Love Plot: an illustrative method to show the imbalance between groups.

Judith Peñafiel[1], Natàlia Pallarés[1], Cristian Tebé[1]

[1]jpenafiel@idibell.cat, npallares@idibell.cat, ctebe@idibell.cat, Biostatistics Unit, IDIBELL

BACKGROUND: A common step in observational studies and randomised controlled trials is to assess the balance of relevant covariates between groups. P-values are often used as a measure of between-group imbalance. However, the limitations associated with the (mis)use of p-values are well known. A simple alternative is to compute the standardized mean difference (SMD). For continuous covariates, SMDs provide the same scale and calculation methods are well known. For categorical variables, however, there is less agreement on how to calculate SMDs.

OBJECTIVE: To evaluate the different methods used to calculate balance statistics for a categorical variable.

METHODS: Data were simulated with different assumptions of between-group imbalance in a set of categorical variables and different sample sizes. Common balance statistics for categorical variables were estimated: raw difference (RD) across levels, pooled standardized difference (PSD) across levels and Mahalanobis distance (MD). In addition to the most common ones, the maximum of the RD (MRD) was calculated. The results were presented and compared graphically with the balance statistics on the x-axis and the variable categories on the y-axis, also known as the love plot.

RESULTS: First results suggest that the balance statistics tend to increase as the sample size decreases. However, in most cases these statistics lead to the same conclusions about between-group balance. Statistics across levels, such as RD and PSD, make it possible to observe at which levels the differences are present, but are difficult to visualize graphically. On the other hand, an overall statistic like MD and MRD is easier to visualize graphically. Nonetheless, if there are imbalanced covariates, MD alone may not be sufficient to determine the levels at which differences occur. Similarly, MRD only shows the most unbalanced level. In addition, if a particular level is unbalanced, MD may be overshadowed by the other levels.

CONCLUSIONS: The balance statistics approaches are heuristic approximations, and their exact values are not crucial specially if they are close to 0. A graphical approach, such as love plot, could be an excellent alternative to p-values in assessing imbalance. The choice may depend on the researcher's aims, preference, and the way in which the degree of imbalance between the groups is to be shown.