

## Short Report

### Kellgren & Lawrence grade 1 osteophytes in the knee—doubtful or definite?

D. J. Hart and T. D. Spector

*The Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, London SE1 7EH, U.K.*

#### Summary

The most universally accepted methods of classification of radiographic osteoarthritis (OA) is the Kellgren & Lawrence (K & L) grading system. However, there have been many criticisms of this including the grading of a doubtful osteophyte as a grade 1. We report the natural history of a given prevalent grade 1 osteophyte in a 10-year follow-up knee study. 62% of women graded at baseline with a doubtful osteophyte went on to develop true osteophytic knee OA compared with only 22% of controls with no sign of disease. So called 'doubtful' osteophytes appear to be 'real' and significantly related to OA knee and therefore cannot be ignored or classified as normal. © 2002 Osteoarthritis Research Society International. Published by Elsevier Science Ltd. All rights reserved.

*Key words:* Osteoarthritis, Radiographic classification, Epidemiology.

Epidemiological studies of osteoarthritis (OA) thrive on correct classification of individuals being with or without disease. The K & L grading system has been accepted for many years as a gold standard in classification. However, there has been criticism over its inconsistencies<sup>1</sup> and no clear consensus in the interpretation of grade 1 osteophytes<sup>2</sup>. Is it correct to classify these people as a normal group, or treat them as an affected group with early changes of OA? It has also been suggested that since no clear consensus exists as to whether grade 1 subjects are cases or controls, they should be treated as a separate grade. This could be achieved by either excluding grade 1 from the analysis or treating them as a separate subgroup. These uncertainties often lead to these patients being excluded from trials of early disease or dropped from the analysis resulting in considerable loss of power. We investigated whether K & L grade 1 was a reliable indicator of OA or future disease in a longitudinal population based study.

In 1990, 1000 women were recruited to the Chingford Study at baseline and had AP weight-bearing knee radiographs taken. After 10 years women were invited to follow up and 800 of the original cohort were successfully recalled and had repeat radiographs.

All women who had a baseline K & L grade 1 were re-read at year 10 to check if the grading was still the same. Using the Kellgren & Lawrence atlas, 79 women had baseline radiographs graded as K & L grade 1 (presence of doubtful osteophyte) in at least one knee (10%), 68 women (8.5%) were unilateral, being grade 0 in the other knee and 11 had grade 1 in both knees (1.4%) giving 90 knees for

analysis. Another 90 women with baseline knee radiographs graded 0 were drawn from the wider cohort and selected as controls. Analysis was performed on the 180 knees. Follow-up year 10 knee radiographs ( $N=180$ ) were read blind to baseline grading to determine rates of change from grade 1 at baseline compared with controls. The reader was not blind to time sequence of the radiographs, but this method has been used in other studies<sup>3</sup> and it has been shown that blinding readers to sequence may increase the error rate. In previous longitudinal reading unblinded to sequence reading we have good reproducibility with kappas ranging from 0.80<sup>4</sup> from this cohort.

A single experienced observer with previously proven good reproducibility using the Kellgren & Lawrence method in grading knee radiographs was used to read the films. Intraobserver kappa statistics were assessed in 40 paired radiographs 10 years apart to determine reproducibility of changes of grade 1 and grade 2 measurements over time. Intraobserver kappas were good and similar for both grades; for grade 1 was 0.86 and for grade 2 was 0.83.

Women were well matched between cases and controls for age (57.8 (4.07) vs 56.4 (5.78)  $P=0.09$ ); weight (67.6 (10.20) vs 68.4 (11.70)  $P=0.58$ ); height (1.62 (5.5) vs 1.61 (6.2)  $P=0.94$ ) and knee symptoms (26% vs 22%  $P=0.55$ ).

To assess the reliability of grade 1 assessments over time we assumed that OA knee osteophytes are unlikely to regress radiographically and we therefore checked numbers with grade 1 at baseline returning to 0 at year 10 (Table I). From 90 women with baseline grade 1 only five (5.5%) cases reverted to a grade 0 at 10-year follow-up. As a comparison we examined the radiographs of 96 women with grade 2 K & L (presence of a definite osteophyte without definite joint space narrowing) at baseline in either knee to see how many had reverted to a grade 1 K & L at follow-up. Only 2% of the grade 2 knees had reverted to a grade 1 over 10 years. The stability and therefore reliability

Received 15 August 2002; accepted 3 September 2002.

Address correspondence to: Dr D. Hart, The Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, Lambeth Palace Road, London SE1 7EH, U.K. Tel.: +44 20 79 28 9292; Fax: +44 20 79 22 8154; E-mail: [deborah.hart@kcl.ac.uk](mailto:deborah.hart@kcl.ac.uk)

Table I  
Outcome of baseline K & L grades between controls and Grade 1 at 5 and 10-year follow-up by number and percentage

K & L Grade	Controls Grade 0 at baseline		Cases Grade 1 at baseline	
	Grade at year 5	Grade at year 10	Grade at year 5*	Grade at year 10
0	80 (89%)	60 (67%)	9 (10%)	5 (6%)
1	6 (7%)	10 (11%)	40 (47%)	29 (32%)
2	3 (3%)	18 (20%)	31 (37%)	46 (51%)
3	1 (1%)	2 (2%)	5 (6%)	10 (11%)

\*Five women did not have year 5 radiograph.

of the grade 1 was quite good over time with only 5.5% cases regressing, only slightly worse than for grade 2, 10 years later.

At 5 years 40% of individuals with a baseline grade 1 had progressed to at least grade 2, indicating traditionally accepted definite radiographic disease (Table I). By 10 years this number rose to 62%. This compares in the controls with only 4% at 5 years, and 22% by 10 years. If a grade 1 were to be considered the same as grade 0 for OA, i.e. considered 'normal', then of those with a grade 1 or less at baseline a little over one-third remained 'normal' at 10 years (38%). In the control group, however, over three-quarters (79%) remained graded as normal (0/1) (Table I). Furthermore grade 1 knees that progressed were associated with obesity [OR 3.84 (95%CI 1.60–9.21)] which suggests they behave the same way as grade 2 knees [OR 2.38 (95%CI 1.29–4.39)] as shown by previous analysis in this cohort<sup>5</sup>.

There were no obvious differences in frequency of baseline knee pain between the two groups. However of those grade 1 cases going on to develop a grade 2 K & L, 31% had experienced knee pain lasting more than a month compared to 19% of those remaining at grade 1 ( $P=0.16$ ). In the control group 58% of those developing grade 2 K & L experienced knee pain compared to 21% having symptoms in those who remained disease free. Of those developing grade 2 disease 22% had also developed joint space narrowing, but this rate was similar to the 21% who had narrowing in those remaining at grade 1, showing this was not a clear predictor.

The conclusions of this study are that in epidemiological studies using this grading system that grade 1 disease is not grouped with grade 0 as a non-case. Grade 1 cases should be treated either as an early disease sub-group or excluded from analysis in case control studies of knee OA. The group acts similarly to grade 2 in terms of risk factors

for progression and pain and has reasonable reliability. Grade 1 K & L is only based on the presence of osteophytes, however both cross-sectional and longitudinal studies have shown that in the knee, osteophytes may be a more reliable predictor than a grade 1 narrowing as an indicator of early disease<sup>3,6</sup>. There may also be a therapeutic potential in using this sub-group in clinical trials in the early prevention of OA progression, as they are currently excluded.

## References

1. Spector TD, Cooper C. Radiographic assessment of osteoarthritis in population studies: whither Kellgren & Lawrence? *Osteoarthritis Cartilage* 1993;1:203–6.
2. Spector TD, Hochberg MC. Methodological problems in the epidemiological study of osteoarthritis. *Ann Rheum Dis* 1994;53:143–6.
3. La Valley MP, McAlindon TE, Chaisson CE, Levy D, Felson DT. The validity of different definitions of radiographic worsening for longitudinal studies of knee osteoarthritis. *J Clin Epidemiol* 2001;54:30–9.
4. Hart DJ, Cronin C, Daniels M, Worthy T, Doyle DV, Spector TD. The relationship of bone density and fracture to incident and progressive radiographic osteoarthritis of the knee: The Chingford Study. *Arthritis Rheum* 2002;46:92–9.
5. Hart DJ, Doyle DV, Spector TD. Incidence and risk factors for radiographic knee osteoarthritis in middle-aged women. *Arthritis Rheum* 1999;42:17–24.
6. Spector TD, Hart DJ, Byrne J, Harris PA, Dacre JE, Doyle DV. Definition of osteoarthritis of the knee for epidemiological studies. *Ann Rheum Dis* 1993;52:790–4.