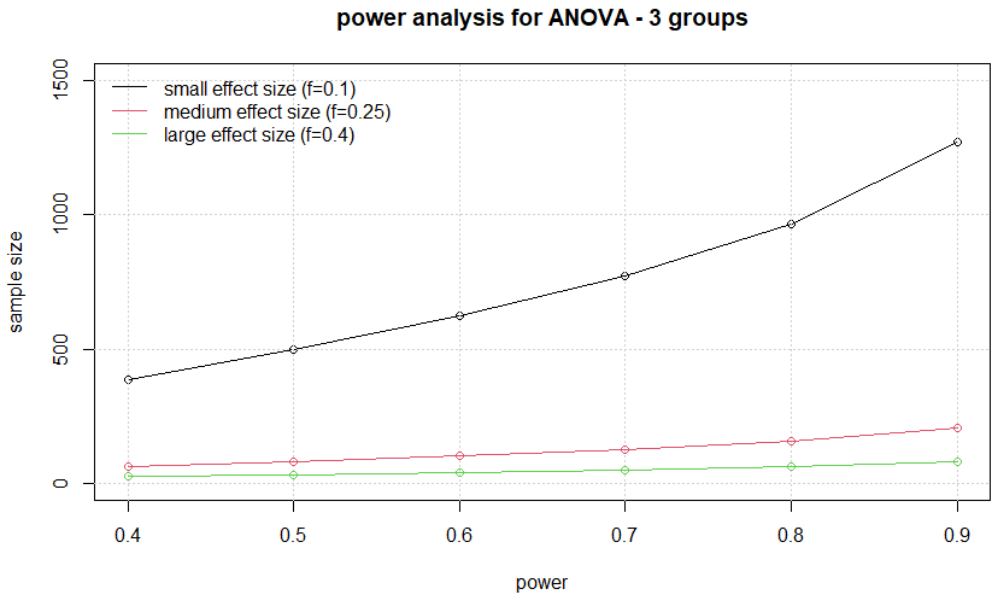
**Statistical methodology plan.** Three groups of patients will be available for the analysis, and a set of continuous measurements will have to be compared against them, along with a small set of demographic variables. Statistical significance will be assessed regarding the miRNA set of continuous measurements. For that purpose, and since miRNA variables will be estimated once, 1-way analysis of variance (ANOVA) will be applied for the detection of differences. Also, pairwise post-hoc comparisons will be held via Bonferroni or Tukey test. Due to the fact that the three groups of patients imply a logical ordering, moreover one may divide the patients in two groups, by merging neighboring groups: for example, as in A-B versus C, or A versus B-C. The latter scenario is useful in the sense that one may employ ROC curve analysis, in order to identify optimal cutoffs (with an estimated sensitivity and specificity) of the candidate miRNA biomarkers. Finally, multiple candidate biomarkers will be used to predict any potential 2-class variable (as obtained by merging two groups), using the traditional logistic regression method.

**Statistical sample size calculation.** According to the aforementioned statistical plan, one may compute the estimated maximal sample size of the study, in order to achieve a given statistical power. Assuming that the statistical significance threshold will be set at 5%, the only described hypothesis testing procedure is ANOVA, and power analysis will be based on this procedure. Using the R statistical language along with the RStudio software interface, both of which are well-known open-source products, and the library “pwr”, one may yield the following results.




According to our power calculations and given budget constraints, one must obey to the “large effect size” scenario (the green line in the figure above); i.e., **to assume that the differences among the groups will be strongly considerable**. In specific, 51 patients (17 per group) suffice for a 70% of statistical power, **while 63 patients (21 per group) suffice for an 80% of statistical power**. A 90% of statistical power, finally, can be completed with 81 patients (27 per group).