Understanding the development of broncho-pulmonary dysplasia (BPD) in preterm infants: mechanisms and different forms of BPD: The PATH-BPD study

Statistical analysis plan

Descriptive phase

All data distribution will be tested for normality with Shapiro-Wilk test and distribution characteristics. Data will be then described and treated accordingly. When needed, the same analyses will be done for residuals in order to choose the most suitable statistical strategy of analysis. Basic clinical variables possibly associated with outcome development will be preliminarily described and examined with univariate analysis using χ^2, Fisher, Student or Mann-Whitney test, as appropriate. Outcomes will be BPD and moderate-to-severe BPD.

Outcome analysis

Outcome data will be analyzed with repeated measures-ANOVA or the non-parametric equivalent (i.e. Friedman test), followed, if needed, by Student or Bonferroni or Sidak or Dunnett or Dunn or Conover-Iman *post-hoc* test, as appropriate. *Post-hoc* tests will be done between-patients and within-patients (if clinically relevant). Their non-parametric equivalents will be used if needed (i.e. in case of non-normal distribution). Depending on the pathobiological plausibility, the data distribution and the other necessary assumptions, MANOVA or ANCOVA or related procedures will be also performed.

Subsequently multivariate binary logistic regressions will be performed having as the outcome as dependent variable and the pathophysiology and pathobiology traits as independent variables (e.g. eLUS aeration score, SpO_2/FiO_2 and $PtcO_2/FiO_2$ ratio, $PtcCO_2$, non-exhaustive list). Several different models will be created based on the risk of multi-collinearity, the pathobiological plausibility and results of univariate analysis. Multi-collinearity will be evaluated with the variable inflation factor and the goodness-of-fit will be investigated with the Hosmer-Lemeshow test.

Finally, as additional analysis, receiver operator characteristics curve (ROC) procedure was performed to investigate the reliability of some pathophysiological variables in respect to the outcomes.

The analysis will be amended or expanded based on Reviewers' suggestions, once data will be submitted for publications to major scientific journals. The areas under the ROC curves (AUC), will be compared with DeLong test. Analyses will be performed with SPSS, JASP or MEDCALC and $p < 0.05$ were considered statistically significant.