

Results/Conclusion

Carlo van Buiten

10/6/2021

Research question

Is cell uniformity (shape and size) a good predictor for the malignancy of breast cancer?

Results

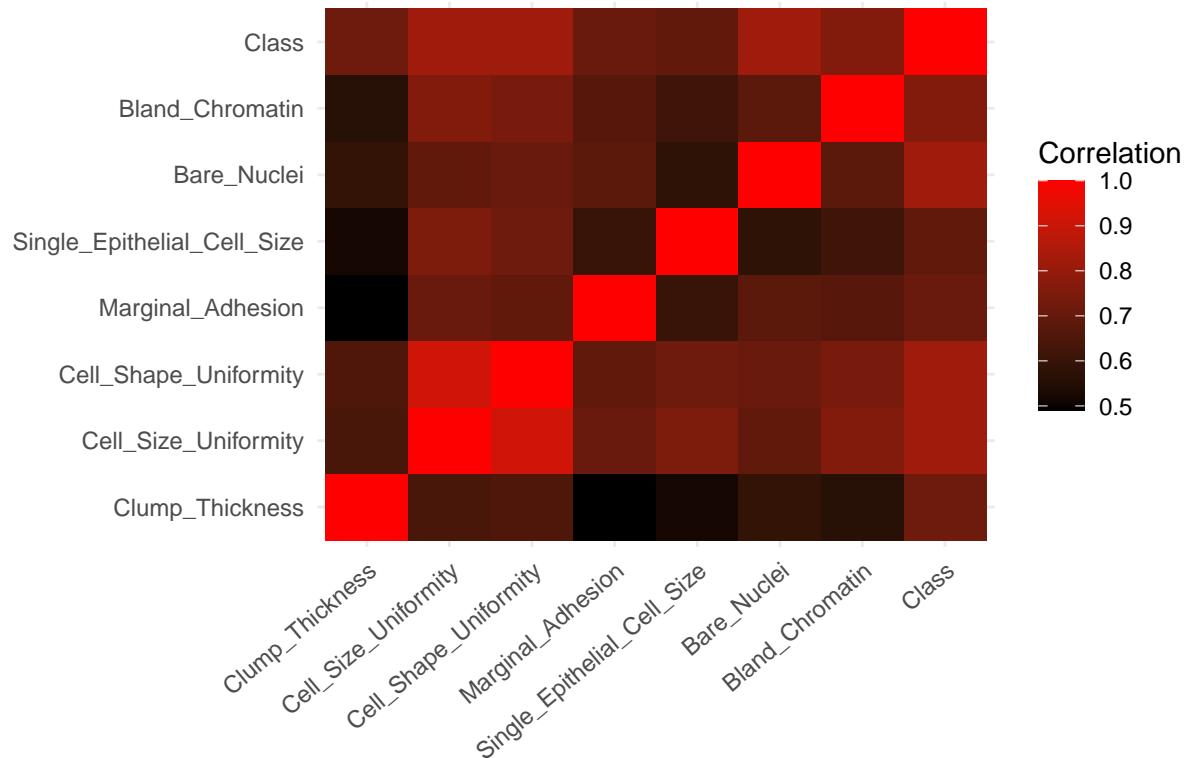


Figure 1: Heatmap showing the correlations between all attributes.

As can be seen in figure 1, there are 4 attributes that correlate with malignancy (class) for more than 75%, 2 of which are pertaining to cell uniformity. These attributes are cell size uniformity, cell shape uniformity, bare nuclei and bland chromatin. The spreads for these attributes can be seen in figures 2 through 5.

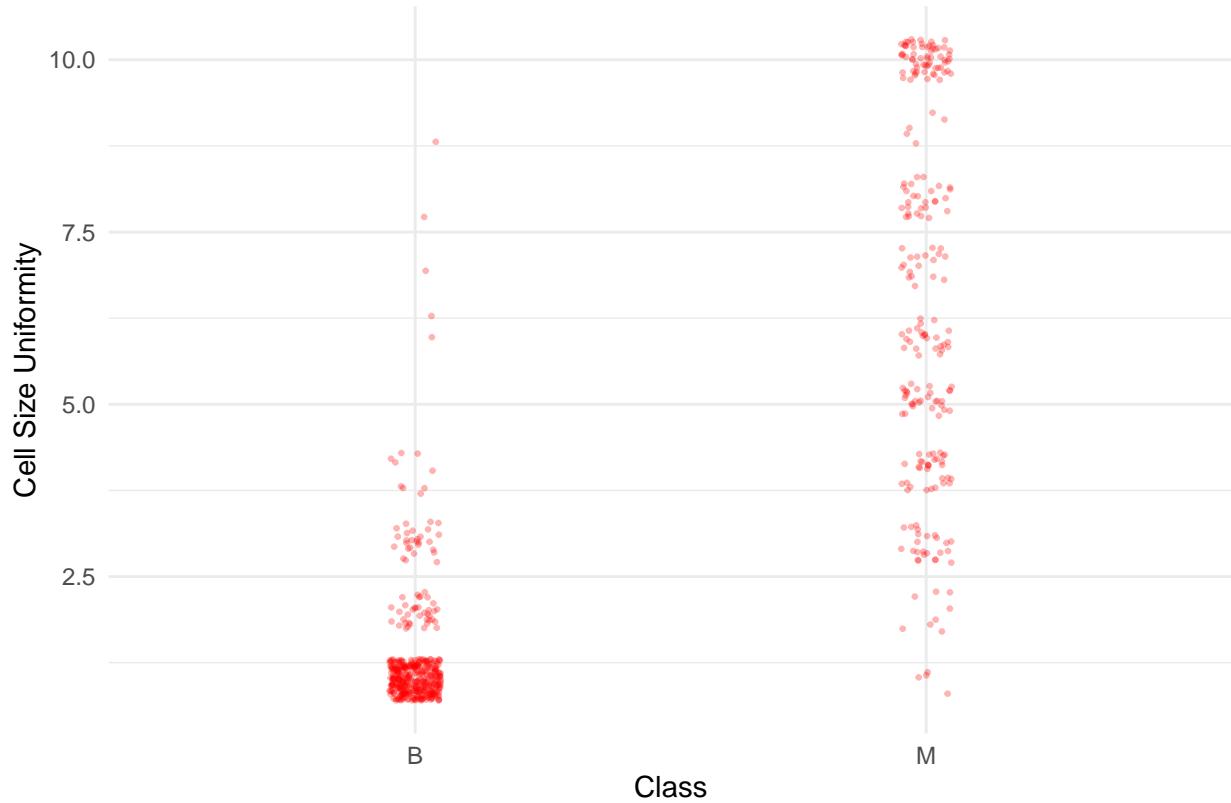


Figure 2: Cell size uniformity spread per class.

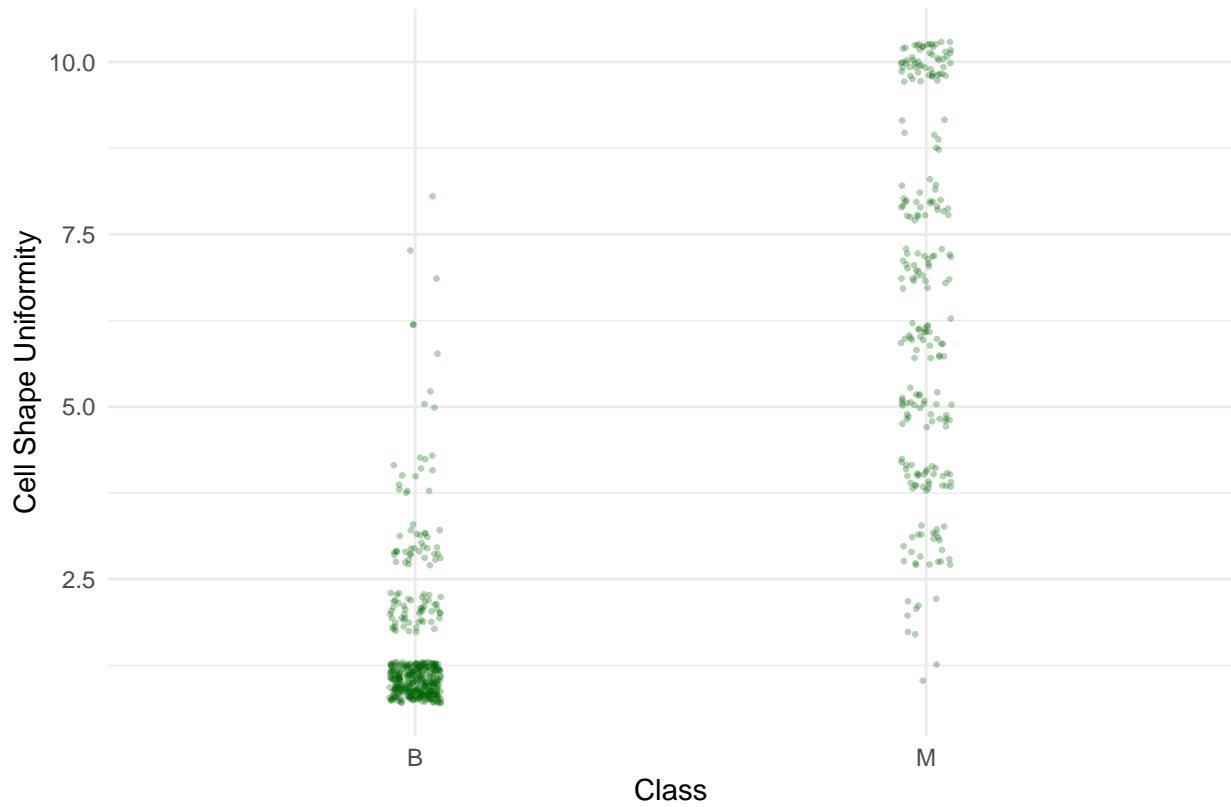


Figure 3: Cell shape uniformity spread per class.

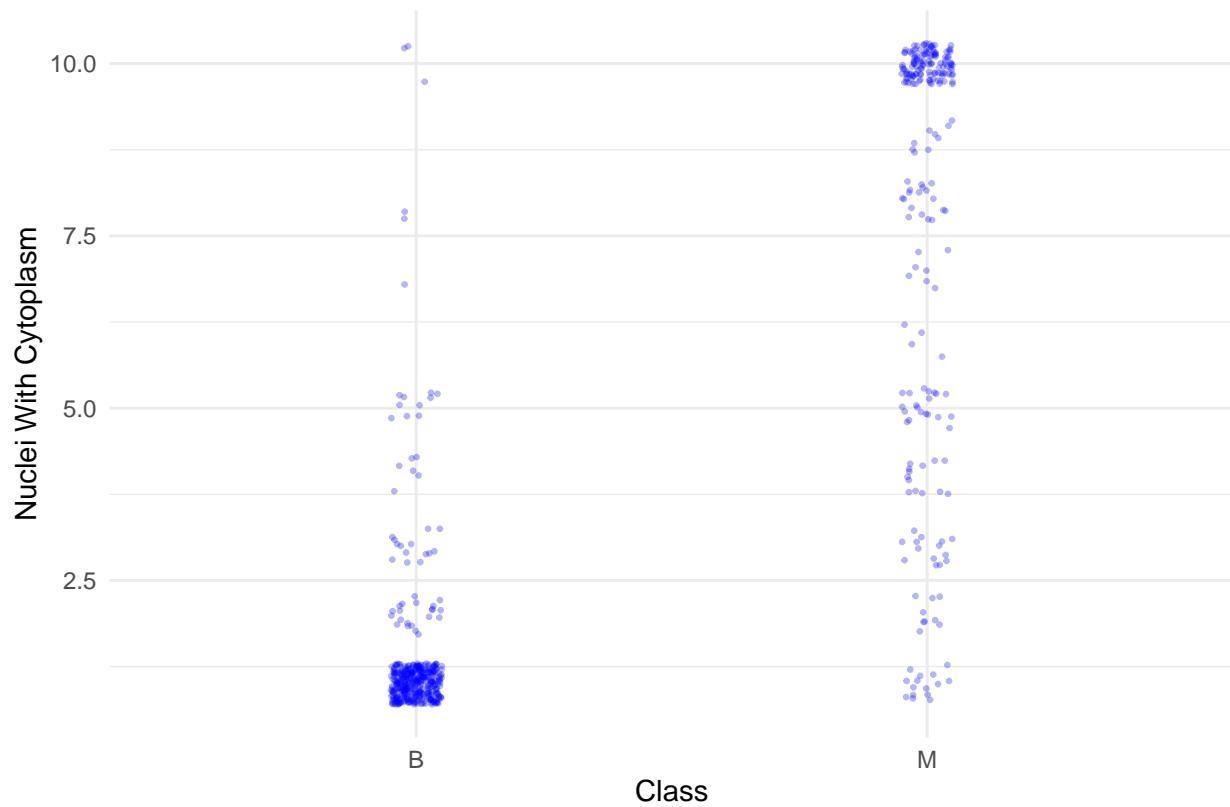


Figure 4: Bare nuclei spread per class.

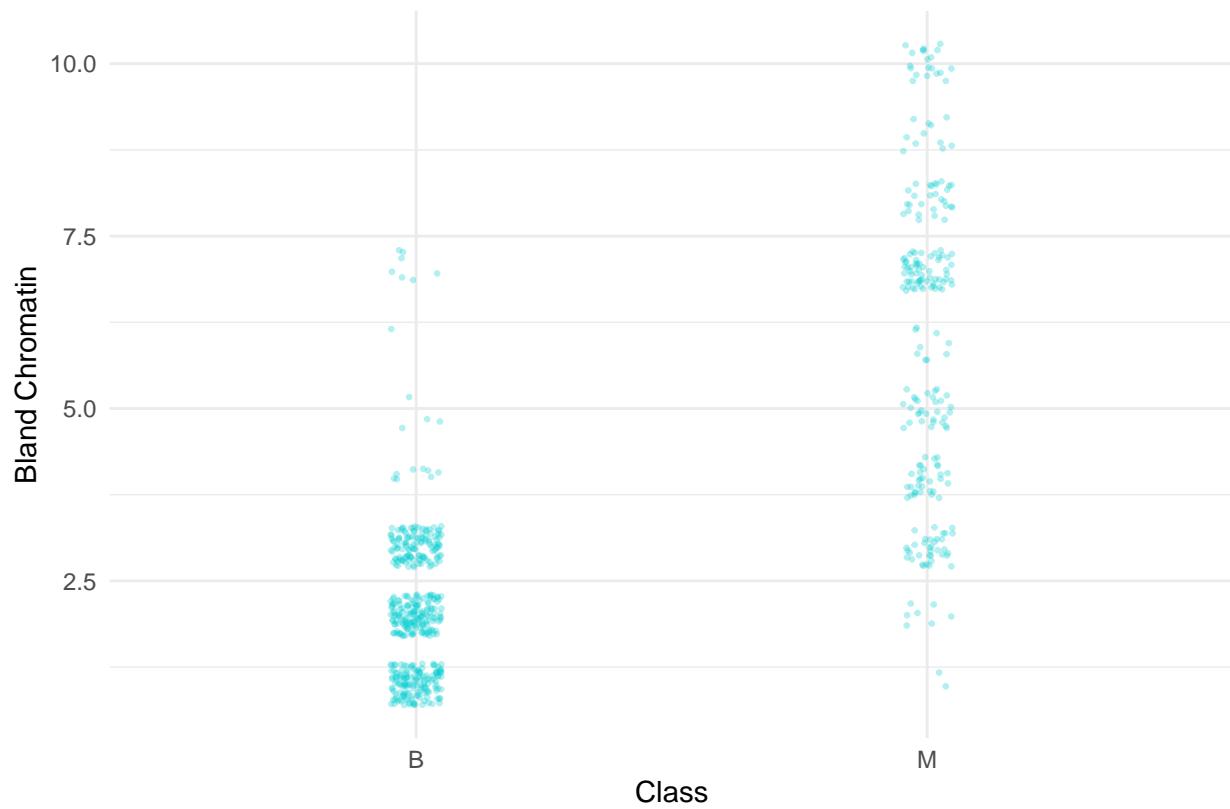


Figure 5: Bland Chromatin spread per class.

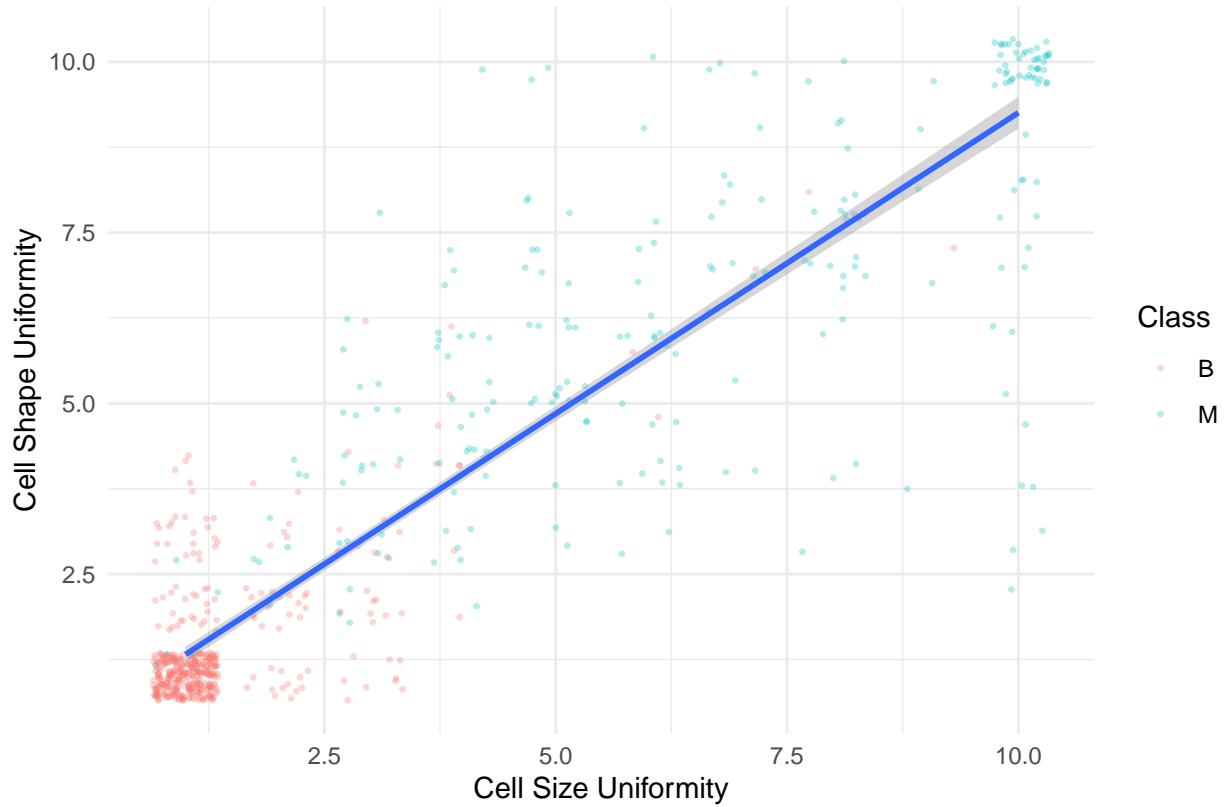


Figure 6: Plot of the correlation between cell shape and size discrepancies

Cell size uniformity and cell shape uniformity (see figure 6) seem to correlate quite a lot, although this correlation seems to wane a little at the higher end. This correlation entails that the larger the differences in size between cells, the larger the differences in shapes as well.

Discussion

Most features that have to do with erratic cell growth seem to correlate strongly with the malignancy of the cases of breast cancer, this makes sense as one of the defining aspects of cancer is the abnormal growth and multiplication of cells.

Taking all the results and graphs at face value, one might say cell uniformity is a strong predictor of malignant breast cancer. While this might be the case, some assumptions have to be made in order for us to be able to draw this conclusion. This is because in the user manual the grades (ranging from 1 to 10) are defined as “Completely uniform”, followed by an amount of percentages (90% to 10%) and eventually ending at “Cells are inconsistent with their uniformity”. This means that at the very least some rounding has been done, which is automatically a loss of information. It is also unclear how one would measure uniformity for cells in a tissue sample in terms of percentages, or specifically; how these were measured in this study.

The unspecific nature of the data is a glaring issue when it comes to drawing hard conclusions, but assuming the integrity of the data is in order, we may potentially verify the hypothesis.

Conclusion

Barring the accuracy of the data (or rather the potential lack thereof), we can safely conclude that cell uniformity is a good predictor for the malignancy of breast cancer.