Attempt to classify patients with arthritis of the hip suitable for prosthetic replacement, and their femoral heads

P. D. BYERS, B. A. ROPER,* AND B. GLENNIE
From the Department of Morbid Anatomy, Institute of Orthopaedics, Royal National Orthopaedic Hospital, 234 Great Portland Street, London W1N 6AD

Byers, P. D., Roper, B. A., and Glennie, B. (1975). Annals of the Rheumatic Diseases, 34, 298–302. Attempt to classify patients with arthritis of the hip suitable for prosthetic replacement, and their femoral heads. Using clinical, radiological, and morbid anatomical data about patients with arthritis of the hip and their femoral heads resected during total hip replacement, an attempt has been made first, to classify patients and femoral heads by a cluster analysis technique, and secondly and more realistically, to use existing clinical and radiological classification as a basis for seeking discriminating features in the data from the femoral head. Both types of analysis failed to give results, and led to a conclusion that arthritis of the hip is a spectrum of disease. This may not pose problems to the clinician, but it does to the investigator studying the pathogenesis of osteoarthrosis. This problem is briefly discussed.

Investigation into degenerative arthritis has steadily mounted over the past 10 years, as the increased use of hip prostheses has made available for study numerous arthritic femoral heads. There is uncertainty about the exact clinical diagnosis in these cases, however, even though prosthetic replacement is indicated. Appeals to the morbid anatomist are often in vain since his perceptions are confused by uncertainty about some basic questions: is the susceptibility to osteoarthrosis (OA) of the unaffected hip joint of a patient with rheumatoid arthritis (RA) altered?; is the rheumatoid process so pervasive as to modify the pathogenesis of OA in such a joint?; does OA secondary to RA mean that features of both diseases are present?; does the surface fibrous tissue that arises during the natural progression of OA have features that distinguish it from the pannus of RA?

As a first step towards answering these questions, an attempt was made to classify on a simple comparative basis more than 100 femoral heads resected because of arthritis (P. D. Byers, T. A. Farkas, and V. Fornasier, 1968, unpublished), using intact heads and fine-detail x-rays of 3 mm slabs. The results were unsatisfactory in that no stable groups could be established by one observer on successive occasions, nor by two observers working independently. Much of the difficulty lay in the complex inter-related nature of the changes. Therefore it became necessary to add clinical observations of the patients to the data. However, guidance from clinical and radiological studies is difficult to obtain due to lack of standardization in data collection and to the fact that methods of evaluation and reliability vary. It became necessary then to standardize the data from patients as well as specimens.

It is commonplace now to regard the pathogenesis of osteoarthrosis as complex, with a range of possible starting points and routes. Kellgren (1961) hypothesized that osteoarthrosis is a mixture of diseases, and this was used to form the basic hypotheses of this study, namely:

1. That a collection of standard data from the history, physical examination, and radiology of patients with arthritis of the hip, and of the resected femoral heads, could be analysed to reveal unsuspected groupings and their principal characteristics.

2. That current diagnostic categories of arthritis represent specific disease entities whose pathogenesis effects changes in the femoral head, whereby the diseases can be recognized.

These hypotheses are independent, and are placed in the order in which they were conceived, and attempt...
in the words of Harris (1970), to 'restructure a system already articulate in some degree', and as such may well be in conflict at some points.

In the planning and in the initial stages of execution we were advised and given considerable practical assistance by Dr. P. H. N. Wood and the members of the ARC Epidemiology Research Unit, Manchester.

#### Material

The class of patient studied was that with arthritis of the hip suitable for prosthetic replacement admitted by several clinicians to the Stanmore section of the Royal National Orthopaedics Hospital during 1969. Selection from among this group was dependent upon the time available to the clinical investigator. Complete data were collected from 89 patients.

#### Methods of data collection

Questionnaires with coded answers for recording on computer punch cards were drawn up for clinical history, physical examination, and radiological examination of the patient, and for gross examination and fine-detail slab x-rays of the femoral head.

(A) CLINICAL HISTORY AND PHYSICAL EXAMINATION

The history questionnaire was based primarily on the New York criteria for rheumatoid arthritis, gout, and ankylosing spondylitis (Bennett and Wood, 1968). It was concerned with frequency and distribution of pain, stiffness, and swelling of small joints, and with detailed features of frequency, distribution, and cause of hip pain, as well as response to analgesics.

In the physical examination details were recorded of soft tissue swelling, pain, tenderness, subluxation, and contractures of all joints and their range of movement. On completion of the history and physical examination clinical diagnosis was recorded as certain or probable for one of the categories.

(B) RADIOLOGY

X-rays of the hands, feet, and pelvis were used, and the joint changes were classified using the Atlas of Standard Radiographs (Kellgren, Jeffrey, and Ball, 1963). After studying the x-rays the examiner recorded the diagnostic category without the option of certain or probable. The following categories list the features by which the clinical investigator was most influenced.

1. **Primary monarticular osteoarthritis**
   Disease in one hip, satisfying radiological criteria for OA.

2. **Primary polyarticular osteoarthritis**
   Disease in either both hips, or one hip and one or both knees, and/or terminal joints of fingers or carpometacarpal joint of thumb. Affected joints show pain and limitation of movement and satisfy radiological criteria for OA.

3. **Secondary osteoarthritis**
   Disease in the hip with clear evidence of predisposing mechanical cause, e.g. congenital dislocation of the hip.

4. **Rheumatoid arthritis**
   Disease in the hip associated with repeated episodes of pain, swelling, and stiffness of several joints of hands, especially in early morning, with swelling, tenderness, and limitation of movement. Radiological criteria of RA are satisfied, and all A.R.A. exclusions are negative.

5. **Mixed osteoarthritis/rheumatoid arthritis**
   Criteria of RA in hands and feet, and either OA alone in hip or features compatible with both OA and RA in x-rays of hip.

6. **Others**
   Evidence of some other generalized disorder of joints in addition to OA changes, e.g. gout.

#### Morbid anatomy

No histological sections of either synovium or femoral heads were used, and the work was based on the gross appearance of the intact head and the fine-detail x-ray of as many 3 mm slices as could be obtained using a standard method. Synovium was not included because the lack of control over sampling, the nonspecific character of the changes, and problems of grading were judged likely to diminish the value to the point where the effort was not worthwhile. The use of a fluorescent antibody method for rheumatoid factor was not considered, but this factor was looked for in the patient's serum. In the femoral heads the findings with whose interpretation we are concerned are reflected in the gross appearance; the small samples provided by histological sections add little to the information obtained by studying the whole head.

A total of 53 observations on each femoral head and 46 on the fine-detail x-ray of 3 mm slabs cut in a standardized way formed the basis of the anatomical investigation. These were concerned with the shape and size of the head; the size and distribution of osteophytes; the size, location, and other characteristics (i.e. protruding or sunken fibrous tissue, their amounts and proportions) of any exposed bone; the amount of residual cartilage, and the proportion that was unchanged; the number, size, and distribution of osteolytic foci; and the extent of sclerotic bone.

#### Observer error

Observer variability in clinical and radiological diagnoses was not studied. Gross appearances of the femoral heads were recorded by one observer, and another observer examined the fine-detail x-rays; both observers repeated observations on 14 samples, giving a measure of intraobserver variation. Both registered a difference between assignment to positive and negative categories in 5% of the 700 observations; in a further 19% of observations there were variations in the grading of severity, the error in exact measurements being of the order of 10%, though differences of 30% were occasionally recorded.

#### Methods of testing hypotheses and results

**AGREEMENT BETWEEN CLINICAL AND RADIOLOGICAL DIAGNOSES (TABLE)**

Since the clinical and radiological diagnoses were made independently, an essential step in the analyses, which was also an indirect test of the hypotheses, was to compare the two sets of diagnoses to assess the
agreement between them. It was found that the two agreed in only 50% of the cases. Surprisingly, the x-rays did not reveal more polyarticular disease than did the clinical study. There was a strong tendency for radiology to identify features predisposing to secondary osteoarthrosis in cases where the clinical impression favoured monarticular osteoarthrosis.

There was some uncertainty of clinical diagnosis in roughly 20% of the cases. However, when this 20% of cases was removed the agreement between the two sets of diagnoses changed only slightly, to 55%.

**Cluster Analysis: Test of Hypothesis 1**

It was decided to attempt a grouping of the data by cluster analysis (on advice from Dr. Wood). High-speed electronic computers have made possible the consideration of large numbers of characteristics in classification, and have revived questions about the merit of one classification as opposed to another (Sokal, 1966). A selection of several different types of cluster analysis were chosen from a suite called Clustan (Wishart, 1972). However, no consistent groupings were formed by the cluster analyses. This technique has not been very successfully used in medicine so far, and therefore the value of these negative attempts with the Clustan program is uncertain.

**Discriminant Analysis: Test of Hypothesis 2**

Subsequent to the failure of cluster analysis to formulate any meaningful classification of arthritis, it was suggested that the data should be subjected to a discriminant analysis. A Stepwise Discriminant Analysis program was used—BMD07M (1971 revision).

Given the six diagnostic groups the program carries out a stepwise analysis, each variable being examined in turn for its discriminatory value. The process is then repeated with combinations of the variables, variables being added according to the improvement they contribute to differentiation, much as is done in multiple regression. This process determines the importance of each variable in distinguishing between the diagnostic groups. The method then derives two canonical variables, which are the mathematical combinations of the original variables which best distinguish between the groups. Next a graph is produced in which the canonical variables are the two axes, and points are plotted to represent each specimen and the mean of each group (Figure). The extent of overlap of the groups is also shown.

Finally, the specimens are reclassified to different groups, when necessary, on the basis of position on the two-dimensional plot.

The method was applied to the two independent classifications, of clinical and radiological diagnoses, for each of the two sets of observations, i.e. gross examination of the femoral heads and fine-detail x-rays. As the clinical diagnoses were made with varying degrees of probability, only those cases with a certain diagnosis were used.

The most striking finding is that the analyses failed to provide consistent and definitive discrimination between the groups. It is noteworthy, too, that whereas well-defined groups were identified on the two-dimensional plot (Figure), these were very small, consisting of three or at most five specimens. Nevertheless, secondary OA and the residual group ‘other’ were well-defined on each of the two analyses based on clinical diagnoses. This suggests that disorders affecting the hip joint that are not due primarily to local OA or RA can be distinguished from these latter groups.

An additional discriminant analysis was carried out on data from only those specimens diagnosed clinically as primary OA, both mono- and polyarticular, or as RA, and restricted to the four most important variables identified when relating fine-detail x-rays to clinical diagnosis. Thirty-five per cent. of specimens were reclassified to different groups, and the overlap...
between all three diagnostic groups was considerable, so that no useful conclusions resulted.

Conclusion and discussion

This study fails to support the idea that a collection of standard clinical and radiographic data on patients with hip arthropathy could be analysed to reveal unsuspected groupings. This does not refute the idea that there may be unsuspected groupings within the material, but indicates only that the methods used and the data collected did not show this possibility. In particular, the failure of the discriminant analysis based on radiological diagnosis and fine detail x-ray is assumed to be due to the restricted information in the latter as compared with the former. It may be that acetabular changes influence the radiological diagnosis, and that these are not reflected in the changes in the head as seen in fine-detail slab x-rays, because either they are independent of head alterations or the data have been inadequately collected.

As has already been noted, the value of cluster analysis in medicine remains indeterminate, though perhaps our failures were related to limitations in the Clustan suite of programs rather than to shortcomings in the basic approach. Discriminant analysis has a better defined and more limited objective, so that the negative results from this approach may appear to be more important. However, there are no a priori grounds for assuming that any particular method can be the arbiter on the question of heterogeneity; one's efforts must be heuristic, and undue weight cannot be attached to negative findings unless a variety of approaches have been tried.

Since in a collection of femoral heads recognizable types of head and very characteristic features are seen, it is surprising that discriminant features between different groups could not be identified by numerical methods. If the method of analysis was acceptable, then the validity of the data and method of collection should be questioned. However, a great deal of care was devoted to the observations on anatomical and fine-detail x-ray features. While this is no guarantee of the worth of the data, satisfactory earlier experience with similar observations (Byers, Contepomi, and Farkas, 1970) indicate that the observations themselves were reasonable. The possibility remains, therefore, that certain key features may be missing, be they observations on aspects not so far perceived or adjustments of the observations made in the light of age, sex, or some other variables, perhaps to be derived from histological study.

The study also failed to support the idea that current diagnostic varieties of hip arthropathy represent specific disease entities leading to characteristic appearances of the femoral head; only if this were true could one claim to identify the disease on the basis of the observations made on the specimens. The failure in this part of the work highlights the difficulties of classifying disease, and the overlapping of the diagnostic groups also suggests that one may in fact be dealing with a spectrum of disease, rather than with a number of discrete disease entities.

However, before concluding that there may be a continuum of pathological changes, the origin of the data must be reappraised. Clinicians seem to recognize patients with types of hip disease who are suitable for prosthetic joint replacement, but often there are
difficulties in specifying what the disease is, thus imposing severe limitations on pathological studies since the interpretation of the latter must necessarily relate to clinical observations.

This draws attention to the need for caution. Many nonclinical investigators undertake elaborate analyses of cartilage from resected joint material classified on clinical and radiological grounds. It must be that the uncertainties of diagnosis and classification have a bearing on the outcome of this research (Byers, 1971). Clearly, research workers must study cases with whose diagnosis they are completely satisfied, and on grounds they can specify, to obtain meaningful and comparable results. Such a requirement may appear to create an impasse: meaningful pathology is dependent on meaningful clinical classification, and the latter is at present open to question. But there are clinical classifications that can be made with some confidence, a grouping of patients according to whether they do well or not after joint replacement or osteotomy. We have initiated a study in the hope that patterns of disease can be found with which the appearances of the femoral head will correlate.

Difficulties with classifying patients notwithstanding, the type of disease is but one variable in the categorization of material for cartilage research. Types of cartilages are recognizable in resected femoral heads: the fibrillated zone around the exposed bone; the cartilage of the osteophytes; varying amounts of grossly unaltered cartilage; and cartilage undergoing changes other than fibrillation. In addition, tissues are formed in the area of exposed bone as the disease develops, which may variously be regarded as reactive or reparative (Meachim and Osborne, 1970). The development of micromethods (Mankin, 1974) makes it more realistic to expect the actual material analysed to be in some way classified.

Dr. P. H. N. Wood (ARC Epidemiology Research Unit) offered encouragement, advice, and technical assistance at the outset of this study, and valuable criticism of the results and their presentation, all of which served to make it viable. Peter Fayers (MRC Statistical Unit, University College London), Mr. Brian Newman (Computer Unit, Middlesex Hospital Medical School), and Mr. David Wishart all helped. Mr. Richard Coles and Mr. Bill Heard provided skilful and essential technical help. To each we extend our thanks and appreciation. The work was supported by the Arthritis and Rheumatism Council.

References


Meachim, G., and Osborne, G. V. (1970) J. Path. 102, 1 (Repair at the femoral articular surface in osteo-arthritis of the hip)


Wishart, D. (1972) 'Clustan 1A Manual'. Programmed Library Services Publication No. 8, University of Edinburgh