

- 4 Accarino A, Richart C, Lucas A, Pedreira J D, Alomar C. Neumonia intersticial linfoidea y periarteritis nodosa. *Rev Clin Esp* 1983; 171: 363-4.
- 5 Homma S. A histopathological study on vascular and interstitial changes in the lungs of collagen diseases. *Japanese Journal of Thoracic Disease* 1985; 23: 332-47.
- 6 Pankow W, Pohl-Reuling B, Wichert P. Ausgeprägte Lungenbeteiligung bei Periarteritis nodosa. *Med Klin* 1988; 83: 231-4.
- 7 Carratalá J, Vidaller A, Maña J, Manresa F, Fernandez-Nogues F. Polyarteritis nodosa associated with idiopathic pulmonary fibrosis: report of two cases. *Ann Rheum Dis* 1989; 48: 876-9.
- 8 Matsumoto T, Yoshimine T, Shimouchi K, et al. The liver in systemic lupus erythematosus: pathological analysis of 52 cases and review of Japanese autopsy registry data. *Hum Pathol*. In press.
- 9 Katzenstein A-L A, Askin F B. *Surgical pathology of non-neoplastic lung disease*. Philadelphia: Saunders, 1990: 35-9.
- 10 Shiokawa Y. Summarized report of studies. *Annual report of malignant rheumatoid arthritis and periarteritis nodosa research committee of Japan*. Tokyo: Ministry of Health and Welfare of Japan, 1975: 21-2.

Radiographic assessment of the knee joint in OA

Sir: The article of Cooper *et al*, on reproducibility of radiological assessment of the knee joint in osteoarthritis,¹ considers a subject integral to the science of rheumatology. The surprisingly poor reproducibility of the cardinal radiological sign of osteoarthritis (osteophytes)² in this Bristol Study contrasts with the observations of Altman's 15 centre study.² Experienced rheumatologists took part in the latter, whereas four of the five observers in the Bristol study were trainees. The poor reproducibility from three of the five observers raises an intriguing possibility. Does the study actually consider the effectiveness and reproducibility of rheumatologists in interpreting radiographs, or solely the efficacy of a particular training programme in teaching the skills in question?

Failure to observe differential reproducibility of joint space assessment between medial and lateral knee compartments was most perplexing. This suggests that the Bristol team observers were no more accurate in assessing the 'normal' compartment of the joint than the 'affected' compartment. Combination of data from 'normal' and osteoarthritic compartments seems even to under-

mine the original premise that this was a study of osteoarthritis.

If Samuel Clemens (Mark Twain) is correct that statistics can be considered as one of the various forms of lies, the failure to achieve statistical significance may be misleading. The blinded nature of the analysis may not overcome an apparent fundamental statistical premise violation, related to non-random selection of who reviewed which films. β Error (related to limited sample size) and phenomenon infrequency (making use of χ^2 inappropriate) further compromise interpretation.

The comments on observations of a difference in order of appearance and of clustering of findings and use of the peculiar ill-defined phrase 'abnormality of bony contour'¹ raise an intriguing possibility, as only osteophytes have actually been demonstrated as a reliable radiological sign of osteoarthritis.² Do the differential appearance order and clustering indicate that perhaps we have been grouping phenomenon with osteoarthritis that are at best tangential? Are we dealing with the traditional overlapping circles of phenomenon which so punctuate the epidemiological chapters in McCarty's text?³

BRUCE M ROTHSCHILD
The Arthritis Center of
Northeast Ohio
5701 Market Street
Youngstown
Ohio 44512, USA

- 1 Cooper C, Cushnaghan J, Kirwan J R, Dieppe P A, McAlindon T, McCrae F. Radiographic assessment of the knee joint in osteoarthritis. *Ann Rheum Dis* 1992; 51: 80-2.
- 2 Altman R, Alarcon G, Appelrouth D, et al. Criteria for classification and reporting of osteoarthritis of the knee. *Arthritis Rheum*. In press.
- 3 Masi A T, Medsger T A Jr. Epidemiology of the rheumatic diseases. In: McCarty D J, ed. *Arthritis and allied conditions*. 11th ed. Philadelphia: Lea and Febiger, 1989: 16-54.

AUTHORS' REPLY: We were interested to read Dr Rothschild's comments on our study of radiological assessment of the knee joint in osteoarthritis. We agree that the reproducibility of these assessments is a subject integral to rheumatological research and practice. Although we cannot be expected to comment on the American College of Rheumatology study which is still in press, we are uncertain that the reproducibility with which osteoarthrosis was assigned in our study was markedly different from that reported using a

similar scale by Altman *et al*.² The lowest intraclass correlation between observers for spurs in that study was 0.56, as compared with a κ statistic of 0.58 in our study. Indeed, we emphasised in our paper that assessment of joint space narrowing and osteophyte in the tibiofemoral compartment performed rather better than did sclerosis and cyst formation. Dr Rothschild is correct that one of the reasons why reproducibility in our study might have been generally worse than in others relates to the observers chosen for the study. One of our specific aims, however, was to examine the assessment of these features in routine practice rather than to ascertain the optimum performances which could be attained by highly trained observers. Indeed, we highlighted the potential role of training to reduce observer variation as an area for further research.

We were unable to understand Dr Rothschild's statistical comments. If he is lamenting the widespread use of statistical inference in biomedical research, space does not permit us to do more than profoundly disagree with him. The study was designed so that all the observers viewed all the radiographs unaware as to their own or colleagues' previous assessments. The statistical analysis of observer variation is controversial, but we chose to use the κ statistic (rather than the χ^2 statistic as your correspondent implies). We took pains to emphasise the pitfalls of this analytical technique, but ultimately decided that a summary index of the concordance for various measurements justified its use. It remains one of the most widely used tools to assess the reproducibility of categorical scales.³

CYRUS COOPER
MRC Environmental Epidemiology Unit
Southampton General Hospital
Southampton SO9 4XY
United Kingdom

JOHN KIRWAN
PAUL DIEPPE
Rheumatology Unit
Bristol Royal Infirmary
Bristol BS2 8HW

- 1 Cooper C, Cushnaghan J, Kirwan J R, Dieppe P A, McAlindon T, McCrae F. Radiographic assessment of the knee joint in osteoarthritis. *Ann Rheum Dis* 1992; 51: 80-2.
- 2 Altman R D, Fries J F, Bloch D A, et al. Radiographic assessment of progression in osteoarthritis. *Arthritis Rheum* 1987; 30: 1214-25.
- 3 Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ* 1992; 304: 1491-4.