D E Stableforth
P Weller
J M Littlewood

Department of Mathematics and Statistics, University of Leeds
D N Joanes

Scoring systems are widely used to monitor changes in the involvement of the lungs in patients with cystic fibrosis and for comparisons between patients and to assess different treatment regimens. The degree of pulmonary involvement is the main factor determining prognosis in cystic fibrosis.[1-3] The chest radiograph reflects past and current pathology[4 5] and is part of patient assessment, particularly in children.[6 7] The Chrispin–Norman chest radiograph score is a numerical score developed for children and is widely used.[5] The scoring is as follows: 0, absent; 1, changes present but not marked; and 2, marked changes. It categorises overexpansion on the lateral film, and bronchial line shadows, nodular shadows, ring shadows, and large shadows in each of four quadrants on the posteroanterior film. The maximum score is 38. It was designed to highlight changes from normal to abnormal but does not discriminate between moderate to gross changes. Chest radiograph scores must accurately reflect acute and chronic clinical status and lung function, and be reproducible within and between observers. Only the Brasfield score fulfils these criteria.[8] It scores air trapping on the lateral film, and linear markings, nodular cystic lesions, large lesions, and general severity on the posteroanterior film. Twenty five points represent a normal chest radiograph with lower scores indicating increasing disease severity. Because this system assesses the chest radiograph as a single unit, it lacks flexibility when features of cystic fibrosis are unevenly distributed in the lung fields.

This study describes a new uncomplicated chest radiograph score – the Northern score – designed for rapid assessment by a single physician. The Northern score was compared with the Chrispin–Norman and Brasfield scoring systems, and related to the Shwachman–Kulczycki clinical score,[6] an overall assessment which includes the chest radiographic appearance.

Methods

Forty five radiographs were selected to give three ranges of disease severity based on respiratory function tests performed on the same day as the radiograph. The ranges were an FEV_1 <40%, 41–60%, and >60% of predicted normal values in 15 cases each. The original radiographs were scored on the same day by 10 physicians with expertise in cystic fibrosis (six consultants and four research fellows) according to the Northern and Brasfield criteria. The Northern system was first scored on the posteroanterior view only, and a second score produced with the inclusion of a lateral film. Five pairs of physicians also scored the radiographs by the Chrispin–Norman method. After an interval of 3–5 months the radiographs were rescored by the three methods. Only three individual physicians and two pairs were available from the original group for this rescoring.

The Northern score is derived by dividing each lung into an upper and lower zone by drawing a horizontal line outwards from the

Table 1 Northern score for each lung quadrant

Scale	Radiological changes
0	Normal: no cystic fibrosis lung disease evident
1	Mild: minimal increase in linear markings and/or nodular cystic lesions up to 0·5 cm diameter
2	Moderate: more pronounced linear markings and/or more widespread nodular cystic lesions
3	Severe: prominent increase in linear markings, profuse nodular cystic lesions, large areas of collapse/consolidation
4	Very severe: little or no area of normal lung seen, dense infiltration
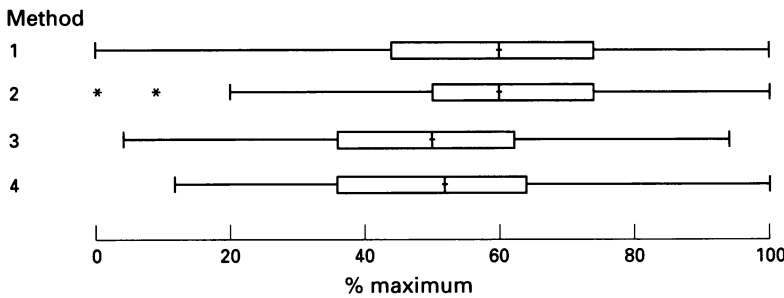
Boxplot of all assigned scores, expressed as % maximum score, for each scoring method. 1 = Northern, 2 = Northern + lateral, 3 = Brasfield, 4 = Chrispin–Norman. *Outliers. The scores are displayed as a percentage of the maximum score for each system so that comparison can be made on a single scale. The middle 50% of the data is shown as a box, the median marked by a cross within the box, and the lines either side of the box show the spread of the data.

Table 2 Statistical test results for (1) analysis of agreement of observers on ranking of radiographs, and (2) the ability of the scoring methods to show variability in radiographs

Scoring method	(1) Kendall's coefficient of concordance	(2) F ratio for radiographs*
Northern without lateral	0·770	45·49
Northern with lateral	0·781	46·16
Brasfield	0·733	36·99
Chrispin–Norman	0·844	28·86

* The higher the F ratio the greater is the variability detected between the radiographs.

Table 3 Correlations with respiratory function tests and Shwachman–Kulczycki (SK) score

Scoring method	Respiratory function tests					
	VC	FVC	FEV_1	log FEF_{25-75}	log PEFR	SK score
Northern without lateral	−0·74	−0·77	−0·87	−0·82	−0·46	−0·79
Northern with lateral	−0·74	−0·77	−0·87	−0·82	−0·45	−0·80
Brasfield	0·63	0·65	0·81	0·77	0·38	0·79
Chrispin–Norman	−0·71	−0·72	−0·83	−0·78	−0·38	−0·81

VC = vital capacity; FVC = forced vital capacity; FEV_1 = forced expiratory volume in one second; FEF_{25-75} = mid forced expiratory flow; PEFR = peak expiratory flow rate.

Table 4 Results of analysis of variance comparing consistency between and by observers (the "pure error" variance) for the four methods (with 95% confidence limits)

Scoring methods	Variance for radiographs	Variance for observers	"Pure error" variance
Northern without lateral	0·718 (0·478, 1·256)	0·097 (0·043, 0·408)	0·218
Northern with lateral	0·725 (0·483, 1·268)	0·091 (0·040, 0·382)	0·240
Brasfield	0·711 (0·476, 1·250)	0·052 (0·024, 0·229)	0·226
Chrispin–Norman	0·703 (0·443, 1·224)	0·173 (0·052, 2·083)	0·231

0–4 based on the increasing severity of linear, nodular cystic (up to 0·5 cm diameter) and large or confluent shadows (table 1). A further 0–4 points are allocated according to the observer's perception of overall severity. This allows an assessment of both acute changes and additional complications such as hilar lymphadenopathy, cardiac size, hyperinflation, pneu-

mothorax. The score ranges from 0 to 20, a higher score reflecting more severe radiological change.

DATA ANALYSIS

Kendall's coefficient of concordance was used as a numerical measure, on a 0–1 scale, of the degree of agreement between observers in their ranking of the radiographs. An analysis of variance was used to indicate the ability of the scores to discriminate between different radiographs. The result was expressed as an F ratio statistic, the greater the F ratio, the more the variability detected between the radiographs. The consistency in scoring within and between observers for the different methods was determined by an analysis of variance model which split the total variability in scoring into subject (radiograph), observer, and residual components (random unexplained variation). To give comparable results for each of the four methods standardised scores were analysed — that is, (score − mean)/standard deviation. This eliminated the effect of the different scales used by the different methods. Correlation coefficients were calculated to relate the mean scores to the Shwachman–Kulczycki clinical status score and to measures of respiratory function. Results for FEV_1 and FEF_{25-75} were log transformed to linearise the plots for these two measures.

Results

A boxplot display of the original scores shown in the figure illustrates for each system where the bulk of the distribution lies. The Northern score without consideration of the lateral film showed the widest spread of data. Both Northern scores rated the radiographs more severely than either the Brasfield or Chrispin–Norman score. This is likely to result from five points being allocated by the former to "complications," and eight by the latter to "large shadows" – scores that are rarely allocated.

The results of the statistical analysis are summarised in tables 2–4. For each analysis the Northern score performed equally well with or without a lateral film. Kendall's coefficient of concordance (table 2), a descriptive statistical term, cannot be interpreted in terms of significance, but shows the Northern score as intermediate for agreement between observers in their ranking of the radiographs. The higher value for the Chrispin–Norman score is not strictly comparable with the others, since it was based upon five paired, rather than 10 individual, observations. The F ratio is also a descriptive statistical term and cannot be interpreted in degrees of significance. The score of 46 for the Northern system, compared with 37 and 29 for the Brasfield and Chrispin–Norman systems respectively (table 2), suggests that any greater capacity of the former to reflect variability between radiographs is slight and unlikely to be clinically relevant.

Correlations with pulmonary function tests suggest that all the radiograph scores reflect the impact of pulmonary disease but were con-

sistently higher for the Northern score. These data are not independent of each other because they do not come from distinct populations. Standard tests of significance were therefore not applied to the differences between the correlation coefficients. The greater correlation values of the Northern score for all meassures of respiratory function suggest a difference. The correlations were virtually identical for all methods with the Shwachman–Kulczycki score.

Equal consistency in scoring occurred between observers for each of the scoring methods (table 4). An estimation of the effect of the radiographs and the observers on variability in scoring showed no differences.

Analysis of the scores of those physicians who provided replicated data showed a low "pure error" variance and equal consistency by observers for all scoring methods.

Discussion

The radiological appearances of the lungs can be interpreted in terms of the impact of underlying pathology,[4 5] and may provide information about patient responses to treatment.[2] A numerical chest radiograph score allows easy comparison between different radiographs but must fulfil certain criteria. The ideal chest radiograph scoring system needs to reflect chronic and acute radiological changes, clinical status, and respiratory impairment as shown by respiratory function tests. For repeated comparison within one patient and between different patients it must be consistent within and between observers. It must also have a sufficient range and sensitivity to differentiate between different degrees of mild, moderate, and severe disease.

This study showed the Northern score to be equal to the Chrispin–Norman and Brasfield scores in the consistency of scoring within and between observers, and in its reflection of overall clinical status as determined by the Shwachman–Kulczycki score. It was as sensitive as the other methods in the differentiation between radiographs of varying severity.

Excellent correlations between respiratory function tests and chest radiograph scores have been reported, confirming that the latter can reflect abnormalities due to pulmonary injury.[3 8–10] The Northern score consistently showed a slightly higher correlation with measurements of respiratory function, although this is unlikely to be clinically important. The importance of these statistical results is that they show that the much simpler and more rapid Northern score performs at least as well as the established scoring systems. The Northern score needs only one observer and does not necessarily require a lateral film. The additional 0–4 scale for overall assessment allows for changes and complications, both acute and chronic, and gives the score more flexibility.

The chest radiograph in cystic fibrosis provides information essential for patient management. The Northern chest radiograph scoring system fulfils the demands made of an effective scoring system. It can be readily learnt and applied immediately by a single physician working in a typical cystic fibrosis outpatient clinic.

1 Penketh ARL, Wise A, Mearns MB, Hodson ME, Batten JC. Cystic fibrosis in adolescents and adults. Thorax 1987; 42:526–32.
2 Doershuk CF, Matthews LW, Tucker AS, Spector S. Evaluation of a prophylactic and therapeutic program for patients with cystic fibrosis. Pediatrics 1965;36:675–88.
3 Mearns MB. Simple tests of ventilatory capacity in children with cystic fibrosis. Arch Dis Child 1968;43:528–39.
4 Hodson CJ, France NE. Pulmonary changes in cystic fibrosis of the pancreas. Clin Radiol 1962;13:54–61.
5 Chrispin AR, Norman AP. The systematic evaluation of the chest radiograph with cystic fibrosis. Pediatr Radiol 1974; 2:101–5.
6 Shwachman H, Kulczycki LL. Long-term study of one hundred and five patients with cystic fibrosis. Am J Dis Child 1958;96:6–15.
7 Taussig LM, Kattwinkel J, Friedewald WT, diSant'Agnese PA. A new prognostic score and clinical evaluation system for cystic fibrosis. J Pediatr 1973;82:380–90.
8 Brasfield D, Hicks G, Soong S, Tiller RE. The chest roentgenogram in cystic fibrosis: a new scoring system. Pediatrics 1979;63:24–9.
9 Reilly BJ, Featherby EA, Weng T, Crozier DN, Duic A, Levison H. The correlation of radiological changes with pulmonary function in cystic fibrosis. Radiology 1971;98: 281–5.
10 Matthew DJ, Warner JO, Chrispin AR, Norman AP. The relationship between chest radiographic scores and respiratory function tests in children with cystic fibrosis. Pediatr Radiol 1977;5:198–200.

BOOK NOTICE

Asthma and COPD in general practice. Studies on the quality of care. FW Dekker. (Pp 135). Department of General Practice, University of Leiden, The Netherlands. 1993.

In the Netherlands, as in the UK, there is a well developed system of primary care in which general practitioners play a major role in the care of patients with asthma and COPD with 80–90% of care for these two conditions taking place outside hospital. As in the UK, Dutch general practitioners fulfil a gatekeeper role, with responsibility for referral to secondary care.

This short but useful book, written by a Dutch academic general practitioner, is presented in the form of a thesis, with a literature review followed by a related series of studies from general practice in the Netherlands on the quality of diagnosis and treatment for asthma and COPD, with some assessment of outcomes. The book provides a useful reference manual on topics for interested general practitioners, but some of the studies are several years old and not all are of direct relevance to UK general practice.

Chapter 2 provides useful summaries of the issues surrounding definitions of diagnosis for asthma and COPD and quality of care. It brings out nicely the contrasts between the technical aspects of care for asthma and COPD, and the "interpersonal exchange" which characterises primary medical care. Chapter 3 consists of a review of studies of quality of care in asthma and COPD, but would be more useful if restricted to those which focused on outcome measures rather than process. There is a quite distinct (and from the UK viewpoint, incongruous) chapter on the evaluation of a desktop in vitro allergy diagnostic test. There are no data presented to indicate that this (presumably expensive) test is superior to careful history taking or skin prick testing in the diagnosis of bronchial allergy.

Somewhat dated chapters on the usefulness of peak flow meter readings and compliance with medication are followed by more interesting studies on patient self-management, and the quality of life in patients with COPD.

The final chapter provides a general discussion on the studies reported in the book, and concludes that there is much room for improvement in the general practice care of asthma and COPD – a message of equal relevance to the UK. Four developments in quality of care are highlighted: the emergence of guidelines and quality standards for management; the important role of practice nurses in patient education; and the potential roles of patient self-management plans and patient satisfaction surveys in care of asthma and COPD. As in the UK, all of these developments are being espoused enthusiastically, although we await convincing evidence of their effectiveness.

The book is a useful addition for general practitioners with a special interest in the subject, but not one which offers important new messages for the true generalist for whom a quick scan of the five page summary chapter will be sufficient. – SH

CORRECTIONS

The chest radiograph in cystic fibrosis

In the paper entitled "The chest radiograph in cystic fibrosis: a new scoring system compared with the Chrispin–Norman and Brasfield scores" by S P Conway *et al* which appeared on pages 860–862 of the September issue a line of text was inadvertently omitted. The last paragraph on page 860 should read: "The Northern score is derived by dividing each lung into an upper and lower zone by drawing a horizontal line outwards from the middle of each hilum. Each quadrant is scored 0–4 based on the increasing severity of linear, nodular cystic (up to 0·5 cm diameter) and large or confluent shadows (table 1)."

Survival of patients with severe α_1-antitrypsin deficiency

In the paper entitled "Survival of patients with severe α_1-antitrypsin deficiency with special reference to non-index cases" by N Seersholm *et al* which appeared on pages 695–698 of the July issue the labelling of the keys of figures 1 and 2 was reversed. The figures are reproduced here with the keys correctly labelled.
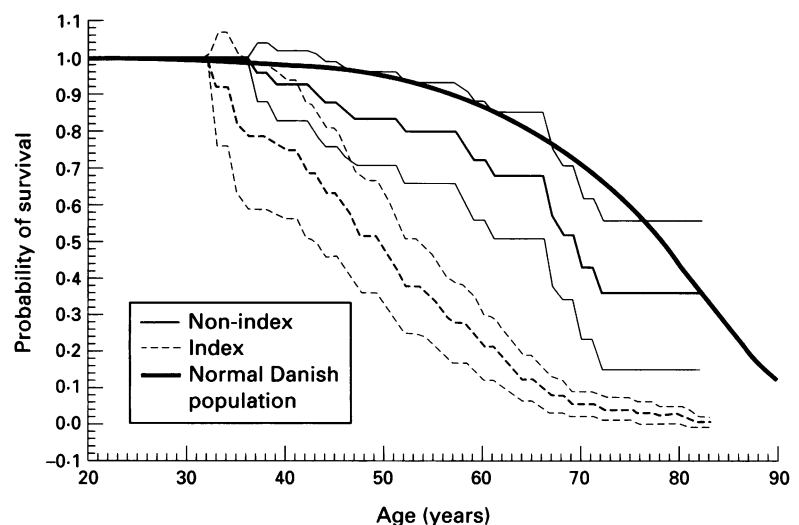

Figure 1   Cumulative probability of the survival time of index cases and non-index cases with 95% confidence intervals. Survival of the normal Danish population is shown for comparison.
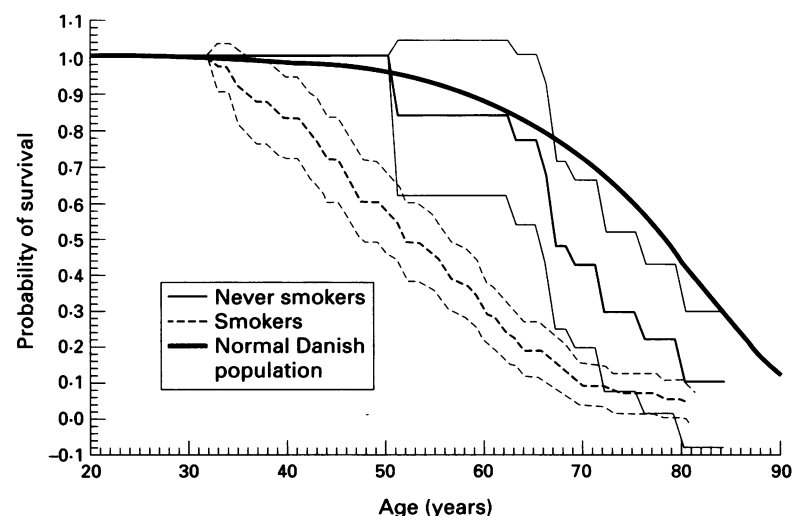

Figure 2   Cumulative probability of the survival time of smokers and non-smokers with 95% confidence intervals. Survival of the normal Danish population is shown for comparison.