Updated statistical analysis plan

(finalized January 20th 2018, and then endorsed by the Trial Steering Committee before the last inclusion, locking the database and beginning statistical analysis)

1.1 Study outcomes

The primary study outcome is a collapsed composite of all postoperative pulmonary complications (PPC) developing within the first five postoperative days. Patients who develop a least one complication are considered as meeting the primary endpoint. Whereas single PPC have been defined elsewhere these include mild, moderate and severe respiratory failure, ARDS; bronchospasm, new pulmonary infiltrates, pulmonary infection, aspiration pneumonitis, pleural effusion, atelectasis, cardiopulmonary edema and pneumothorax. Secondary outcomes include: 1) collapsed severe PPC (any adverse pulmonary events except mild respiratory failure); 2) postoperative extra-pulmonary complications; 3) intra-operative adverse events as defined in the original protocol; 4) unexpected need for intensive care unit (ICU) admission; 5) postoperative wound healing (defined as interruption in the timely and predictable recovery of mechanical integrity in the injured tissue); 6) in-hospital mortality and; 7) hospital-free days at postoperative day 90.

1.2 Data collection and management

Data collection is performed using electronic case report forms in Research Electronic Data Capture (REDCap™) hosted at the Coordinating Center for Clinical Trials of the University of Dresden, Germany. REDCap is a Secure Sockets Layer (SSL)-encrypted, password-protected, web-based application designed to support data capture for research. The system had the following functions: randomization, patient registration, data input, data cleaning, and data export for statistical analysis. Local investigators entered data directly into the system. Instructions for using the system are available to investigators at all times. Electronic files are archived on the University of Dresden–based server in a secure and controlled environment to maintain confidentiality. Electronic documents are controlled with password protection according to best practices.

The objective of the clinical data management plan is to provide high–quality data by adopting standardized procedures to minimize the number of errors and missing data, and consequently, to generate an accurate database for analysis. Two independent monitors are installed to perform study monitoring. Remote monitoring is performed to signal early aberrant patterns, issues with consistency, credibility and other anomalies. On–site monitoring comprised controlling presence and completeness of the research dossier and the informed consent forms, and source data checks will be performed in the files of 25 % of the patients.

1.3 Cleaning and locking of the database

The database will be locked as soon as all data are entered and all discrepant or missing data are resolved – or if all efforts are employed and we consider that the remaining issues cannot be fixed. At this step, the data will be reviewed before database locking. After that, the study database will be locked and exported for statistical analysis. At this stage, permission for access to the database will be removed for all investigators, and the database will be archived.

1.4   Missing data

No or minimal losses to follow–up for the primary and secondary outcomes are anticipated. Complete–case analysis will be carried out for all the outcomes, that is, excluding patients with missing data in the outcome of interest. However, if more than 1% of missing data were found for the primary outcome, a sensitivity analysis using multiple imputations and estimating–equation methods will be carried out.

1.5   Predefined statistical analysis plan

There were no major adjustments from the preliminary analysis plan, as reported previously. In accordance, all statistical analyses will be conducted according to the modified intention–to–treat principle, considering all patients in the treatment groups to which they were randomly assigned, excluding cases lost to follow-up due to withdrawal of consent or cancellation of surgery.

Continuous distribution of the data will be assessed by visual inspection of histograms and D'Agostino–Pearson's normality tests. For both arms, the baseline characteristics will be expressed as counts and percentages, means and standard deviations (SD), or medians and interquartile ranges (IQR) whenever appropriate.

Hypothesis tests will be two–sided with a significance level of 5% with except of the primary outcome, due to the correction for the interim analyses. We will not adjust p–values for multiple comparisons. Analyses will be performed using the R (R Core Team, 2016, Vienna, Austria) program.

Trial profile: Patient flows will be presented in a CONSORT flowchart.

Baseline comparisons: Patient's baseline characteristics will be presented by study arm.

Adherence to study interventions and ventilatory variables: Surgical and perioperative characteristics will be reported. Ventilatory variables and vital signs will be reported after intubation, one hour of surgery and last hour of surgery and compared between the two groups.

Primary outcome: The effects of the intervention on incidence of PPC will be reported as number and percentages and estimated with risk ratio and 95% confidence intervals calculated with Wald's likelihood ratio approximation test and with χ^2 tests for hypothesis testing. The two–sided α–level for the primary outcome is 0.044 to account for the interim analyses. Kaplan–Meier curves will be used to report time to PPC. Curves will be compared with the log–rank tests.

Secondary outcomes: The number and percentages of severe PPC, intra-operative adverse events, unexpected need of ICU admission, postoperative wound healing, postoperative extra-pulmonary complications and in-hospital mortality will be reported. The effect of the intervention on these outcomes will be assessed with risk ratio and 95% confidence intervals calculated with Wald's likelihood ratio approximation test and with χ^2 tests for hypothesis testing. The effects of the intervention on hospital-free days at day 90 will be estimated with a Student's t–test and reported as the mean difference between the two groups. The consistency of the findings of the Student's t-test for the hospital-free days at day 90 will be confirmed according to the mean ratio calculated by a generalized additive model considering a zero-inflated beta distribution.

Subgroup analyses: Treatment effects on incidence of PPC will be analyzed according to the following subgroups: 1) non-laparoscopic *versus* laparoscopic; 2) BMI < 40 kg/m^2 *versus* BMI ≥ 40 kg/m^2; 3) baseline SpO_2 < 96% *versus* SpO_2 ≥ 96%; 4) peripheral procedures *versus* upper abdominal (surgical incision involves the abdominal wall cranial of the

umbilicus; and 5) waist-to-hip ratio < 1.0 *versus* waist-to-hip ratio ≥ 1.0. The effects on subgroups will be evaluated according to the interaction effects between each subgroup and the study arms by generalized linear models and presented in a forest plot.

Other exploratory analyses: As a sensitivity analysis, the effect of the intervention on the primary outcome will be re-estimated using a generalized linear mixed-effect model with stratification variable (center) as random effects. Since the primary outcome of the present study is a composite one, the choice of the statistical method is an important part of designing because various methods provide different power, depending on the situation. In addition to the standard analysis described above, the following analyses will be performed:

- Count analysis: the number of positive component events (i.e., "count") across the composite will be assessed. The groups will be compared on the count using a Mann-Whitney test, and the odds ratio with the 95% confidence interval will be assessed with a proportional odds logistic regression model;

- Individual component analysis: the effect of the intervention in each component will be analyzed using a generalized linear model using a Bonferroni correction for multiple comparisons. The 99.58% Bonferroni-corrected confidence intervals will be reported ($1 − 0.05/12 = 0.9958$);

- Common effect test: A multivariate (i.e., multiple outcomes per subject) generalized estimating equations (GEE) model will be used to estimate a common effect odds ratio across the components;

- Average relative effect test: The average relative effect test will be assessed by averaging the component-specific treatment effect from the distinct effects model, and testing whether the average is equal to zero. In the GEE distinct effect model a distinct treatment effect is estimated for each component;

- Heterogeneity of treatment effect: Heterogeneity of treatment effect across components will be assessed by a treatment-by-component interaction test in the distinct effects GEE model; and

- Clinical severity weight: Each component will be weighted by a clinical severity weight determined *a posteriori*. A multivariate (i.e., multiple outcomes per subject) GEE model will be used to estimate a common effect odds ratio across the components while applying the severity weights.