Psoriatic arthritis (PsA) is a distinct chronic inflammatory arthritis associated with psoriasis. There are a variety of clinical phenotypes, resulting in historic variability in case definition, potentially confounding research. The original classification criteria were developed by Moll and Wright\textsuperscript{1}; however, despite the proposal of a number of other criteria\textsuperscript{2}, none have been universally accepted\textsuperscript{3}. In response to this, the CASPAR study group (ClASsification criteria for Psoriatic ARthritis) developed classification criteria specifically for use in clinical research\textsuperscript{4}. Their large prospective multinational study found that the CASPAR were simple to use, and demonstrated a sensitivity of 0.914 and specificity of 0.987. Subsequent studies confirmed these findings\textsuperscript{5,6,7,8}.

The CASPAR criteria set can be used with confidence in prospective trials; however, its performance when used retrospectively has not yet been assessed. Much of the observational research in PsA relies on cohorts established prior to the development of the CASPAR criteria. It is essential that such cohorts can be appropriately classified. We have set out to investigate the feasibility and performance of the CASPAR criteria when used to retrospectively classify an existing research cohort.

**MATERIALS AND METHODS**

The CASPAR criteria were applied to 480 patient records from the Royal National Hospital for Rheumatic Diseases PsA cohort and 100 consecutive controls with inflammatory arthritis from a general rheumatology clinic. The CASPAR criteria consist of confirmed inflammatory arthritic disease (joint, spine, or entheseal) with at least 3 points from the following features: current psoriasis (assigned a score of 2 points; all other features are assigned a score of 1), a history of psoriasis or a family history of psoriasis (unless current psoriasis is present), dactylitis, juxtaarticular new bone formation (hands or feet), rheumatoid factor (RF) negativity (except latex test), and psoriatic nail dystrophy. We adapted this score for retrospective use from medical records; “inflammation” and “current psoriasis” were scored if they had ever been confirmed in the PsA clinic. To optimize feasibility, once 3 points had been
scored the records were not scrutinized further. Two raters (WT and LC) reviewed all records. An interrater validation exercise was undertaken on 20 records. The diagnostic “gold standard” was expert clinical diagnosis by a clinician with longstanding expertise in PsA. In cases of doubt where the clinical diagnosis was of PsA and did not fulfill CASPAR criteria, the records and radiology were reviewed in full by WT, LC, and NM for consensus. Two sensitivity and specificity analyses were performed. The first excluded those with a clinical diagnosis of PsA but who did not have sufficient data available to confirm the classification with the CASPAR criteria. The second analysis was performed on all records, including those with missing data.

RESULTS

Nine case records were missing or destroyed, leaving 471 cases from the PsA cohort. On review, 456 cases had PsA. The diagnosis had changed over time in 15 cases and they were entered into the control group. Of the 115 controls, 96 had rheumatoid arthritis, 5 osteoarthritis, 3 reactive arthritis, 3 seronegative arthritis, 3 undifferentiated arthralgia, 2 ankylosing spondylitis, 1 spondyloarthropathies, and 2 systemic sclerosis (Table 1). We found 4% of the control group had a current or personal history of psoriasis, but none had a clinical diagnosis of PsA or fulfilled the CASPAR criteria. This is consistent with the current prevalence estimates for psoriasis of between 0.6% and 4%9. Twenty-eight (4.6%) of the records were reviewed by LC, WT, and NM for consensus assessment.

Of the 456 physician diagnoses of PsA, 8 did not fulfill the CASPAR criteria. Only 1 had all the CASPAR data available. A further 7 had either no radiographs or no RF tests performed to enable complete assessment. Sensitivity was 99.7% and specificity was 99.1%. Sensitivity remained high even after inclusion of the 7 case records with missing data (sensitivity 98.2%, specificity 99.1%).

The objectivity of the CASPAR assessment resulted in excellent interrater reliability. Of the 20 records reviewed, there was full agreement in 18 cases. Both raters identified the same 2 records as cases of doubt for consensus review. Both raters found the CASPAR to be easy and quick to use. It was found that 52% of the PsA cases fulfilled criteria based on current psoriasis and a negative RF, both readily accessible from the case records.

DISCUSSION

The CASPAR criteria could be applied prospectively to established research cohorts, but that would take years. The ability of a classification tool to be used retrospectively is therefore important. We have found the CASPAR criteria to be easy and practical for 3 reasons. First, the majority of cases achieve the required 3 points quickly from inflammation, current psoriasis, and a negative RF. Second, the objectivity of the criteria made reliability high with minimal training. Third, removing the requirement to scrutinize the records after sufficient points had been reached further improved the feasibility. Missing data points are a reality of observational research; however, we have found the sensitivity and specificity to be similar to those in previous reports even when data points are missing.

This study is potentially weakened by 3 factors. First is the lack of a gold standard, which has been a concern in prior reports. Our approach of the experienced physician’s diagnosis is a recognized method, and was the approach taken in previous studies. Second, through modification of the CASPAR criteria to allow “current psoriasis” and inflammation to be scored at any time rather than at the same clinical assessment, we have potentially increased the sensitivity. The initial sensitivity of the criteria was 91%, reported by Taylor, et al; however, subsequent reports have evaluated sensitivity between 97% and 100%, limiting the degree to which we can overestimate it. Further, specificity is of greater importance when classifying a research cohort. Finally, we have not compared the performance of CASPAR to that of other criteria. These comparisons are already in the literature and the purpose of our study was to assess whether the feasibility and performance characteristics were maintained when used retrospectively.

Our study demonstrated that the feasibility, specificity, and sensitivity of the CASPAR are maintained when adapted for retrospective use to classify an established research cohort from medical records, even in those with missing data.

REFERENCES


