Histopathology

Quantitative determination of the diagnostic accuracy of the synovitis score and its components

Journal:	*Histopathology*
Manuscript ID:	Draft
Manuscript Type:	Original Article
Keywords:	Synovitis , Receiver operating characteristic analysis, Rheumatoid arthritis, Histopathology

Quantitative determination of the diagnostic accuracy of the synovitis score and its components

[1]Slansky, E., [3]Li, J., [4]Häupl, T., [5]Morawietz, L., [6]Krenn, V., *[2]Pessler, F.

[1]Medical Faculty "Carl Gustav Carus" and
[2]University Children's Hospital, Technical University of Dresden, Germany
[3]Dept. of Biostatistics and Applied Probability, National University of Singapore
[4]Dept. of Rheumatology and [5]Inst. of Pathology,
Charité Universitätsmedizin, Berlin, Germany
[6]Institute for Pathology, Trier, Germany

*Corresponding author

University Children's Hospital
Div. of Rheumatology/Immunology
Technical University Dresden
Fetscherstr. 74
01307 Dresden, Germany
Email: frank.pessler@uniklinikum-dresden.de
Tel.: +49 (0)351 458-2440
Fax: +49 (0)351 458-5895

Key words: osteoarthritis, receiver operating characteristic analysis, rheumatoid arthritis, synovitis.

Running title: *Diagnostic accuracy of the synovitis score*

Summary

Aims. To assess the diagnostic accuracy of a 3-component synovitis score and to determine the relative contribution of each of its components to its overall discriminatory power.

Methods and results. The synovitis score was determined in 666 synovial specimens: normal synovium, n=33; post-traumatic arthropathy (PtA), n=29; osteoarthritis (OA), n=221; psoriatic arthritis (PsA), n=42; and rheumatoid arthritis (RA), n=341. The discriminatory abilities of the score and its components were quantified with standard and multi-category receiver operating characteristic (ROC) analysis. The score differentiated all arthropathies accurately from normal tissue (area under the ROC curve, AUC: 0.87-0.98), and RA from OA or PtA (AUC: 0.85 for both), but could not distinguish well within pairs of inflammatory or degenerative arthropathies. AUCs of the intimal hyperplasia and stromal cellularity components (r=0.94 and 0.91, respectively) correlated with the AUCs of the complete score markedly more strongly than the inflammatory infiltration component (r=0.60). Multi-category ROC analysis ranked the score several-fold higher than any of its components, and the components in the order stromal cellularity>intimal hyperplasia>infiltration.

Conclusion. Combining three distinct histological parameters into one 3-component score led to greatly increased overall diagnostic power. The discriminatory ability of the score stems more from measuring proliferative than infiltrative aspects of synovitis.

Introduction

We have previously described a 3-component score for the grading of the histological severity of synovitis (1-3). Each of the three components (intimal hyperplasia, stromal cellularity, and inflammatory infiltration) is graded on a scale from 0 to 3, yielding a final score from 0 to 9. A large-scale analysis of 559 synovial specimens was used to establish median values for normal (1.0), post-traumatic (2.0), osteoarthritis (OA, 2.0), psoriatic arthritis (PsA, 3.5), reactive arthritis (ReA, 5.0) and rheumatoid arthritis (RA, 5.0) specimens and to detect a high correlation between two independent observers in scoring these specimens (r=0.94) (1). In an unrelated study of synovial specimens characterized by a broad range of inflammation, the synovitis score correlated strongly with expression of the proliferation marker Ki-67 and the macrophage marker CD68 (a well validated tissue marker for disease activity in RA (4)) in the synovial subintima, thus supporting the notion that it reflects disease activity of inflammatory arthropathies at the tissue level (5). This score is based on hematoxylin-eosin stained tissue sections, can be applied to specimens covering a wide range of inflammation, and can be learned relatively easily (for online instruction see www.patho-trier.de). Indeed, when OA and RA synovia were scored by an investigator who had learned the score using this website, the resulting mean values were remarkably close to those previously reported by Krenn et al. (6). The synovitis score should therefore lend itself well to synovial tissue classification and to assessing histological disease activity in clinical trials where synovial biopsies are available. However, neither the diagnostic ability of the score to differentiate among various cohorts of specimens nor the relative contribution of each of its three components to its overall discriminatory power have been quantified systematically.

Receiver operating characteristic (ROC) analysis is often used to evaluate diagnostic tests. The ROC curve is based on sensitivities and specificities from a set of test results, where sensitivity (true positives) is plotted on the y-axis and 1-specificity (false positives) along the x-axis (7-10). The area under this ROC curve (AUC) is a direct measure of the ability of the test to distinguish between the two outcomes (diagnoses), with an AUC of 1 corresponding to a test with a perfect discriminatory ability, an AUC of 0.5 to a test without any discriminatory power, and an AUC of <0.5 corresponding to a

negative association with the outcome in question. The trade-off value is the optimal compromise between sensitivity and specificity and is usually defined as the point along the ROC curve possessing the maximal sum of sensitivity plus specificity. The accuracy of the test at this trade-off value can be expressed as the Youden index (i.e., sensitivity[fraction] + specificity[fraction] -1), where an index of 1 designates a perfectly accurate test (11).

ROC analysis with multiple classes and multiple tests ("multi-category ROC analysis") was developed recently to measure the diagnostic performance of a test in the discrimination among several classes (diagnoses) independently of the prevalence of the study populations (12). Here, the discriminatory ability of the test is measured as the hyper-volume under the ROC manifold (HUM), which, conceptually, corresponds to the AUC of binary ROC curve analysis. In conventional ROC analysis an AUC of 0.5 indicates that there is no difference between the two sample groups. Similarly, when multi-category ROC analysis is used to evaluate one test in the discrimination among K classes, the HUM corresponding to acceptance of the null hypothesis is defined as 1/K!, or less when ties are present (12). Thus, the diagnostic accuracy of a test for simultaneously differentiating among multiple categories increases the higher its HUM value is above this non-discriminatory value.

Here, we have assessed the diagnostic performance of the 3-component synovitis score and each of its components with conventional (binary) and multi-category ROC analysis and have determined the relative contribution of each of the three components to the overall discriminatory power of the score. We find that the score differentiates accurately between inflammatory and degenerative arthropathies and that the overall discriminatory ability of the complete score is greatly superior to that of any of its components.

Patients and Methods

The study included raw data of the previously published study of 559 specimens (1), plus 115 additional specimens which had become available since then. Reactive arthritis specimens (n=9) were excluded due to the small sample size, as were specimens from patients with concomitant crystal arthropathy (n=6; OA, 5; RA, 1) or necrotic tissue

(n=2, both RA). Thus, the following specimens were analyzed: "normal" synovium obtained from individuals without joint disease at the time of autopsy (n=33), post-traumatic arthropathy (PtA, n=29), OA (n=221), psoriatic arthritis (PsA, n=42) and RA (n=341), yielding a total of 666 specimens. Clinical diagnoses were defined according to standard criteria as described previously (1). All tissues were obtained at the time of surgical synovectomy or arthroplasty. Paraffin-embedded sections were stained with hematoxylin-eosin and evaluated for the three components of the synovitis score, i.e. lining hyperplasia, stromal cellularity, and inflammatory infiltrates. Each component is assigned a grade from 0 to 3, resulting in a minimum score of 0 and a maximum of 9. Most specimens (87%) had been scored by two examiners (L. M. and V. K.). In case of discrepancy between the observers the average of the two measurements was used. As opposed to the original score, which consists of integers, this sometimes resulted in scores carrying one decimal point, i.e. 1.5, 2.5, etc. Statistical significance of differences in the synovitis score between arthropathies was determined with the Mann-Whitney U test. Since results from both binary and multi-category ROC analysis are largely prevalence independent, we considered these analyses appropriate instruments to compare sample groups of divergent sizes, as in the present study.

Binary ROC analysis was carried out with the biostatistical software program SPSS (version 15.0). Paired comparisons were set up for ROC analysis such that the presumably more highly inflamed arthropathy was the state variable, according to the following order of presumed increase in inflammation: normal synovium<PtA <OA<PsA<RA. Multi-category ROC analysis was done as described (12) with a code written in the R software environment for biostatistical computing (http://www.r-project.org/).

Results

Differences in the synovitis score among arthropathies. The box plots in fig. 1 show the distribution of values of the complete score (panel A) or its components (panel B) in the sample cohorts. In agreement with our previous results (1), median values of the complete score were lowest in normal controls (median score, 0.5), whereas PtA and OA corresponded largely to a mild synovitis (median score, 2), PsA to a synovitis

intermediate between the degenerative arthropathies and RA, and RA to a moderate to severe synovitis (score, >4). More highly inflamed outliers were seen in both PtA and OA. Evaluation of the individual score components showed that, as expected, values for all three components were substantially higher in RA and PsA than in the normal controls or degenerative arthropathies (fig. 1B). It also revealed that the median synovitis score of the normal specimens was greater than zero mostly because of a mild lining hyperplasia.

Evaluation of the 3-component score with ROC curve analysis. ROC curves for representative comparisons are shown in fig. 2, and AUCs, sensitivities and specificities at optimal trade-off values, and Youden index values in table 1. The synovitis score differentiated with great precision between RA or PsA and the "non-inflammatory" (degenerative) arthropathies or normal tissues. The AUC for the comparison RA vs. normal synovium approached 1, with the optimal trade-off value of a score of 1.25 possessing a sensitivity and specificity of 0.92 and 0.97, respectively (Youden index, 0.89 for both), for a diagnosis of RA. The AUC of the clinically more relevant distinction RA vs. OA was somewhat lower (AUC, 0.85; sensitivity, 0.70; specificity 0.86; Youden index 0.56), and was the same as for the comparison RA vs. PtA. This result revealed a substantially higher diagnostic accuracy of the score than was suggested by the apparent overlap of the score values by the boxes and tick marks in fig. 1A. The synovitis score could not differentiate accurately between the inflammatory arthropathies RA and PsA. Likewise, the AUC for the differentiation of OA from PtA was near 0.5, supporting our previous report of a mild synovitis, similar to OA, in post-traumatic arthropathies (13). The score differentiated OA well from normal tissue with an AUC of 0.87. Surprisingly, the AUC for the comparison PtA vs. normal synovium was nearly 1. This could be explained by the observation that, although the ratio of the median scores of the PtA and normal cohorts (2.0/0.5, i.e. 4.0, $p<0.01$) was much smaller than that of the RA and normal cohorts (5.0/0.5, i.e. 10.0), the individual values of the PtA and normal cohorts did practically not overlap.

Evaluation of the components of the score. We then evaluated the relative contributions of the three components of the score to its diagnostic accuracy, namely by performing the ROC curve analysis for the same comparisons as above with each of the three components (results included in table 1). Linear regression analysis showed that the

AUCs of the intimal hyperplasia and stromal density components correlated better with the AUCs of the complete score than the infiltration component (table 2). As shown in fig. 3, component AUCs tended to be lower than AUCs of the complete score, but the curves of the intimal hyperplasia and stromal density components paralleled that of the complete score extensively. On the other hand, significant discrepancies were seen between the infiltration component and the complete score. This was most pronounced in the distinctions PtA vs. normal, OA vs. normal and PtA vs. OA, where the infiltration component was markedly inferior to the complete score, and PsA vs. PtA, in which the infiltration component actually had a higher AUC than the complete score. All 3 components possessed high AUCs for the distinctions between RA and the non-inflammatory arthropathies, and the stromal density component separated RA from OA essentially as well as the complete score.

Multi-category ROC analysis. Multi-category ROC analysis was then used to quantify and compare the abilities of the complete score and each of its components to simultaneously differentiate among all six sample groups. All HUMs were markedly higher than the HUM of the null hypothesis, indicating that even the weakest test did have a significant overall discriminatory ability. Most importantly, the score and its components could be ranked easily according to their HUM values (table 3). In agreement with the AUC values derived with binary ROC curve analysis (e.g., fig. 3), HUMs from multi-category ROC analysis affirmed the complete score as the best diagnostic test, with a HUM several fold larger than any of the component HUMs. Stromal density was the best test among the individual score components, followed by intimal hyperplasia, whereas the inflammatory infiltration component had the least overall discriminatory ability. In contrast to the relatively small differences between the complete score and the lining hyperplasia and stromal cellularity components, which the binary ROC curve analysis had suggested (e.g., fig. 3), multi-category ROC analysis thus revealed a substantial gain in overall diagnostic accuracy through use of the complete 3-component score. A hypothetical marker with no overall discriminatory ability would have a HUM of 0.000076 (table 3 legend). The HUM of the inflammatory infiltration component was several fold larger, thus demonstrating that even it, the weakest of the three components, has a considerable overall discriminatory power.

Discussion

The present study extends previous results demonstrating that the synovitis score can classify synovial specimens into those characterized by a high degree ("high-grade synovitis," e.g., RA and PsA) or low degree ("low-grade synovitis," e.g., OA and PtA) of inflammation (1). We have now quantified the discriminatory power of the score according to objective measures used in the evaluation of diagnostic tests and found that it is an accurate diagnostic tool not only in distinguishing between inflammatory and non-inflammatory (degenerative) arthropathies but also between degenerative arthropathies and normal synovium. Previous studies featuring this score have used the complete score, and we have now quantified the relative contributions of the individual components to its overall diagnostic ability. This analysis revealed that the stromal cellularity and intimal hyperplasia components correlated more strongly with the diagnostic accuracy of the complete score than the infiltration component. Moreover, multi-category ROC analysis ranked these two components above the inflammatory infiltration component. Together, these results suggest that the diagnostic power of this 3-component synovitis score stems to a large extent from measuring proliferative rather than inflammatory features of synovitis. This finding likely explains why the score delineates RA so accurately from non-inflammatory tissues, since the synovitis of RA is characterized by a high degree of cellular proliferation, as measured by objective endpoints such as expression of the proliferation-associated marker Ki-67 (14) (indeed, the synovitis score correlated positively with Ki-67 expression in that study). Along these lines, one may also speculate that the intimal hyperplasia component was inferior to the stromal cellularity component because the intimal overgrowth seen in the inflammatory arthropathies results only partly from cellular proliferation, but also from immigration of CD68+ cells. It appears remarkable that the score could distinguish accurately between normal tissue and PtA, an arthropathy generally not considered to be proliferative and characterized by a low degree of Ki-67 expression (13). One possible explanation is that part of the increased cellularity in the PtA is due to inhibition of apoptosis rather than increased proliferation and that, as in OA (13), migration of CD68+ cells into the lining is responsible for some of the intimal hyperplasia that was measured by this component of the score in the PtA samples.

Comparison with other synovitis scores. Several other grading systems for synovitis have been published, consisting of a single component based on the degree of inflammatory infiltration (e.g., (15)), or combinations of two (16), four (17), or even seven (18) components. However, none of these, or other, grading systems have been validated with ROC curve analysis or other instruments used in the evaluation of diagnostic tests, and it is unknown how their diagnostic accuracies compare to that of the 3-component synovitis score. Having identified the proliferation-related aspects of the synovitis score as its most powerful components, it will now be important to compare it to some of these other grading systems, for instance scores that do not contain proliferation-related components or scores that incorporate features that are more common among highly inflamed arthropathies such as fibrin deposition or vascular degeneration (18). It is also possible that combining the 3-component score with one of these latter features would increase its diagnostic ability in some instances. A further improvement of the score might become particularly meaningful in potentially difficult diagnostic scenarios, such as differentiating among inflammatory arthropathies, or between the inflamed end of the OA spectrum and RA.

Value of binary and multi-category ROC analysis in the present study. Binary ROC analysis identified several scenarios in which the score, despite the apparent overlap of values between sample cohorts that was suggested by the tick marks in fig. 1A, revealed a high diagnostic precision. Remarkably, the synovitis score differentiated between PtA and normal tissue with a near perfect AUC. The fold difference in median synovitis scores between these two specimen cohorts was relatively small and would not have suggested a particularly high potential of the score to differentiate between these cohorts. However, due to the very low degree of overlap in scores between the two specimen groups, the AUC was remarkably high, revealing a sensitivity of 0.97 and specificity of 0.91 (at the trade-off value) for this distinction. Thus, ROC analysis revealed a particular discriminatory strength of the synovitis score that would not have become apparent by comparing fold differences in median scores between the two sample groups. The present study represents the first application of multi-category ROC analysis since its initial validation evaluating microarray data to classify tumor tissues (12). We used it successfully to evaluate the relative contributions of the three components to the

score, and the results agreed well with those obtained independently by linear regression analysis. Moreover, it demonstrated that the overall diagnostic gain of using the complete 3-component score, instead of any of its components by itself, was substantially larger than gleaned from the results of the binary ROC analysis. Specifically, binary ROC analysis showed that each of the individual components had diagnostic strengths and weaknesses depending on the diagnostic comparison (e.g., fig. 2). In contrast, multi-category ROC analysis demonstrated clearly that combining these three separate histological parameters of synovitis into one multi-component score greatly increased the overall diagnostic power of the score among all five sample cohorts, thereby widening its applicability. Multi-category ROC analysis may thus be a powerful tool to evaluate multi-component histopathological grading systems in general, for instance in the field of tumor pathology.

Acknowledgements

We thank Prof. Angela Rösen-Wolff (Technical University Dresden) for helpful discussion and a critical reading of the manuscript. Support is gratefully acknowledged from International Collaboration Seed Grant SGP 08/A01 from the German Ministry for Education and Science.

Conflict of interest

None of the authors report a conflict of interest.

References

1. Krenn V, Morawietz L, Burmester GR, Kinne RW, Mueller-Ladner U, Muller B, et al. Synovitis score: discrimination between chronic low-grade and high-grade synovitis. Histopathology 2006;49:358-64.
2. Krenn V, Morawietz L, Burmester GR, Häupl T. [Synovialitis score: histopathological grading system for chronic rheumatic and non-rheumatic synovialitis]. Z Rheumatol 2005;64:334-42.
3. Krenn V, Morawietz L, Häupl T, Neidel J, Petersen I, König A. Grading of chronic synovitis--a histopathological grading system for molecular and diagnostic pathology. Pathol Res Pract 2002;198:317-25.
4. Haringman JJ, Gerlag DM, Zwinderman AH, Smeets TJ, Kraan MC, Baeten D, et al. Synovial tissue macrophages: a sensitive biomarker for response to treatment in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64:834-8.
5. Pessler F, Ogdie A, Diaz-Torne C, Dai L, Yu X, Einhorn E, et al. Subintimal Ki-67 as a synovial tissue biomarker for inflammatory arthropathies. Ann Rheum Dis 2007.
6. Diaz-Torne C, Schumacher HR, Yu X, Gomez-Vaquero C, Dai L, Chen LX, et al. Absence of histologic evidence of synovitis in patients with Gulf War veterans' illness with joint pain. Arthritis Rheum (Arthr Care Res) 2007;57:1316-1323.
7. Zweig MH. Evaluation of the clinical accuracy of laboratory tests. Arch Pathol Lab Med 1988;112:383-6.
8. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561-77.
9. Zweig MH, Broste SK, Reinhart RA. ROC curve analysis: an example showing the relationships among serum lipid and apolipoprotein concentrations in identifying patients with coronary artery disease. Clin Chem 1992;38:1425-8.
10. Faraggi D, Reiser B. Estimation of area under the ROC curve. Statistics in Medicine 2002;21:3093-3106.
11. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32-5.
12. Li J, Fine JP. ROC analysis with multiple classes and multiple tests: methodology and its application in microarray studies. Biostatistics 2008;9:566-76.
13. Pessler F, Dai L, Diaz-Torne C, Gomez-Vaquero C, Paessler ME, Zheng DH, et al. The synovitis of "non-inflammatory" orthopaedic arthropathies: a quantitative histological and immunohistochemical analysis. Ann Rheum Dis 2008;67:1184-7.
14. Pessler F, Ogdie A, Diaz-Torne C, Dai L, Yu X, Einhorn E, et al. Subintimal Ki-67 as a synovial tissue biomarker for inflammatory arthropathies. Ann Rheum Dis 2008;67:162-7.
15. Pearle AD, Scanzello CR, George S, Mandl LA, DiCarlo EF, Peterson M, et al. Elevated high-sensitivity C-reactive protein levels are associated with local inflammatory findings in patients with osteoarthritis. Osteoarthritis Cartilage 2007;15:516-23.
16. Haywood L, McWilliams DF, Pearson CI, Gill SE, Ganesan A, Wilson D, et al. Inflammation and angiogenesis in osteoarthritis. Arthritis Rheum 2003;48:2173-7.
17. Smeets TJ, Dolhain RJ, Breedveld FC, Tak PP. Analysis of the cellular infiltrates and expression of cytokines in synovial tissue from patients with rheumatoid arthritis and reactive arthritis. J Pathol 1998;186:75-81.

18. Flagg SD, Meador R, Hsia E, Kitumnuaypong T, Schumacher HR, Jr. Decreased pain and synovial inflammation after etanercept therapy in patients with reactive and undifferentiated arthritis: an open-label trial. Arthritis Rheum 2005;53:613-7.

Figure Legends

Fig. 1. Quantitative determination of the synovitis score in the specimen cohorts. Each of the three components (intimal hyperplasia, stromal cellularity and inflammatory infiltration) was graded from 0 to 3, yielding a final score between 0 and 9. Boxes represent the 25th to 75th percentiles; horizontal lines represent the median; upper and lower whiskers represent the maximal and minimal values, respectively. Specimens examined were: normal synovium, n=33; post-traumatic arthropathy (PtA), n=29; osteoarthritis (OA), n=221; psoriatic arthritis (PsA), n=42; and rheumatoid arthritis (RA), n=341. **A.** Results for the complete synovitis score. *P* values for statistical significance of differences between medians for all possible paired comparisons are listed in table 1. **B.** Results for the 3 individual components of the score.

Fig. 2. ROC curves for selected paired comparisons. ROC curves were generated with SPSS 15.0 biostatistical software using synovitis scores from the following specimens: **A.** RA (n=341) vs. normal synovium (n=33); AUC, 0.97; **B.** RA (n=341) vs. OA (n=221); AUC, 0.85; and **C.** RA (n=341) vs. PsA (n=42); AUC, 0.63.

Fig. 3. Comparing AUCs of the synovitis score and its 3 components. Paired comparisons (identified below the x-axis) were arranged by decreasing AUCs of the synovitis score (blue line) and compared to AUCs derived with the stromal cellularity (yellow), intimal hyperplasia (orange) and inflammatory infiltration (green) components. The graph was generated with the results of binary ROC analysis listed in table 1.

Table 1. Evaluation of the synovitis score and its components with binary ROC analysis.

Comp.[1]	Ratio of medians[2] (*p* value)	Complete Score					Components		
		AUC (95% CI)	Trade-off value	Sens.	Spec.	Youden index	AUC Stroma	AUC Lining	AUC Infiltr.
PtA vs. Normal	4.0[#]	0.98 (0.95 – 1.01)	1.25	0.97	0.91	0.87	0.96	0.87	0.58
OA vs. Normal	4.0[#]	0.87 (0.82 – 0.92)	1.25	0.69	0.91	0.60	0.76	0.77	0.72
PsA vs. Normal	6.5[#]	0.90 (0.83 – 0.97)	1.75	0.76	0.97	0.73	0.87	0.80	0.81
RA vs. Normal	10.0[#]	0.97 (0.95 – 0.98)	1.75	0.91	0.97	0.88	0.96	0.90	0.89
PtA vs. OA	1.0[§]	0.59 (0.50 – 0.67)	1.25	0.97	0.31	0.27	0.70	0.60	0.37
PsA vs. OA	1.6[#]	0.68 (0.57 – 0.78)	4.25	0.40	0.95	0.35	0.69	0.59	0.66
RA vs. OA	2.5[#]	0.85 (0.82 – 0.88)	3.25	0.74	0.83	0.57	0.84	0.75	0.76
PsA vs. PtA	1.6[§]	0.63 (0.50 – 0.76)	3.25	0.50	0.90	0.40	0.56	0.54	0.74
RA vs. PtA	2.5[#]	0.85 (0.80 – 0.90)	3.25	0.74	0.90	0.64	0.77	0.72	0.83
RA vs. PsA	1.5[#]	0.63 (0.53 – 0.73)	2.75	0.86	0.43	0.29	0.66	0.67	0.55

[1]Pairs of sample cohorts were set up for ROC curve analysis such that the presumably more inflamed cohort was the state variable (listed first) and the presumably less inflamed one, the test variable (listed second).
[2]Median synovitis score of state variable/median synovitis score of test variable. [§] $p<0.05$, [#] $p<0.001$ (Mann Whitney U test).
Abbreviations: AUC, area under the ROC curve; CI, confidence interval; Comp., comparison; Infiltr., infiltration; PtA, post-traumatic arthropathy; OA, osteoarthritis; PsA, psoriatic arthritis; RA, rheumatoid arthritis; ROC, receiver operating characteristic, Sens., sensitivity; Spec., specificity.

Table 2. Correlations between the synovitis score and its components.

	Synovitis score	Intimal hyperplasia	Stromal density	Infiltration
Synovitis Score	1	---	---	---
Intimal hyperplasia	0.94 (0.017)	1	---	---
Stromal density	0.91 (0.016)	0.95 (0.0018)	1	---
Infiltration	0.60 (0.008)	0.43 (0.004)	0.34 (0.0025)	1

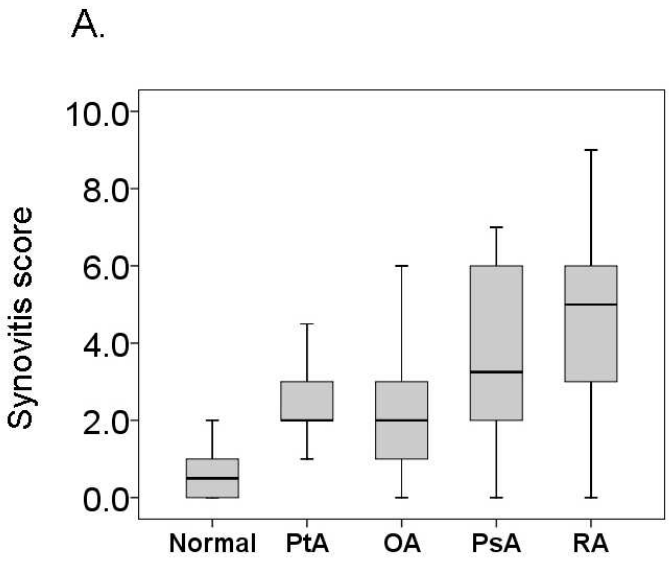
Values represent correlation coefficients (r) obtained by linear regression (*p* values).

Table 3. Diagnostic ranks of the synovitis score and its components according to multi-category ROC analysis

Rank	Test	HUM[1]
1	Synovitis score	0.062
2	Stromal cellularity	0.014
3	Intimal hyperplasia	0.0075
4	Infiltration	0.0005
5	Nondiscriminatory marker[2]	0.000076

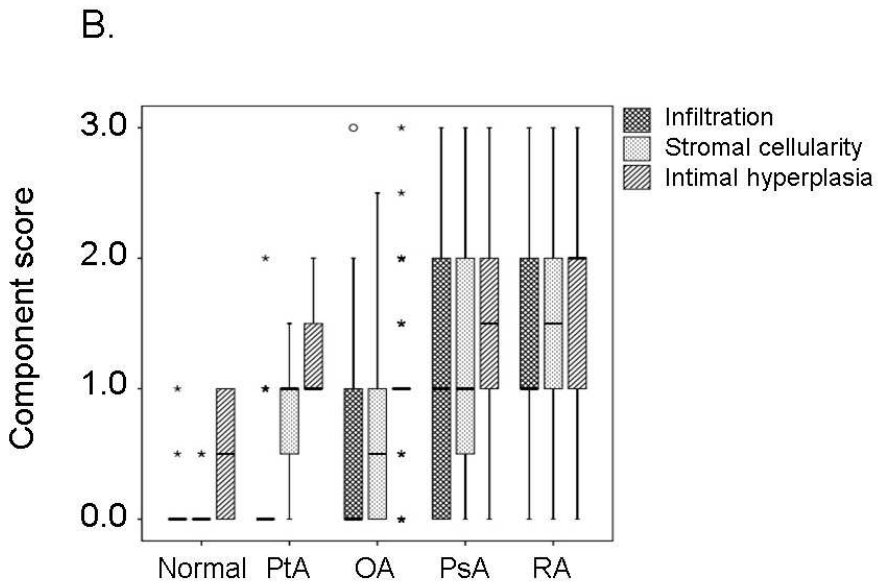
[1]Values represent hyper volumes under the ROC manifolds (HUMs), calculated according to ref. (12).

[2]Hypothetical marker corresponding to the null hypothesis, considering the number of classes (n=5) and the presence of ties in this analysis.



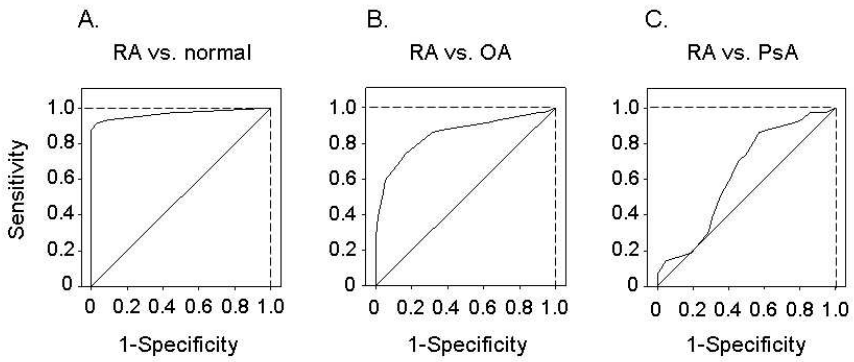
Slansky et al., Fig. 1A

Quantitative determination of the synovitis score in the specimen cohorts. Each of the three components (intimal hyperplasia, stromal cellularity and inflammatory infiltration) was graded from 0 to 3, yielding a final score between 0 and 9. Boxes represent the 25th to 75th percentiles; horizontal lines represent the median; upper and lower whiskers represent the maximal and minimal values, respectively. Specimens examined were: normal synovium, n=33; post-traumatic arthropathy (PtA), n=29; osteoarthritis (OA), n=221; psoriatic arthritis (PsA), n=42; and rheumatoid arthritis (RA), n=341. A. Results for the complete synovitis score. P values for statistical significance of differences between medians for all possible paired comparisons are listed in table 1.
254x190mm (96 x 96 DPI)



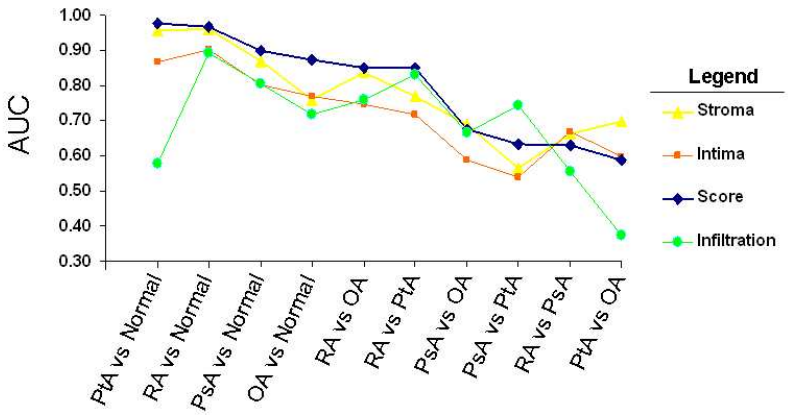
Slansky et al., Fig. 1B

Results for the 3 individual components of the score.
254x190mm (96 x 96 DPI)



Slansky et al. Fig. 2

ROC curves for selected paired comparisons. ROC curves were generated with SPSS 15.0 biostatistical software using synovitis scores from the following specimens: A. RA (n=341) vs. normal synovium (n=33); AUC, 0.97; B. RA (n=341) vs. OA (n=221); AUC, 0.85; and C. RA (n=341) vs. PsA (n=42); AUC, 0.63.
254x190mm (96 x 96 DPI)



Slansky et al. Fig. 3

Comparing AUCs of the synovitis score and its 3 components. Paired comparisons (identified below the x-axis) were arranged by decreasing AUCs of the synovitis score (blue line) and compared to AUCs derived with the stromal cellularity (yellow), intimal hyperplasia (orange) and inflammatory infiltration (green) components. The graph was generated with the results of binary ROC analysis listed in table 1.
254x190mm (96 x 96 DPI)