An analysis of 14 molecular markers for monitoring osteoarthritis. Relationship of the markers to clinical end-points

Objective: To investigate whether any of 14 serum and urine molecular markers (MMs) used to monitor osteoarthritis (OA) would be associated with particular clinical end-points.

Design: Thirty-nine OA patients were bled and urine collected at five time points: at baseline visit and at visits 1, 3, 6 and 12 months later. Twelve clinical measurements were made and the concentrations of each of 14 MMs were determined. Principal component analysis, stepwise linear regression with backward elimination, and logistic regression were used to determine the correlations between MMs and clinical measures.

Results: Principal component analysis was used to reduce the 12 clinical measurements into three independent clinical clusters: baseline clinical assessments, changes in clinical assessments and signal joint measurements. The 14 MMs were similarly reduced to five MM clusters. Each of the three clinical clusters was correlated with a single but different MM cluster. Baseline clinical assessments were correlated with bone markers typified by hydroxylysyl pyridinoline (HP) crosslinks, swelling of the signal joint was correlated with inflammation markers, especially CRP, and the change in clinical assessments over the 1 year evaluation was correlated with TGFβ1. There was no correlation between any of the skeletal markers and the clinical measures, a situation which draws attention to the need for a direct assessment of cartilage damage in OA to validate the use of cartilage markers.

Conclusions: This study demonstrates statistical methodology for analysis of clinical trials using multiple MMs and clinical end-points. The patient numbers are sufficient to test hypotheses of relationships of single MMs such as CRP, TGFβ1 and HP to clinical measures, but larger clinical trials are needed to validate hypotheses. © 2001 OsteoArthritis Research Society International

Key words: Biological markers; Osteoarthritis; Clinical correlations.

Introduction

One of the pressing problems in osteoarthritis (OA) research is the definition of molecular markers (MMs) that will facilitate determination of patients’ prognosis and need of treatment. There are a large number of candidate MMs and we have been exploring the utility of 14 of them. They have been measured in an archival set of serum, plasmas and urines from patients in the placebo group (NSAID-permitted) of a completed clinical trial. It is commonly recognized that the concentrations of cartilage markers in synovial fluid, being closer to their site of origin, may be better correlated with pathologic changes in cartilage metabolism. Nevertheless, this study focused on MMs in blood and urine—markers that can be readily sampled longitudinally and that are amenable to large scale clinical trials—in anticipation that changes in these markers might reflect changes in clinical status.

The MMs that we have so far examined can be divided into skeletal or inflammation markers. Skeletal markers are related to cartilage or bone metabolism. They are keratan sulfate (KS), the C-propeptide of type II procollagen (CPII), bone sialoprotein (BSP), cartilage oligomeric matrix protein (COMP), cartilage proteoglycan aggrecan fetal epitope (epitope 846), and the collagen cross-links hydroxylysyl pyridinium (HP) and lysyl pyridinium (LP). Disease markers related to inflammation that were analysed were C-reactive protein (CRP), tumor necrosis factor receptor type I (TNF-RI) and type II (TNF-RII), interleukin 6 (IL-6), hyaluronan (HA), TGFβ1, and eosinophil cationic protein (ECP).
variables.

obtain individual single marker correlations with clinical correlations. Those correlations were deconvoluted to

MMs obtained from principal component analysis to extract this study we used reduced sets of clinical measures and
cation of statistical methods to a dataset is to generate and clinical
lation. Therefore, in seeking a relationship between MMs

clinical variables could be similarly reduced and an optimal

power of the technique suggested to us that the set of

divided into

in a previous paper we found that the 14 MMs could be

set of variables into clusters where each cluster is made up of variables that change in a similar manner. For example,
in a previous paper we found that the 14 MMs could be divided into five independent clusters of markers. The

power of the technique suggested to us that the set of clinical variables could be similarly reduced and an optimal

method of finding meaningful data might be through exami-
nation of the correlations between the reduced sets.

A rigorous statistical analysis of the relationship between

MMs and clinical data is necessary in order to identify one

or a set of MMs, the rise and fall of which might reflect changes in disease activity. Such methods are necessary for

finding relationships among variables such as clinical measures of disease, patient responses and MMs, changes which display significant intrinsic variation. Finally, in working with a finite patient population, results may or may not represent an adequate sample of the true popu-
lation. Therefore, in seeking a relationship between MMs and clinical findings, the primary value of an initial appli-
cation of statistical methods to a dataset is to generate hypotheses that can be tested through further experimenta-
tion and examination of additional patient populations. In

this study we used reduced sets of clinical measures and MMs obtained from principal component analysis to extract

correlations. Those correlations were deconvoluted to obtain individual single marker correlations with clinical

variables.

Patients

Patients were selected with Grades I to III OA of the knee or hip by the ACR OA criteria24,25. Five males and 34 females with mean age 57, range 45–61 and a primary diagnosis of OA were entered into the study. The knee (31 patients) or hip (eight patients) as appropriate was designated the signal joint. The majority of patients had additional joint involvement. Each patient was removed from therapy (NSAID) for 1 week prior to the baseline (t=0) visit. Five milliliters of blood and a 24 h urine specimen were taken at baseline. After the baseline visit, patients were placed on piroxicam 20 mg/day, and blood and urine specimens collected at visits 1, 3, 6, and 12 months later. A control cohort was also formed consisting of 13 females and eight males who were without joint pain and with a mean age of 50 years and an age range of 45–63 years.

Patient status was determined at each visit. Two physi-


cian clinical assessments and four patient visual analog

assessments were made at each visit. These measures are

listed in Table I with the range of values allowed for each measure. Serum and urine were aliquoted and stored at

−72°C until assayed at Pfizer or shipped to other labora-
tories for assay. Patients were required to have serum and urine taken at all five measurement periods in order to
determine stability and changes in the markers over a 1-year period.

Methods

Keratan sulfate (KS) and carboxypropeptide of type II procollagen (CPII) were measured by competitive radio-
imunoassay (RIA). Cartilage oligomeric matrix protein (COMP), bone sialoprotein (BSP), and aggrecan fetal

epitope (epitope 846) were measured by Elisa. C-reactive protein (CRP), tumor necrosis factor receptor type I (TNF-

RI), tumor necrosis factor receptor type II (TNF-RII), inter-

leukin 6 (IL-6), transforming growth factor beta 1 (TGF-β1), eosinophil cationic protein (ECP), hyaluronic acid (HA)

were measured by commercial Elisas. Hydroxylysyl pyridi-
noline (HP), lysyl pyridinoline (LP) and creatinine were

analyzed biochemically. Each assay was carried out as

previously detailed2.

Statistical analyses were carried out with SAS (SAS

Institute, Cary, NC) and Statistica (StatSoft Inc, Tulsa, OK)
software. Possible associations in clinical data at baseline

were explored by principal component analysis with

orthomax factor rotation, and by Spearman rank order
correlation. Clinical data were grouped into clusters of

related measures by principal component analysis just as

MM data were grouped into clusters previously2. For all

statistical tests that require normal distributions, the use of

logarithms of all independent variables except age was

necessary. Associations between MM clusters and clinical

measurement clusters were determined by stepwise linear

regression using an open procedure that permitted all

possible associations of MM clusters with each clinical

measurement cluster. A combined addition and elimination

method was used until only statistically significant associ-

ations remained. At each step the most non-significant

variable was eliminated and the regression recomputed.

Each of the variables previously removed as non-

significant was placed back into the linear regression to

ensure that it had not become significant with the removal

<table>
<thead>
<tr>
<th>Disease measurement (Range of values)</th>
<th>Patient values geometric mean</th>
<th>Patient values range</th>
<th>Patient change geometric mean†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient self-assessment of disease activity (1–5)</td>
<td>4</td>
<td>2–5</td>
<td>-2.6</td>
</tr>
<tr>
<td>Physician’s global assessment of disease activity (1–5)</td>
<td>3</td>
<td>2–4</td>
<td>-2.2</td>
</tr>
<tr>
<td>Pain at rest (1–5)</td>
<td>2</td>
<td>3–17</td>
<td>-1.2</td>
</tr>
<tr>
<td>Pain on weight-bearing (0–17)</td>
<td>13</td>
<td>3–17</td>
<td>-5.7</td>
</tr>
<tr>
<td>Stiffness (0–17)</td>
<td>10</td>
<td>0–17</td>
<td>-6.8</td>
</tr>
<tr>
<td>Soft tissue swelling (0–3)</td>
<td>1</td>
<td>0–3</td>
<td>-0.8</td>
</tr>
</tbody>
</table>

*Clinical parameters were based on the scoring by the physician for the physician global assessment scale, and soft tissue swelling of the knee. The remainder of the assessments were made by the patients, who globally rated their disease from mild to severe (patient self-assessment) or rated other parameters of their disease using visual analog scales.

†Geometric means of changes were calculated as follows: \( R = \log((X + \Delta X)/X) \), GeoMean(\( \Delta X \)) = GeoMean(\( X \)) * exp(Mean(\( R \))), \( X \) = clinical measure, \( \Delta X \) = change.

Table I

Clinical measures with their range of values*
allowable values, severity. Y Y Y
the concentration of the MM at the half point between PRINCIPAL COMPONENT ANALYSIS OF THE CLINICAL PARAMETERS
Results
where using a logistic regression model using the equation:

\[ Y = \left( \frac{Y_{\text{max}} - Y_{\text{min}}}{1 + (C_{50}/X)^H} \right) + Y_{\text{min}} \]

Y is the clinical value of the parameter, \( Y_{\text{max}} \) and \( Y_{\text{min}} \) are respectively the maximum and the minimum allowable values, \( X \) is the concentration of the MM, \( C_{50} \) is the concentration of the MM at the half point between \( Y_{\text{max}} \) and \( Y_{\text{min}} \), and \( H \) is the maximum slope of the sigmoid curve. Overall statistical significance was computed using an F-ratio in Statistica. Logistic regression was used because of the non-continuous truncated nature of the clinical measures. To follow up the regression, Fisher's exact test was used to determine the significance of the separation of the patients into a 2 x 2 layout: marker level greater or less than the \( C_{50} \), and of greater or lesser clinical severity.

Results

**PRINCIPAL COMPONENT ANALYSIS OF THE CLINICAL PARAMETERS**

Six measures of clinical disease were determined for each patient at baseline and at each of four subsequent visits over a 1-year period. In addition, changes from baseline were calculated at each subsequent visit for each of the six clinical measures, resulting in a total of 12 clinical variables (six status and six change measures). Principal component analysis was used to reduce the number of clinical variables. There was a highly significant separation of the 12 clinical variables into three independent factors. The principal components of each factor and their levels of significance are listed in Table II. The principal components of Factor 1 are measurements for five parameters at baseline (patient global assessment, physician's global assessment, pain on weight-bearing, pain at rest and stiffness). Each measurement is a numerical severity score. The principal components of Factor 2 are the changes in status of each of the five parameters in Factor 1 over the 1-year period. The principal components of Factor 3 were measurements in which the signal knee was evaluated at each visit and both the status and changes in that joint recorded. For clarity in discussing the analysis, the principal components of the three factors are referred to as clusters and called respectively: baseline clinical, change in clinical, and signal joint.

**Association of markers clusters with clinical clusters**

In the related paper\(^2\), a principal component analysis was carried out on 14 MMs in controls (nonOA) and OA patients in this study. Five factors were found to describe best the variation of the MMs. Because the principal components of these five factors were logically segregated, the principal components of each factor could be given a descriptive name: inflammation markers, bone markers, cartilage synthesis markers, cartilage degradation markers, and TGFβ. The possibility of a meaningful association between each of the five MM clusters at baseline and each of the three clinical clusters going forward was explored by open stepwise linear regression\(^*\). In effect all potential combinations of the five MM clusters were explored to see which, if any, sets of markers were correlated with each clinical cluster. All combinations were tested and the most non-significant set was removed with a forward/backward procedure. This process was repeated until only significant associations remained. For each clinical cluster a significant correlation was found with only one MM cluster. The baseline clinical status was associated with the bone markers; the change in clinical was associated with TGFβ, and the signal joint measures were associated with the

*There were no associations between changes in MMs and clinical status or change variables.

---

<table>
<thead>
<tr>
<th>Clinical measures</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain on weight-bearing</td>
<td>0.793</td>
<td>-0.187</td>
<td>-0.043</td>
</tr>
<tr>
<td>Stiffness</td>
<td>0.756</td>
<td>-0.081</td>
<td>0.077</td>
</tr>
<tr>
<td>Patient's self-assessment of disease activity</td>
<td>0.705</td>
<td>0.023</td>
<td>-0.296</td>
</tr>
<tr>
<td>Pain at rest</td>
<td>0.665</td>
<td>-0.242</td>
<td>-0.086</td>
</tr>
<tr>
<td>Physician's assessment of disease activity</td>
<td>0.620</td>
<td>-0.207</td>
<td>0.231</td>
</tr>
<tr>
<td>Change in disease—physician's assessment</td>
<td>0.022</td>
<td>0.813</td>
<td>-0.209</td>
</tr>
<tr>
<td>Change in pain on weight-bearing</td>
<td>-0.455</td>
<td>0.775</td>
<td>0.116</td>
</tr>
<tr>
<td>Change in stiffness</td>
<td>-0.318</td>
<td>0.753</td>
<td>-0.135</td>
</tr>
<tr>
<td>Change in pain at rest</td>
<td>0.104</td>
<td>0.752</td>
<td>0.393</td>
</tr>
<tr>
<td>Change in disease—patient self-assessment</td>
<td>-0.293</td>
<td>0.625</td>
<td>0.111</td>
</tr>
<tr>
<td>Change in soft tissue swelling—signal joint</td>
<td>0.172</td>
<td>0.211</td>
<td>0.886</td>
</tr>
<tr>
<td>Soft tissue swelling—signal joint</td>
<td>0.266</td>
<td>0.151</td>
<td>-0.842</td>
</tr>
</tbody>
</table>

Variance due to factor‡§ 3.03 3.00 1.89

*Clinical values (see Table I) from each OA patient were utilized to determine the correlation matrix. Change values were taken between baseline and the 1 year assessment.

†Principal component analysis coefficients. Matrix rotation and determination of the number of factors were carried out in SAS with the orthomax software and the requirement that the eigenvalues must be greater than one for significance. Three factors were found.

‡Isolation of independent factors. The primary components in each Factor are shown in bold.

§The three factors together account for 66% of the total variance.

Table II

Clinical measurements at baseline and their change at one year*
inflammation markers. The associations and the level of significance are shown in Table III. Neither of the two cartilage MM clusters (cartilage synthesis cluster or cartilage degradation cluster) showed an association with the clinical marker clusters.

Baseline clinical and bone markers

The bone marker cluster consisted of measurements of three markers, hydroxylysyl pyridinoline (HP), lysyl pyridinoline (LP) and bone sialoprotein (BSP). When the three markers were examined individually for their association with baseline clinical, only pyridinoline gave a highly significant positive correlation ($P<0.001$). The logistic regression parameters for equation 1 (Materials and methods) are $Y_{max}=4$, $Y_{min}=3$, $C_{50}=5.15$, $H=8.0$; $N=34$, $R=0.668$, $P<0.001$. Using the mid-point of the logistic regression, a Fisher's exact test gave significance of $P=0.0032$. To gain a better understanding of the nature of the correlation, the values for pyridinoline were examined as a function of age (Fig. 1). The majority of patients (and controls) fall within the bounds of 2–6 μg HP/μg creatinine. In our series, a significant number of OA patients over age 53 have values greater than 6 and the individuals with the highest HP values (≥6) all had the more severe grade of OA and all but one were female.

Change in clinical vs TGFβ1

A single molecular marker, TGFβ1, was found to have a statistical association with the change in clinical cluster ($P<0.5$). To gain a better understanding of its relationship to clinical parameters a further principal component analysis was conducted. All six clinical variables at baseline, 1, 3, 6,

### Table III

<table>
<thead>
<tr>
<th>Clinical cluster</th>
<th>Marker cluster</th>
<th>Slope coefficient</th>
<th>$R^2$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline clinical</td>
<td>Bone markers</td>
<td>0.653</td>
<td>0.462</td>
<td>0.0004</td>
</tr>
<tr>
<td>Change in clinical</td>
<td>TGFβ1</td>
<td>0.410</td>
<td>0.170</td>
<td>0.050</td>
</tr>
<tr>
<td>Signal joint</td>
<td>Inflammation</td>
<td>−0.711</td>
<td>0.321</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*Possible associations between the three clinical clusters and the five marker clusters, individually or combinations, were explored by univariate stepwise linear regression. In the analysis, the dependent variable was the selected clinical cluster, and all five marker clusters were initially included as predictor variables. Non-significant ($P>0.05$) predictors were removed by a backward/forward stepwise procedure, until only significant ($P<0.05$) marker factor(s) remained.
and 12 months were subjected to a single principal component analysis. Unlike the previous analysis, clinical assessments of change were omitted from this analysis. Under this circumstance the clinical variables broke down into two components. The first was composed of patient-related variables, patient self-assessment, pain on weight-bearing and stiffness; the second was composed of primarily physician-related variables, physician’s global assessment, soft tissue swelling and pain at rest (Table IV). The TGF-β1 association remained with the patient-related principal component factor 1 (PC-1) (Table IV) and the association over the year period is given in Fig. 2. When each individual clinical variable within the patient-related factor is examined individually, the correlation was strongest with the change in patient global assessment from baseline. The parameters of best fit of patient global assessment compared with baseline TGF-β1 are $Y_{max}=0$, $Y_{min}=-2$, $C_{50}=37.5$, $H=3.77$; $N=33$, $R=0.590$, $P<0.001$. By principal component analysis we previously found that TGF-β1 was independent of the other MMs. Using the two cases—change and no change—with the division point defined by logistic regression as 38 ng/ml of TGF-β1, Fisher’s exact test gave significance of $P=0.0003$.

### Signal joint vs inflammation markers

The correlations between the signal joint and the change in signal joint with the Inflammation Markers were examined next. A significant correlation was found only with the signal joint at baseline and the inflammation markers ($P<0.005$). The relationship between the inflammation markers and the change in signal joint did not reach statistical significance. We then explored the signal joint at baseline with each of the inflammation markers. CRP was the most highly correlated ($P<0.01$). Of the remaining markers, only TNF-RI reached statistical significance ($P=0.05$).

### Discussion

All clinical data obtained at baseline were examined by principal component analysis to derive independent factors with which to examine possible correlations with MMs. This process served two purposes: first, it strengthened the associations within the clinical data by clustering the measures into independent categories, and second, by reducing the number of clinical variables, it decreased the chances of obtaining spurious correlations. This dataset contained, as predictor variables, both baseline values and change values over the entire year. We maintained both baseline and change values together in this analysis, using the reasoning that the amount of change could be related to the initial values at baseline. In additional analysis (results not shown) the association between TGF-β1 and clinical change did not depend on the inclusion of baseline values as a predictor variable.

The clinical measures were found to cluster into three logical categories. The first category was composed of baseline determinations of patient’s global self-assessment, physicians’ global assessment, stiffness, pain at rest and pain on weight-bearing. This grouping might well have been suspected since a patient with a high score on one of these parameters most probably would have a high score for the others and vice versa. Similarly, the segregation of the changes over the year of the same clinical variables into a related group might also be expected a priori. Improvements or deteriorations in one of the measures over the year will frequently be mirrored in all the related measures. Finally, unlike the other clinical variables, the signal knee requires a determination of swelling by direct measurement. It is not a global assessment of general pain, stiffness, well-being, etc. Thus it is not unexpected that the measurements should segregate independently from the other two clinical clusters. Since no marker correlated with the change in swelling of the signal joint, this association did not enter into the analysis.

The three clusters of clinical parameters were then examined for associations with the five sets of MMs. With each clinical cluster relationships were found, but only to a single MM group. Two groupings, the synthetic and the degradation markers, failed to show any correlation with the clinical parameters. Since in OA the change in cartilage structure over a single year period is normally limited, this finding suggests that subjective clinical evaluations over a 1-year period do not reflect cartilage changes. These results emphasize the great importance of obtaining quantitative measures of the change in joint structure in order to determine the utility of these markers.

It was an unexpected finding that there was a correlation between the clinical condition at baseline and the level of HP. Since the status at baseline reflects cumulative damage over time, we expected there to be no direct correlation between any marker and baseline measurements. However, Thompson et al. have reported a correlation of radiological score and HP levels.
patients often shows abnormal bone scans unrelated to the degree of radiological abnormality. MacFarlane et al. and Dieppe et al. have suggested a correlation with progression of OA and bone scintigraphic activity, i.e. higher bone metabolism will be reflected later in more severe OA. Sharif et al. found that increased synovial fluid osteocalcin levels were correlated with increased abnormality of scintigraphic scans. Petersson et al. have found scintigraphic changes associated with joint pain in early OA. This suggests that, in some patients, bone metabolic changes may be related to clinical changes in OA.

The explanation for the association of HP and OA severity at baseline may be related to a combination of demographics and changes with menopause. Menopausal status, unfortunately, was not available in the retrospective data of this study. A plot of HP vs age in our patient group shows that after age 50, the population splits into those with
References


