Attribute Correlations in BIOSCAD data

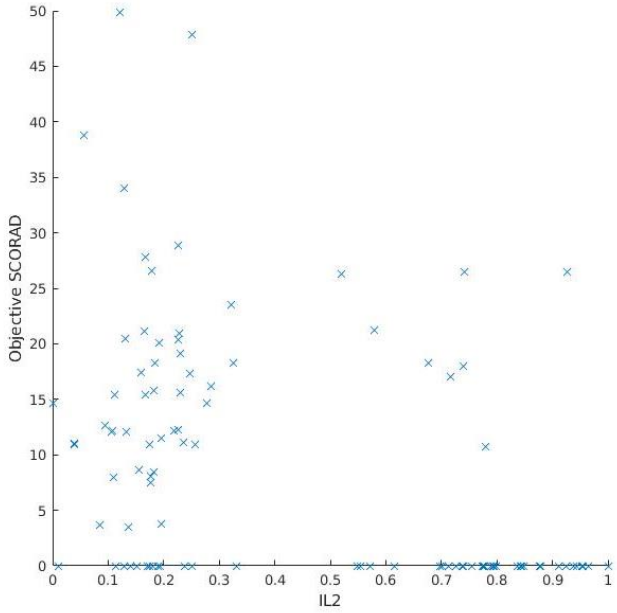
Method:

Correlation was calculated between each pair of attributes. All attributes, with the exception of the three SCORAD attributes, have been logged and normalised during pre-processing. This data has been visualised by colour coding the strength of the correlation – red represents negative correlation, white represents no correlation, and green represents positive correlation. The colours fade in to each other to indicate the strength of the correlation. The darker the colour, the more strongly correlated it is. It should be noted that, to make the colour gradients more distinct, the darkest colours do not necessarily represent strong correlation but rather represent the strongest correlation in the data set. As objective, subjective, and total SCORAD will be highly correlated, they are removed from the colour coding to make the colour gradient more distinct.

Non-lesional data:

In order to create a larger dataset, the non-lesional data from AD sufferers was combined with the non-lesional data from the control group. The SCORAD values for the control group were set to 0. This gave a total data set of 100 points.

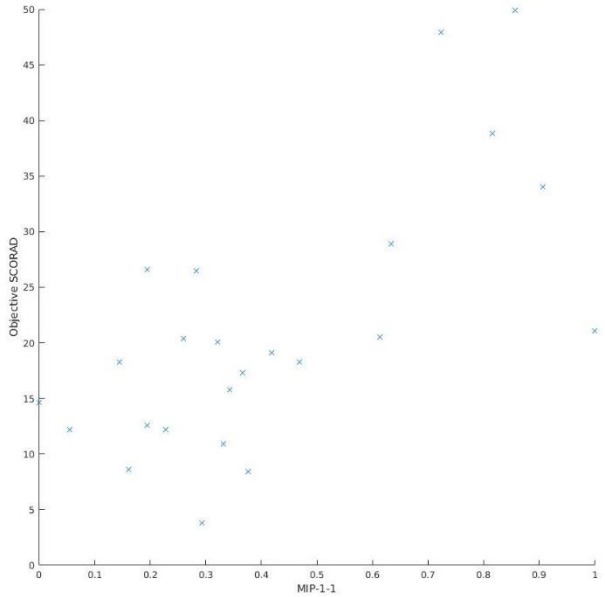
The combined lesional data showed very little correlation between any single biomarker and Objective SCORAD. The biomarker which showed greatest correlation to objective SCORAD was IL_2 which had a product moment correlation coefficient of -0.4019. However, as shown in the graph below, this does not indicate clear correlation between the two.



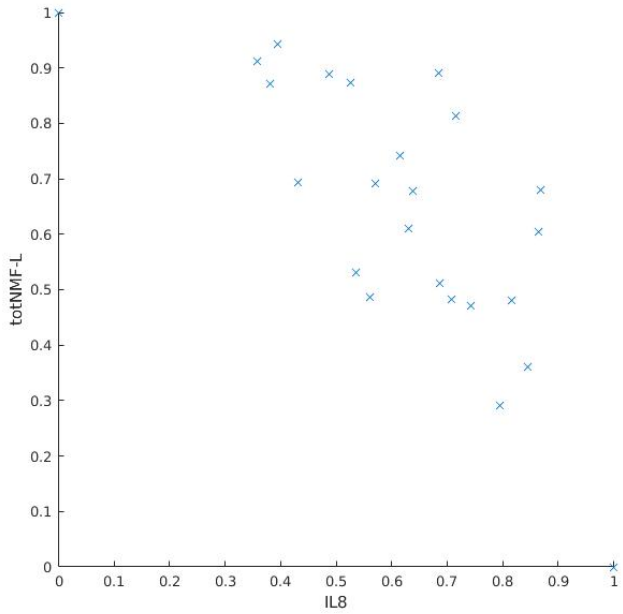
Lesional data:

After pre-processing, the lesional data contained only 17 data points and thus it is difficult to draw any conclusions.

In the lesional data, the strongest correlation to objective SCORAD was with MIP_1_1. Between these two, the product moment correlation coefficient was 0.68237.



There was also evidence of negative correlation between IL_8 and totNMF_NL.



Full data:

Full correlation data between attributes can be found in the excel spreadsheet. The data set has been divided up into non-lesional data for AD patients only, non-lesional combined data, and lesional data for AD patients. For each of these, both the product moment correlation coefficients and the Spearman rank correlation coefficients have been calculated between each pair of attributes.