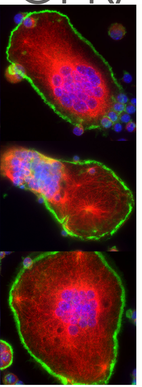
Osteoclast-related Biomarkers in Rheumatoid Arthritis

A. J. de Brum-Fernandes[1], P. Dufort[1], G. Boire[1], J. Dixon[2], M. Durand[1], R. Harrison[3], S. Komarova[4], L. Kurgan[5], M. Manolson[3], R. Rak[5], S. Sims[2]

[1]Division of rheumatology, Department of Medecine, Faculty of Medecine and Health Sciences, Sherbrooke University; [2]Department of Physiology and Pharmacology, University of Ontario; [3]Dental Research Institute, Faculty of Dentistry, University of Toronto; [4]Faculty of Dentistry, McGill University; [5]Department of Electrical and Computer Engineering, University of Alberta

OPRA



Introduction

Osteoclasts play a critical role in the pathophysiology of Rheumatoid Arthritis (RA) and possibly Osteoarthritis (OA). We hypothesized that the ability of peripheral blood mononuclear cells to form osteoclasts (osteoclastogenesis) and the activity of these osteoclasts correlate with pathology in RA and OA patients.

Objectives

To determine if osteoclast function, osteoclastogenesis and other related factors are related to the presence and activity of RA.

Methods

Subjects:

Patients and Controls: 139 patients with RA were recruited from the outpatient clinic at the Division of rheumatology, Centre Hospitalier Universitaire de Sherbrooke. Fourty-one self-reported healthy controls were recruited from the local population.

Osteoclastogenesis:

Peripheral blood mononuclear cells were separated from blood by Ficoll gradient and the number of CD14+ cells was determined by flow cytometry. The cells were cultured in the presence of RANKL and M-CSF and the number of OCs was evaluated after 21 days. Bone resorption was quantified on cortical bone slices stained with toluidine blue. OC apoptosis was evaluated by colorimetric assay (TACS™ TdT Blue Label kit).

Analysis of the results:

We developed descriptive (human-readable) classification models, one for each objective, capable of discriminating patients with regard to those objectives. The advantage of using descriptive methods lies in the simplicity of the model, which is human-interpretable, as opposed to other methods that create a mathematical or probabilistic model. Standard statistical methods were also used, Student's T-Test and ANOVA for continuous parametric variable(s), Mann-Whitney test and Kruskal-Wallis test for continuous non parametric variable(s) and the pearson's chi-square test for categoric variables.

Results

Descriptive analysis:

Mains demographics parameters and medical of controls and patients were analysed by standard analysis to determine differences among groups (see Table 1).

Main descriptives characteristics of the RA patients

Table 1. Baseline characteristics of the patients.†

Characteristics	Control (Ctl) (N=41)	RA-total (RAt) (N = 139)	RA-active (RAa) (N = 77)	RA-inactive (RAi) (N = 62)
Age -- yr	57.1±1.2	60.8±1.0*	62.1±1.3*	59.2±1.5
Female sex -- no. (%)	23 (56.1)	94 (67.6)	59 (76.6)*	35 (56.5)
Menopause -- no. (%)				
Menopause	NA	64 (68.1)	43 (72.9)	21 (61.8)
Pre-menopause	NA	4 (4.3)	2 (3.4)	2 (5.9)
Body-mass index‡	27.6±0.9	26.6±0.4	26.6±0.6	26.5±0.6
Ethnic group -- no. (%)¶				
Caucassian	38 (92.7)	133 (95.7)	72 (93.5)	61 (98.4)
Other	3	6	5	1
Smoking status -- no. (%)				
Ever smoke	15 (36.6)	84 (60.4)**	42 (54.5)**	42 (67.7)**
Alcohol status -- no. (%)	32 (78.0)	67 (48.6)***	35 (45.5)**	32 (51.6)**
Medication -- no. (%)				
NSAIDs	NA	57 (41.0)	32 (41.6)	25 (40.3)
Biphosphonate	NA	42 (30.2)	29 (37.7)	13 (21)*
Anti-TNF	NA	39 (28.1)	24 (31.2)	15 (24.2)
Sulfasalazin	NA	23 (16.5)	16 (20.8)	7 (11.3)
Methotrexate	NA	118 (84.9)	66 (85.7)	52 (83.9)
Antimalarial drugs	NA	101 (72.7)	50 (64.9)	51 (82.3)*
Prednisone	NA	24 (17.3)	17 (22.1)	7 (11.3)
Calcium (supplement)	NA	15 (10.8)	9 (11.7)	6 (9.7)
Vitamin D (supplement)	NA	11 (7.9)	8 (10.4)	3 (4.8)

† Plus-minus values are means ±SEM. CRP denotes C-reative protein, SR denotes sedimendation rate and COPD denotes coronary obstructive pulmonary disease.
‡ Body-mass index is the weight in kilograms divided by the square of the height in meters.
¶ Ethnic group are self-reported.
* P≤0,05 for the comparison between this group and the control group.
** P≤0.01 for the comparison between this group and the control group.
*** P≤0.001 for the comparison between this group and the control group.
* P≤0.05 for the comparison between this group and the RA active group.

Descriptive classificiation models:

The best results were obtained using the alternating decision tree algorithm for both RA presence and RA activity. Although this kind of a decision tree may seem to be more difficult to interpret than a standard decision tree, it is equipped with very useful pieces of information. Following Tables 2 and 3, each node (condition) is associated with a confidence (strength) with which the corresponding feature predicts (affects) the prediction outcome. That way it is easy to analyze which features play the most significant role in the process of classification and as a result would constitute the best candidates for potential markers. An example of a classification is given in Table 4.

Model of RA presence

Table 2. Model of RA presence

Condition	Strength
(0) Root	-0.602
(1a) Age < 66,5	0.157
(1b) Age >= 66,5	-1.070
(2a) Apoptosis < 19%	-0.399
(3a) Weight (kg) < 72,273	-0.835
(3b) Weight (kg) >= 72,273	0.408
(2b) Apoptosis >= 19%	0.517
(9a) Height (cm) < 151,2	-0.765
(9b) Height (cm) >= 151,2	0.515
(4a) Physical activity (min/wk) < 90	-0.548
(4b) Physical activity (min/wk) >= 90	0.385
(5a) Years non-smoker < 12,5	-0.162
(5b) Years non-smoker >= 12,5	-1.030
(6a) Alcohol consumption = no	-0.529
(6b) Alcohol consumption = yes	0.277
(7a) Resorption < 70869,5	0.376
(7b) Resorption >= 70869,5	-0.610
(8a) Physical activity (min/wk) < 330	-1.177
(8b) Physical activity (min/wk) >= 330	0.474

Table 2. Model of the presence of RA. Given a patient, if the sum of branches satisfying the conditions is negative then the patient has RA; otherwise (positive sum), the patient has no RA. The bigger the value (in either direction), the stronger the probability.

Table 3: Model of disease activity

Condition	Strength
(0) Root	-0.086
(1a) Join pain < 1,5	0.493
(1b) Join pain >= 1,5	-0.992
(2a) Synovitis < 3,5	0.245
(2b) Synovitis >= 3,5	-0.959
(3a) Osteoclasts < 1260,5	-0.238
(9a) Femoral BMD < 0,931	-0.345
(10a) Femoral T-score < -1,15	0.444
(10b) Femoral T-score >= -1,15	-0.756
(9b) Femoral BMD >= 0,931	0.592
(3b) Osteoclasts >= 1260,5	0.933
(4a) HAQ < 1.063	0.281
(4b) HAQ >= 1.063	-0.613
(5a) CD14+ < 2317	-0.239
(5a) CD14+ >= 2317	0.781
(6a) Severity < 36,5	0.303
(6b) Severity >= 36,5	-0.448
(7a) PTH < 3.65	0.504
(7b) PTH >= 3.65	-0.295
(8a) Hb < 136,5	-0.266
(8b) Hb >= 136,5	0.468

Table 3. Model of disease activity in patients with RA. Given a patient, if the sum of branches satisfying the conditions is negative then patient's disease is active; otherwise (positive sum), the disease is inactive. The bigger the value (in either direction), the stronger the probability.

Table 4. An example of RA activity classifications.

Features	Patient #	1	2
	Severity	43	0
	Synovitis	13	0
	Joint pain	13	0
	HAQ	1.25	0
	Hb	123	150
	PTH	4.9	7.4
	Femoral T-score	-1.3	0
	Femoral BMD	0.819	1.074
	CD14	1767	3511
	Osteoclasts	1757	974
Classification	Branches	0, 1b, 2b, 3b, 4b, 5a, 6b, 7b, 8a	1a, 2a, 3a, 9b, 4a, 5b, 6a, 6b, 8b
	Sum of branch strenghts	-2.965	2.5444
	Active RA ?	yes	no

Table 4. An example of RA activity classifications. According to the model of RA activity (Table 3) patient #1 has a negative accumulative value of branches that are matched by the given features for that patient, which as the result corresponds to the "active disease" classification. In contrast, patient #2 has a positive accumulative value of branches, which, in turn, corresponds to the "inactive disease" classification.

In order to ensure that the generated model is not biased to a given pair of training and testing sets we imposed a ten-fold cross validation. The classification results for RA presence and RA activity are shown in Tables 5 and 6, respectively.

Table 5: Classification results for RA presence

Test set	Accuracy	Specificity	Sensitivity/Recall	Precision
10-fold CV	76.1%	34.1%	88.5%	88.5%
Entire set	90.6%	75.6%	95.0%	93.0%

Table 6: Classification results for RA activity

Test set	Accuracy	Specificity	Sensitivity/Recall	Precision
10-fold CV	74.6%	66.7%	81.3%	74.4%
Entire set	93.5%	90.5%	96.0%	92.3%

10-fold CV indicates the ability of a model to classify unseen data; whereas entire set shows how well the model describes the data it was trained on. It's important to note that in the latter case it is possible to build a model that would classify data perfectly (all measures at the level of 100%); however that could potentially increase the complexity of the model to a degree that could be hard to understand or analyze.

RA presence study

Building a model describing the RA presence was based on 20 features including demographics and osteoclastogenic information. The model, in the form of a small alternating decision tree with 19 leaf nodes (see Table 2), allowed us to find several markers (feature-value pairs) that are related to the presence of RA.

The model reveals the following associations:

- Older age is strongly associated with RA presence. Low apoptosis is associated with RA presence, but only in low weight patients (below 72 kg), in which association becomes quite strong. High apoptosis is associated with absence of RA, but only for tall patients (above 151 cm).
- Physical activity below 90 min/wk is associated with RA presence while activity above 90 min/wk is somewhat associated with the absence of RA, which may be a consequence of having the disease.
- Being a smoker is always positively associated with RA presence, but much more so for ex-smokers who ceased smoking more than 12 years ago (this group may include older patients).
- Lack of alcohol consumption is relatively strongly associated with the presence of RA, which most likely is a consequence of using MTX (alcohol consumption is forbidden for patients using this drug).

RA activity study

Building a model describing the RA activity was based on 78 features including demographics, clinical observations, lab tests, and osteoclastogenic information. This extended set of features, compared to the RA presence study, is the result of more thorough clinical and lab analysis performed on patients with RA, when compared with controls who are included in the RA presence study. The model, an alternating decision tree with 21 leaves, reveals the following associations:

- The presence of the active disease is strongly associated with joint pain above 1.5 and Synovitis above 3.5.
- The number of osteoclasts (generated in low concentration of RANKL) below 1260 is somewhat associated with active arthritis, whereas the number of osteaclasts above 1260 is strongly associated with inactive arthritis.
- Patients with lower number of osteoclasts (below 1260) and Femoral BMD below 0.931 are more prone to active arthritis, but only when Femoral T-score for those patients is above -1.15.
- HAQ above 1.063 is associated with active arthritis, whereas HAQ below 1.063 is somewhat associated with inactive arthritis.
- The high number of CD14+ (above 2317) is strongly associated with inactive arthritis, whereas the lower number (below 2317) is somewhat associated with active arthritis.
- High PTH and low hemoglobin are somewhat associated with active arthritis, whereas low PTH and high Hb are indicative of inactive disease.

Note that legend of Figure 2 explains the above features.

Standard analysis

Statistical analysis of the numbers of osteoclasts precursors cells ($CD14^+$) and mature osteoclasts for RA patients and control (see Figure 1). No significant statistical difference were found between osteoclasts precusors cells number in RA patient and control. However, there is significantly more osteoclastogenesis for inactive RA patients when compare to both active RA patient and control.

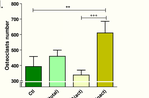


Figure 1. Statistical analysis of $CD14^+$ cells and osteoclastogenesis for control and RA patients. (** P less than 0.01 for a comparison with the group control group. +++ P less than 0.001 for the comparison between active and inactive RA active group.)

Osteoclasts physiology was analyzed using either apoptosis or resorption assay. RA patients (active and inactive) showed less apoptosis than control. Bone resorption was significantly higher in inactive RA patients compare to active RA patient and control group (see figure 2).

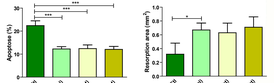


Figure 2. Statistical analysis of osteoclasts apoptosis and resorption. (* P less than 0.05 for a comparison with the control group. *** P less than 0.001 for a comparison with the group control group.)

Conclusions

Our analysis show that:

- Inactive but not active RA is associated with higher levels of *in vitro* osteoclastogenesis. This association was suggested by the descriptive classification models and confirmed by standard statistical analyses.
- Both methods showed that the group of RA patients present more in vitro bone resorption and less apoptosis than controls, but no difference was found between active and inactive RA patients.
- The descriptive classification models suggest that inactive RA is associated with higher numbers of $CD14^+$ cells but this association could not be confirmed by statistical analysis.

The results indicate that there are important differences in in vitro osteoclast biology in RA, as well as between active and inactive disease.

Acknowledgment

IRSC CIHR
Institute of Nutrition, Met. & Diabetes / Institut de la nutrition, du métabolisme et du diabète
Canadian Institutes of Health Research / Instituts de recherche en santé du Canada

Université de Sherbrooke
Faculté de médecine

Pfizer

Canadian Arthritis Network / Le réseau canadien de l'arthrite