Statistical methods – further details

***Sample size considerations***

The initial plan was to estimate the minimum necessary number of people required for a robust multiple logistic regression for a response defined by PPC occurrence as a binary event.

Given the observational nature of the study, a series of confounders (demographics and clinical hypothesized to be potentially associated with the occurrence of post operative pulmonary complications (PPC) needed to be taken into consideration.

Assuming that around 10 explanatory variables would suffice to explain the variability in the outcome and given the nature of this binary outcome, statistical literature suggested that the study should collect between 10 and 16 adverse events per each potential explanatory variable for robust results, i.e. for a power of at least 80% equivalent to a type 2 error of less than 20% as proven by simulation studies (1-4). So, say somewhat of a minimum of 10×10=100 such events (PPC) would be required. The prevalence of such adverse events in this population was between 20% and 30%. Therefore, we aimed at recruiting a total minimum number of between 340 and 500 patients. It turned out that 48% of the patients developed the event which makes our study reliable for statistical estimation in terms of power. This method of approximation the sample size for such models also hold for time-to-event data.

***Methods of analysis***

The collected variables have been explored both graphically and by summary statistics. Descriptive statistics are presented by the outcome and overall; means, standard deviations, percentiles, ranges for continuous variables and proportions for categorical/binary data.

The data exhibited hierarchical structure given the 28 hospitals resources and this type of clustering, reflected in adjusted standard errors, has been accounted for in all analyses. There was not much difference in standard errors whether the analyses accounted for the clustering or not. This suggests that the variability between hospitals did not add much to that of the within hospitals, most probably because this population’s inherent high variability.

Multiple PPC per individual were also observed but the clinical question of interest was whether patients had at least one PPC within next 7 days since surgery as clinically defined in the paper. The timing of these failures were collected as days (1, 2,..,7). That had implications on the choice of the survival setting – that of parametric interval censored data (Weibull as it also exhibits proportional hazard nature as Cox analysis also requires). That was preferred to discrete survival analysis. Multiple failures per individual were also considered but there were no qualitative differences in the results and hence the first failure outcome was chosen for the clarity of the presentation and for clinical rationale.

Crude associations between statistical outcomes of interest and the available explanatory candidates have been performed using Weibull survival setting on all the clinically or epidemiologically hypothesised as potentially associated with the outcome. P-values less than 0.05 were deemed as a statistically significant.

Akaike Information Criterion (AIC) has been used for model choice (on similar number of complete observations). The interpretation was made for the model containing bundle components as predictors (AIC=2198.269) rather than the model containing the bundle as derived variable (AIC=2202.216). Table 3 presents both models.

Predictions are made to highlight the effect of the explanatory variables left in the model.

Data were entered in ACCESS 2013 (Microsoft ®) and analysed in STATA *(Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP).

References:

1. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression” by Eric Vittinghoff and Charles E. McCulloch, American Journal of Epidemiology (2006) Vol. 165, No 6.

2. Concato J, Peduzzi P, Holfold TR, et al. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. J Clin Epidemiol 1995;48:1495–501.

3. Peduzzi P, Concato J, Feinstein AR, et al. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995;48:1503–10.

4. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.

***Missing data analyses***

There are only 2.3% missing observations for the outcome (PPC) and there is no suggesting that any of the available variables is associated to the missing outcome (results of univariate logistic regressions with the outcome defined as 1-missing and 0-not missing not shown).

As for the bundle which exhibits 63 (11.1%) missing observations, two extreme scenarios have been considered in which all the missing observations were assumed to have had the bundle wither applied or not. The results (Table 1 below), now based on 529 (93%) observations, are consistent with those based on complete data analyses.

Imputation methodologies for interval-censored data can be intricate (Patrick Royston, *Multiple imputation of missing values: further update of ice, with an emphasis on interval censoring*, The Stata Journal (2007) 7, Number 4, pp. 445–464). However, further analyses, assuming that the 2.3% of the missing outcome belonged to the patients who either have experienced PPC or not and repeating the scenarios above did not reveal a different qualitative statistical picture.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **All failures (275) model assuming the bundle is NOT applied when the observations are missing** | | | | | | | |
|  | **HR** | **SE** | **Z** | **P-value** | **95%CI low** | **95%CI high** |  |
| **Age (5 year effect)** | 1.059354 | .0120317 | 5.08 | 0.000 | 1.036033 | 1.0832 |  |
| **Intra-operative Fio2 % (5% increase)** | 1.080127 | .0304584 | 2.73 | 0.006 | 1.02205 | 1.141505 |  |
| **Peak inspiratory volume (<=30 vs >30)** | .4665539 | .1023628 | -3.47 | 0.001 | .3034916 | .7172278 |  |
| **Bundle (YES vs. NO)** | 1.075791 | .1763208 | 0.45 | 0.656 | .7802198 | 1.483333 |  |
| **CONSTANT** | .6147455 | .1404311 | -2.13 | 0.033 | .3928699 | .9619267 |  |
| **All failures (275) model assuming the bundle IS applied when the observations are missing** | | | | | | | |
| **Age (5 year effect)** | 1.059402 | .012042 | 5.08 | 0.000 | 1.036061 | 1.083269 |  |
| **Intra-operative Fio2 % (5% increase)** | 1.08 | .0307192 | 2.71 | 0.007 | 1.021439 | 1.141919 |  |
| **Peak inspiratory volume (<=30 vs >30)** | .4700608 | .1044671 | -3.40 | 0.001 | .3040756 | .7266519 |  |
| **Bundle (YES vs. NO)** | .960533 | .1175903 | -0.33 | 0.742 | .7556255 | 1.221007 |  |
| **CONSTANT** | .6161032 | .1409778 | -2.12 | 0.034 | .3934413 | .964777 |  |

Table 1: Extreme scenarios for missing values of the bundle. The estimates do not change and the adjusted p-value for the association between the bundle and PPC