

Appendix IV

Statistical analysis

Introduction

Smeekes et al. [1] concluded in their study of 2017 that “Scoring morphologic features of MoM tissue is not reproducible using the ALVAL score [2] or the Oxford ALVAL score [3]”. The histological slides “were independently examined by three pathologists who were experienced in diagnosing skeletal and soft tissue related diseases, and thus well trained in recognizing different types of inflammation cells and patterns of inflammation”. The variables of the two classifications scored by each pathologist were provided in a supplemental appendix.

The semiquantitative scoring of the descriptive terms used in the two classifications (synovial lining, inflammatory infiltrate, tissue organization, necrosis, perivascular lymphocytic infiltrate) can be discordant even among pathologists with long-standing experience in the field, although a two point difference observed in some cases of this study is unusual; on the contrary, the recognition of specific types of inflammatory cells (macrophages, lymphocytes, plasma cells, eosinophils/neutrophils) should show minimal or no discordance and it is independent from the two classifications of ALVAL and also the area of musculoskeletal pathology. Therefore, all parameters were re-examined for each pathologist to confirm or disprove the hypothesis that the differences in the scoring of the variables of the two classifications is due to the experience/accuracy of the histological analysis of the three pathologists more than the reproducibility of the classifications.

Methods

The analysis was performed on the scoring sheet obtained by the three pathologists separately. All histological parameters were compared among the three pathologists who performed the analysis. Descriptive statistics were presented as means and standard deviation (SD). The normality assumption was assumed satisfied due to the central limit theorem. The means of each histological variable was compared among the three pathologists using the analysis of variance test (ANOVA) with significance at $\alpha = 0.05$. Tukey’s test of multiple comparisons was used for pairwise comparisons of histological data in order to determine the value of the difference among the pathologists. All statistical analysis was performed using SAS 9.4 software.

Results

The means and the corresponding standard deviations for each of the scored histological parameters of the two classifications are shown in [Table 1](#).

Table 1 Descriptive statistics

Variable	Pathologist 1	Pathologist 2	Pathologist 3
	Mean \pm SD	Mean \pm SD	Mean \pm SD
Synovial lining	2.94 \pm 0.23	2.68 \pm 0.53	2.73 \pm 0.56
Inflammatory infiltrate	2.03 \pm 0.94	2.57 \pm 1.07	2.68 \pm 1.06
Tissue organization	2.25 \pm 0.65	2.08 \pm 0.8	2.46 \pm 0.61
Campbell ALVAL (sum score)	7.22 \pm 1.35	7.32 \pm 1.8	7.86 \pm 1.81
Macrophages	2.44 \pm 0.7	2.54 \pm 0.77	2.16 \pm 0.76
Lymphocytes	1.86 \pm 0.64	2.38 \pm 0.83	2.05 \pm 0.78
Plasma cells	0.53 \pm 0.55	0.78 \pm 0.75	1.14 \pm 0.54

Eosinophils/Neutrophils	0.22 ±0.42	0.41±0.60	0.03±0.16
Necrosis	1.81±0.71	1.46±1.07	1.41±0.96
Oxford ALVAL (semiquantitative score)	1.67±0.75	1.90±1.13	2.08±0.84

The results of the ANOVA test of multiple comparisons were used to compare all the scores of the three pathologists in order to detect statistically significant differences and are presented in [Table 2](#). At alpha 0.05, the distributions of the synovial lining (p=0.0379), inflammatory infiltrate (p=0.0179), lymphocytes (p=0.0147), plasma cells (p=0.0003), eosinophils/ neutrophils (p=0.0013) were significantly different. These results confirm that the values obtained by these pathologists for the scoring of the two classifications varied widely.

Table 2 ANOVA test for multiple comparisons

Variable	Pathologist 1 vs Pathologist 2 vs Pathologist 3	
	f-value	p-value
Synovial lining	3.3746	0.0379*
Inflammatory infiltrate	4.18	0.0179*
Tissue organization	2.81	0.0649
Campbell ALVAL (sum score)	1.53	0.2215
Macrophages	2.54	0.0836
Lymphocytes	4.39	0.0147*
Plasma cells	8.75	0.0003*
Eosinophils/Neutrophils	7.05	0.0013*
Necrosis	2.0	0.1407
Oxford ALVAL (semi quantitative score)	1.84	0.1642

*Significant at alpha =0.05

Graphical displays of the distributions and pairwise comparisons of the statistically significant differences of the scoring by the three pathologists are presented in [Fig. 1 through Fig. 5](#).

Figure 1 Distribution of synovial lining

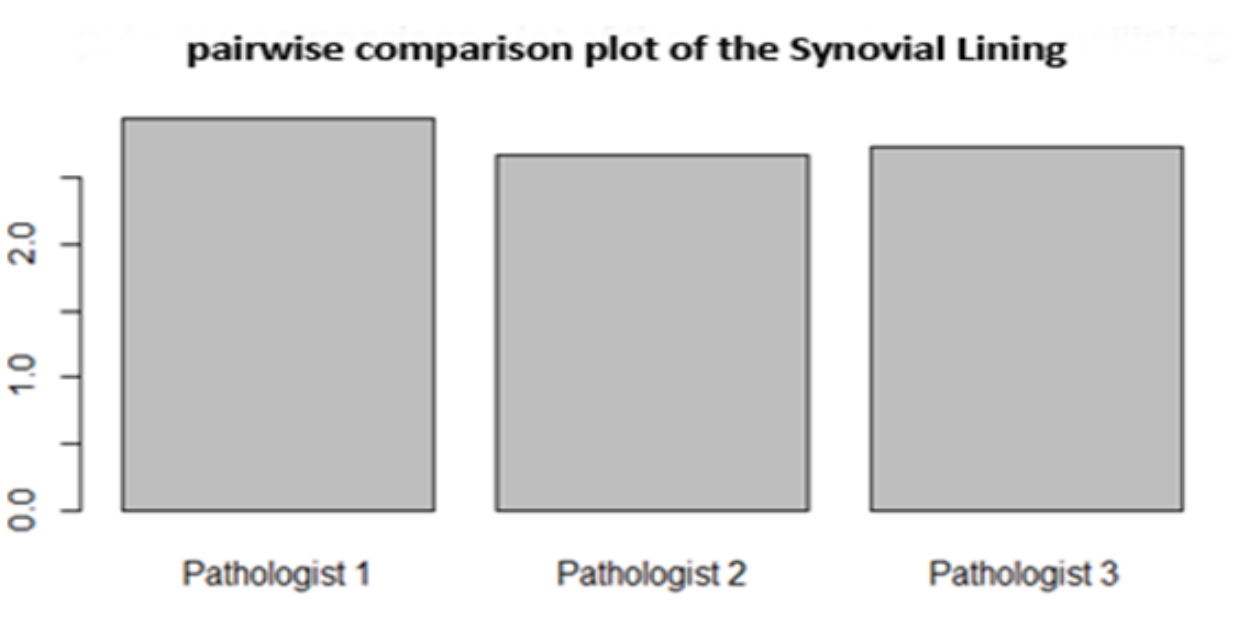
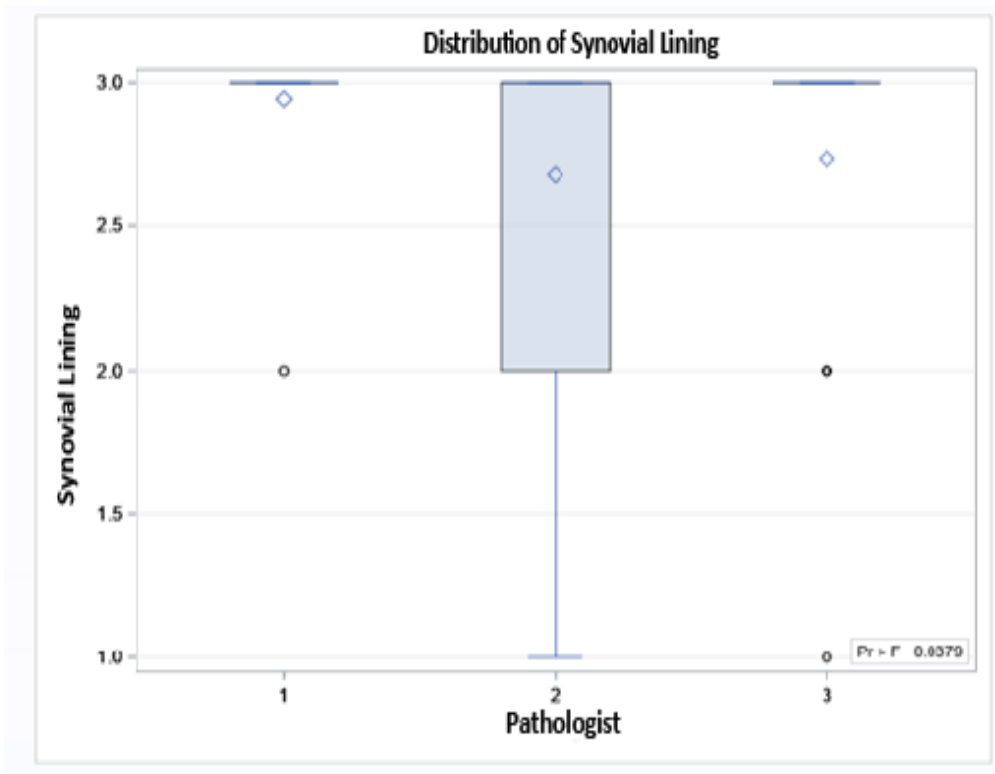


Figure 2 Distribution of the inflammatory infiltrate

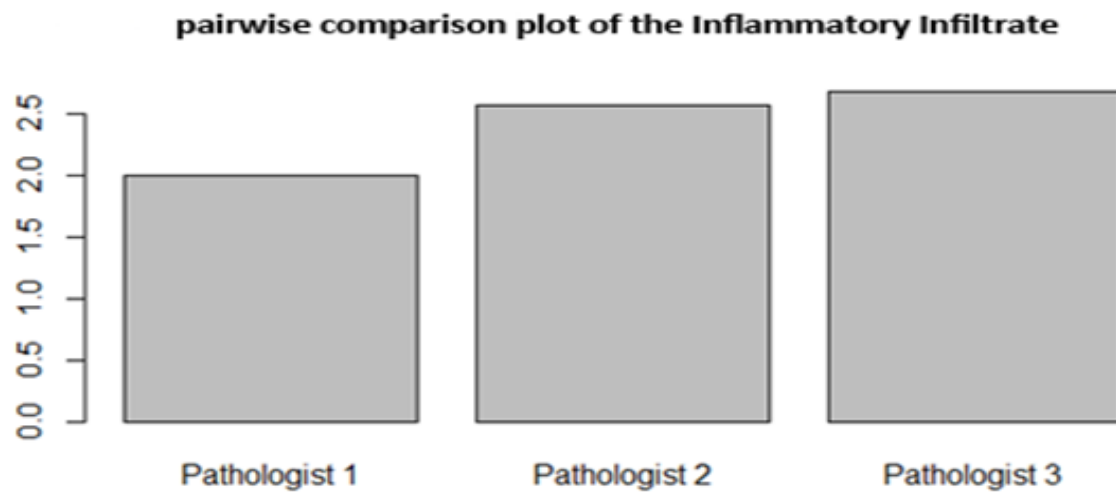
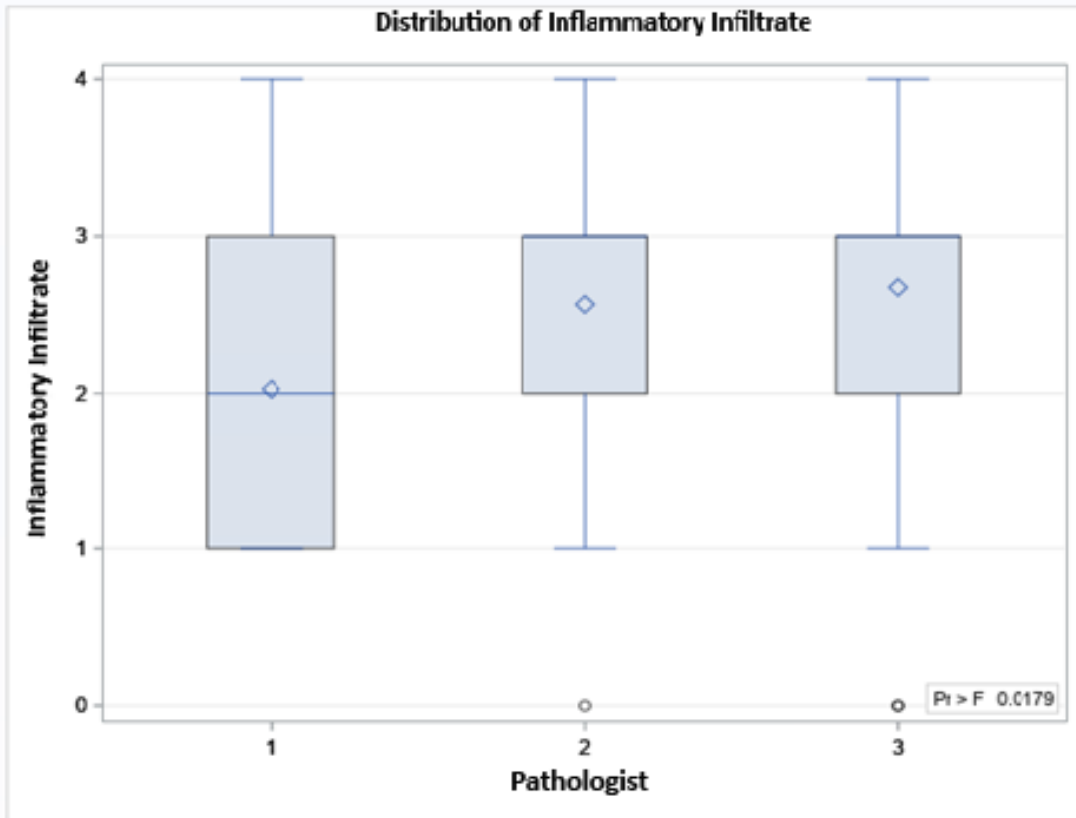
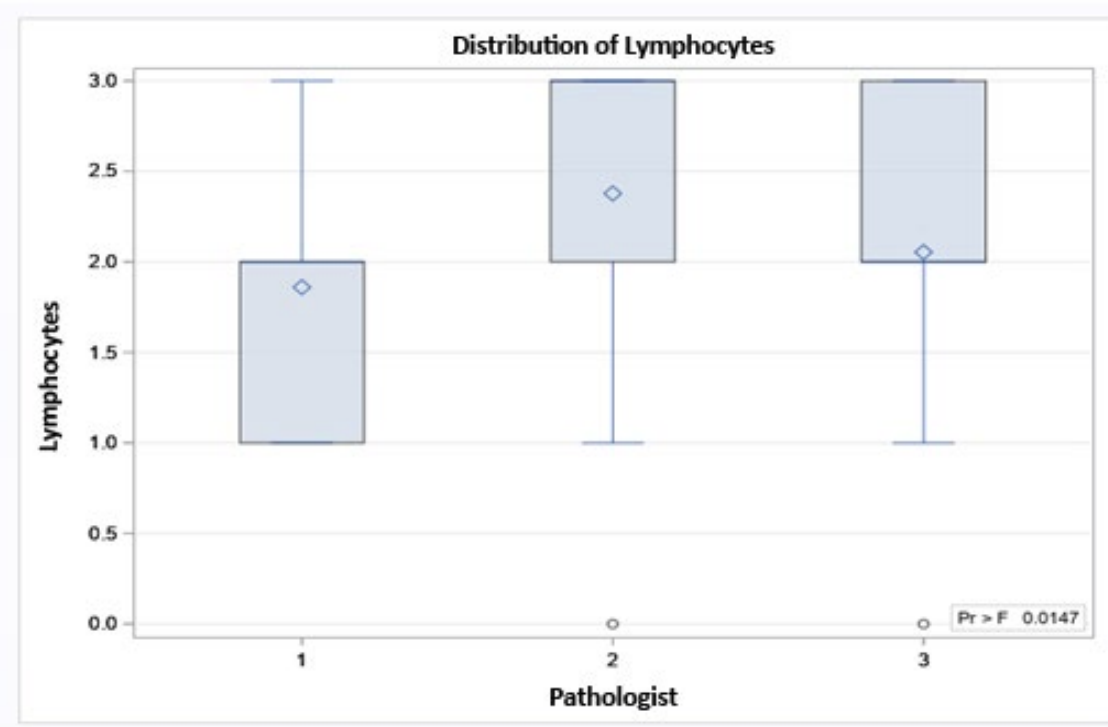


Figure 3 Distribution of lymphocytes



pairwise comparison plot of the Lymphocytes

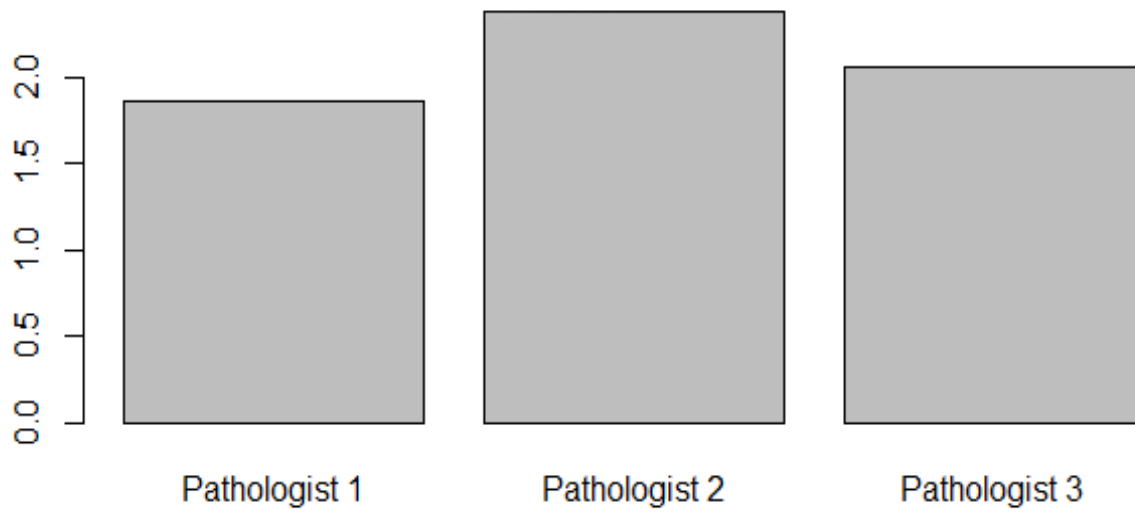
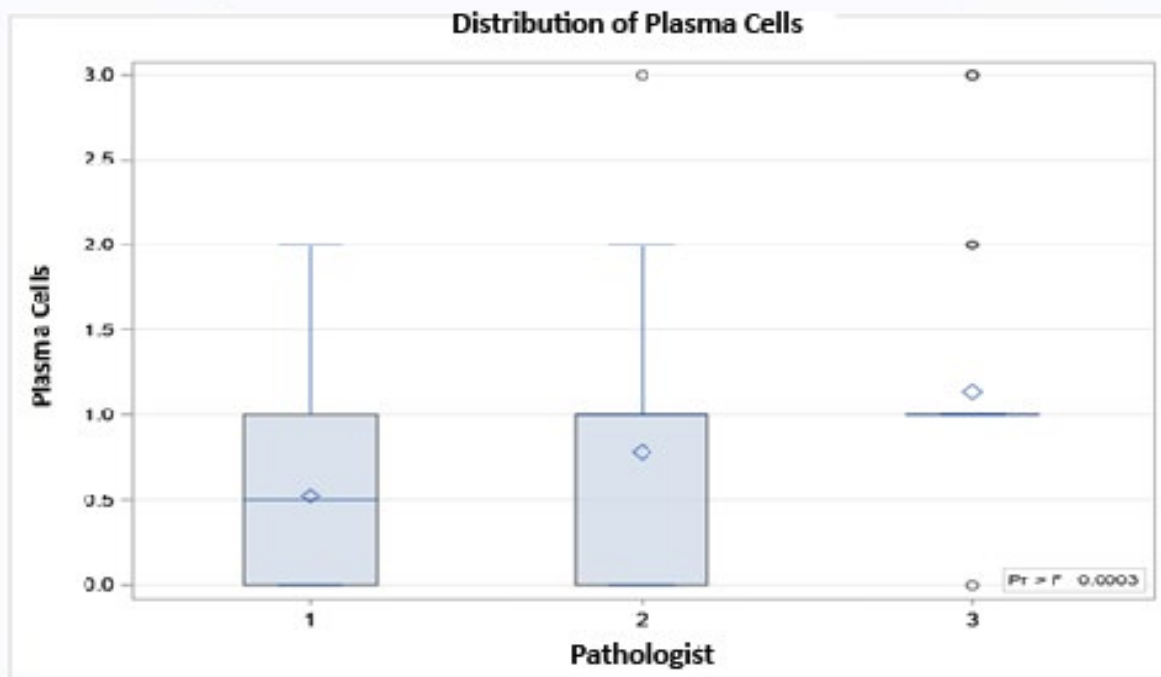


Figure 4 Distribution of plasma cells



pairwise comparison plot of the Plasma Cells

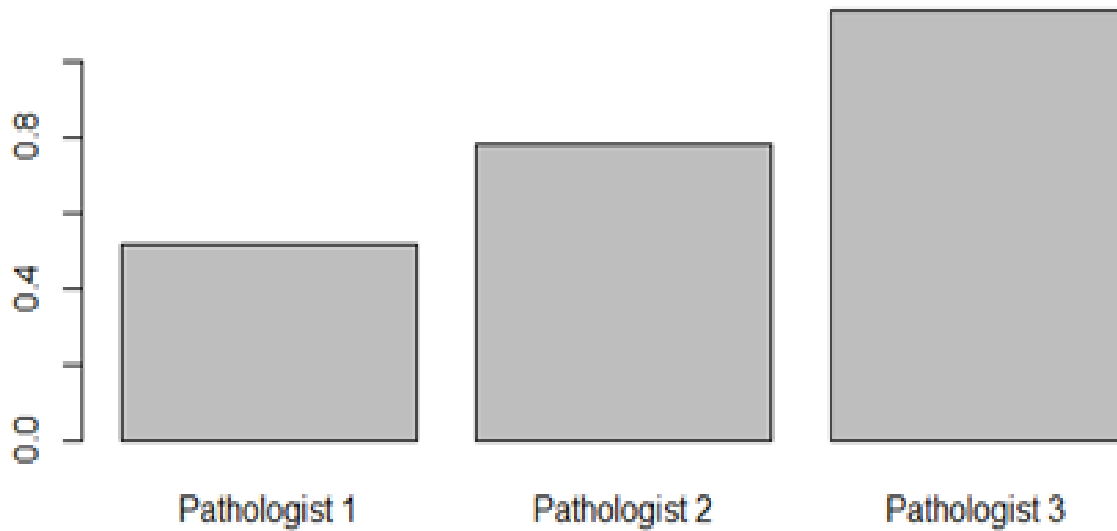
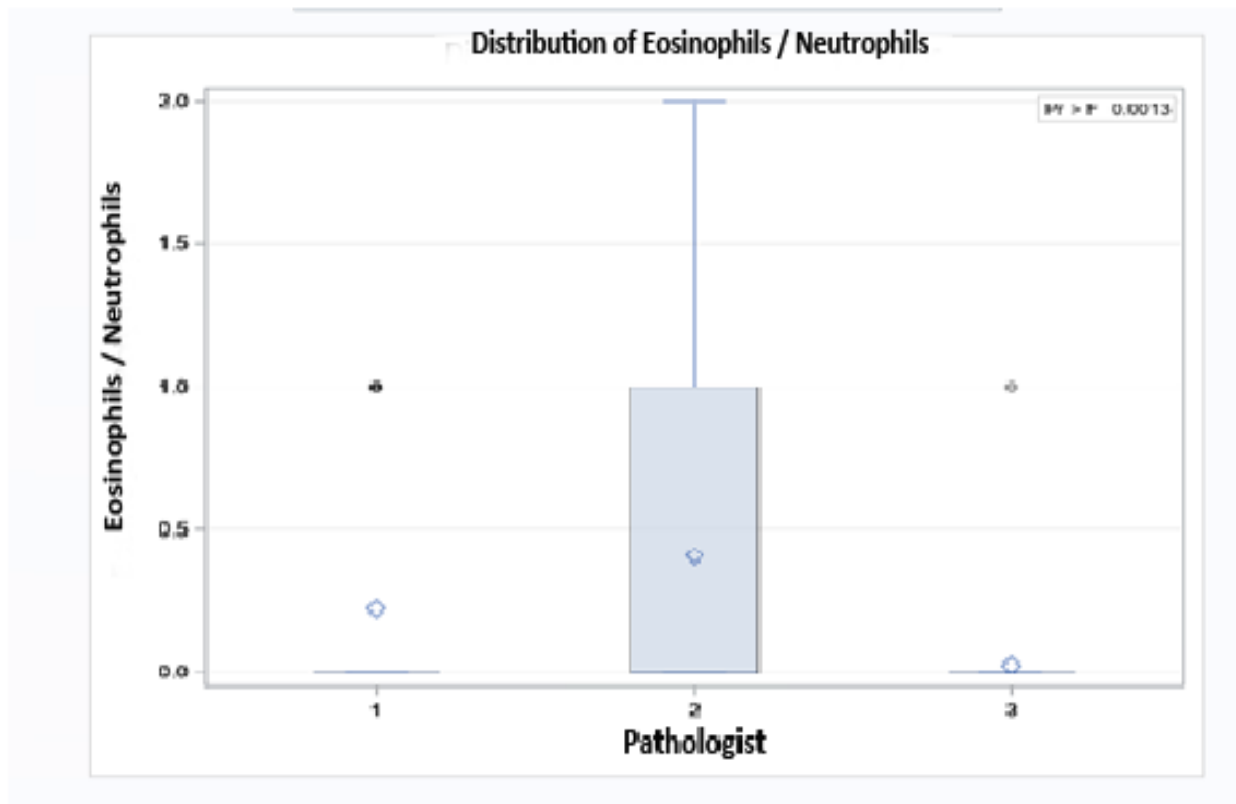
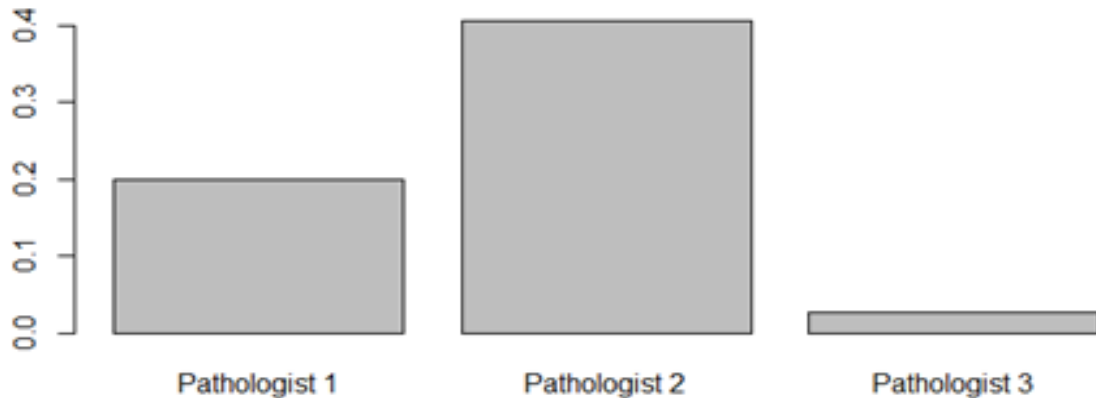


Figure 5 Distribution of eosinophils/neutrophils



pairwise comparison plot of the means of the Eosinophils / Neutrophils



Discussion

The analysis of the individual scoring sheet shows statistically significant differences among the three pathologists in the semiquantitative scoring of lymphocytes, plasma cells and eosinophils/neutrophils and not for macrophages, which are easier to identify because usually abundant and with discreet cytoplasm filled with conventional and corrosion metallic

particles. This difference is particularly evident for the last two categories, which can be present focally and difficult to identify consistently with a single examination of the histological slides. Moreover, the distribution for plasma cells is different for pathologist 3 compared to pathologists 1 and 2 and for eosinophils/neutrophils for pathologist 2 compared to pathologists 1 and 3 indicative of different individual difficulty in the recognition of a specific type of inflammatory cell. Several factors can have contributed to these differences among the three pathologists: (a) a single examination of the slides instead of two examinations at different times to decrease intra-observer variability; (b) different time spent in examining the slides; (c) a high number of cases examined in one session; (d) time spent at high magnification (x400) to identify the inflammatory cells. All these factors usually occur when the experience/accuracy of the pathologists is not comparable.

Conclusions

The identification of specific cell types in an inflammatory infiltrate is dependent on the expertise of the examining pathologist and independent from the reproducibility of the two histological classifications. Thus, statistically significant differences among the three pathologists in the distribution of lymphocytes, plasma cells, and eosinophils/neutrophils are attributed to non-homogeneous expertise among the three pathologists and non-related to differences in the interpretation of the parameters used for the two histological classifications examined. As a consequence, the conclusions of the study by Smeekes et al. [1] based on the comparable experience of the three pathologists must be rejected.

References

1. **Smeekes C, Cleven AHG, van der Wal BCH, Dubois SV, Rouse RW, Ongkiehong BF, et al.** Current pathologic scoring systems for metal-on-metal THA revisions are not reproducible. *Clin Orthop Relat Res* 2017;475:3005-11.
2. **Campbell P, Ebramzadeh E, Nelson S, Takamura K, De Smet K, Amstutz HC.** Histological features of pseudotumor-like tissues from metal-on-metal hips. *Clin Orthop Relat Res* 2010;468:2321-27.
3. **Grammatopoulos G, Pandit H, Kamali A, Maggiani F, Glyn-Jones S, Gill HS, et al.** The correlation of wear with histological features after failed hip resurfacing arthroplasty. *J Bone Joint Surg Am* 2013;95:e81 (1-10).