Biomarker-guided intervention to prevent acute kidney injury after major surgery The prospective multicenter randomized controlled interventional trial BigpAK-2

Statistical Analysis Plan (SAP)

Version: Final version
Date: 12 May 2023

NCT04647396

Acronym: BigpAK-2

Sponsor: University Hospital Muenster

Collaborator: BioMérieux

Biostatistician and author of the SAP
Dr. rer. nat. J. Gerß, Dipl.-Stat.
Institute of Biostatistics and Clinical Research
University of Muenster

Signatures

Principal Coordinating Investigator

Univ.-Prof. Dr. med. A. Zarbock

Münster, 12.05.2023

Place, date

Signature

Biostatistician

Dr. rer. nat. J. Gerß, Dipl.-Stat.

Münster, 12 May 2023

Place, date

Signature

Contents

1 Background of the Study 4
1.1 Study objectives 4
1.2 General study design and plan 4
2 Analysis populations 4
2.1 Full analysis set 4
2.2 Per protocol population 4
2.3 Safety population 4
3 Endpoints 5
3.1 Primary endpoint 5
3.2 Secondary endpoints 5
4 Statistical Analyses 6
4.1 Baseline Characteristics 6
4.2 Concomitant medication 6
4.3 SOFA Score 6
4.4 Hemodynamics 6
4.5 Laboratory parameters 6
4.6 Renal parameters 6
4.7 Adherence to the KDIGO bundle 6
4.8 Clinical outcomes (intention-to-treat) 6
4.9 Adverse events 6
5 Statistical Methods 7
5.1 Randomization and Blinding 7
5.2 Descriptive summary of study data 7
5.3 Primary statistical analysis: Basic concepts 7
5.4 Secondary statistical analyses 8
5.5 Sample size calculation 9
5.6 Primary statistical analysis: Details of the interim analysis, sample size adaptation and final analysis 9
5.7 Operating characteristics 11
6 Treatment of missing values 11
7 Software 11

1 Background of the Study

1.1 Study objectives

The goal of this trial is to investigate the effect of the implementation of the KDIGO bundle[1] in patients at high risk for AKI after major surgery compared to standard of care in the same patient population. This biomarker-guided approach (individualized therapy) enables to treat patients at high risk for AKI prior to a functional damage of the kidneys.

1.2 General study design and plan

International, multicenter, two arm, randomized, controlled, parallel-group clinical trial with masked participants and outcomes assessors

2 Analysis populations

2.1 Full analysis set

The full analysis population contains all patients who were randomized.

2.2 Per protocol population

The per protocol population contains all patients without major protocol deviations.
Major protocol deviations that lead to exclusion from the per protocol population are

1. violation of inclusion/exclusion criteria
2. noncompliance to the randomized intervention

2.3 Safety population

The full analysis population contains all patients who were randomized.

[1] Bundle recommended by the "Kidney Disease: Improving Global Outcomes Group" (KDIGO bundle) for at least 12 hours:

1. discontinuation of all nephrotoxic drugs when possible
2. optimization of volume status and hemodynamic parameters (consideration of a functional hemodynamic monitoring)
3. close monitoring of serum creatinine, fluid balance and urinary output
4. avoidance of hyperglycemia
5. considerations of alternatives to radiocontrast agents
6. discontinuation of angiotensin converting enzyme inhibitors and angiotensin receptor blockers in the perioperative period
7. avoidance of HES, gelatin, and chlorid-rich solutions

3 Endpoints

3.1 Primary endpoint

Occurence of moderate or severe AKI within 72 hours after start of the intervention

3.2 Secondary endpoints

1. Adherence to the KDIGO bundle
2. Occurrence and severity of any stage of acute kidney injury within 3 days after surgery
3. Change in biomarker values during 12 hours after initial measurement of [TIMP2]*[IGFBP7]
4. Free-days of mechanical organ support through to day 3
5. Free-days of vasopressors through day 3
6. Need of RRT at day 30 and 90
7. Duration of RRT at day 30 and 90
8. Renal recovery at day 90
9. 30-day and 90-day mortality
10. ICU length-of-stay and Hospital length-of-stay
11. Major adverse kidney events (MAKE90) (defined as the composite of death, use of RRT and persistent renal dysfunction (defined as serum creatinine ≥2x to baseline value at hospital discharge)) at day 90

4 Statistical Analyses

4.1 Baseline Characteristics

4.1.1 Number of randomized patients
4.1.2 Demographic data (age, sex, BMI, comorbidities, medication)
4.1.3 Admission diagnosis and category of admission (elective, emergency)
4.1.4 Surgical parameters (nature of surgical procedure, duration of surgical procedure, fluid input and output, intraoperative transfusion)
4.1.5 APACHE II
4.1.6 Medical history

4.2 Concomitant medication

4.3 SOFA Score

4.4 Hemodynamics

4.4.1 MAP
4.4.2 HR
4.4.3 CVP
4.4.4 CI

4.5 Laboratory parameters

4.5.1 Glucose
4.5.2 Lactate
4.5.3 Hemoglobin
4.5.4 Na
4.5.5 K

4.6 Renal parameters

4.6.1 BUN
4.6.2 Creatinine
4.6.3 GFR
4.6.4 Urine output
4.6.5 Fluid balance

4.7 Adherence to the KDIGO bundle

4.8 Clinical outcomes (intention-to-treat)

4.8.1 AKI (yes/no, severity, diagnosis by, transient/persistent, RRT dependency, type of RRT, RRT days)
4.8.2 Mortality
4.8.3 Duration of mechanical organ support (Begin and end of RRT, ECMO, and mechanical ventilation)
4.8.4 Length of stay (ICU, Hospital)
4.8.5 Renal Recovery
4.8.6 Need of RRT
4.8.7 MAKE

4.9 Adverse events

5 Statistical Methods

Statistical analyses will be performed according to the principles of the ICH guideline E9 'Statistical Principles for Clinical Trials'.

5.1 Randomization and Blinding

Treatment assignment will be accomplished using REDCap and randomization will be performed centrally in a 1:1 proportion, stratified by center. Patients will enter the observation period immediately after randomization.

Neither the patient, nor the study personnel can be blinded in regards to the treatment assignment. The individuals involved in adjudication of endpoints or complications will be blinded to treatment assignment.

Randomization will be checked by comparing demographical and clinical baseline variables between the randomized groups. Differences between treatment groups will be quantified as follows.

1. Normally distributed continuous variables:
   Student's two-sample t test, absolute mean difference between groups
2. Non-normally distributed continuous variables:
   Wilcoxon two-sample test (t approximation), Hodges-Lehmann estimator of location shift
3. Categorical variables: Absolute difference of percentages, odds ratio
4. Censored variables: Absolute difference of percentages, hazard ratio

5.2 Descriptive summary of study data

All continuous variables will be summarized using the following descriptive statistics: n (non-missing sample size), number of missing values, mean, standard deviation, median, first and third quartile, minimum, maximum.

For categorical variables the frequency and percentages of observed levels will be reported based on the non-missing sample size. All summary tables will be structured with a column for each treatment and will be annotated with the total population size relevant to that table/treatment, including any missing observations.

Censored variables will be summarized using Kaplan Meier estimation.

5.3 Primary statistical analysis: Basic concepts

The primary endpoint occurrence of moderate or severe AKI (stage 2 or 3) within 72 hours after surgery will be evaluated. It is intended to show superiority of the interventional treatment compared to standard care.

The null hypothesis H_0: $AKI_1=AKI_0$ will be tested at a significance level of 5% against the two-sided alternative H_1: $AKI_1 \neq AKI_0$, where AKI_1 denotes the primary endpoint rate under the interventional treatment and AKI_0 under standard care.
Formally the two-sided test will be decomposed into two one-sided tests of the null hypotheses H_{0a}: $AKI_1 \geq AKI_0$ and H_{0b}: $AKI_1 \leq AKI_0$ against the respective alternatives. The significance level of each of the one-sided tests will be 2.5%.

A group sequential adaptive plan with one interim analysis will be applied. The interim analysis will be performed applying the inverse normal method based on a group sequential plan with O'Brien and Fleming type alpha spending function without futility stop. The interim analysis is planned to be conducted when half of the total number of patients have been recruited (information rate 0.5). The primary statistical analysis will include all randomized patients (full

analysis set) and will be performed according to the intention-to-treat principle (ITT) in order to prevent attrition bias. The randomized groups will be compared with a Cochrane Mantel Haenszel test[2] with stratification by center. The results of the interim analysis will be compared to the final results in order to find out whether changes in practice may already have occurred during the study period.

The decision bounds of the group sequential adaptive plan for the one-sided tests are shown below.

Information rate	bounds accept H_0	bounds reject H_0	sign.level one-sided	α spent
0.5	-	2.963	0.0015	0.0015
1.0	1.969	1.969	0.0245	0.0250

The primary statistical analysis provides confirmatory statistical evidence.

5.4 Secondary statistical analyses

Secondary outcomes will be evaluated in the full analysis set according to ITT. Secondary statistical analyses are intended to be exploratory (hypothesis generating) and will be interpreted accordingly. A two-sided p-value of ≤ 0.05 will be considered as noticeable ("significant") without adjustment for multiplicity.

Statistical analysis of the pre-specified secondary endpoints will be performed with descriptive and inferential statistical methods. In inferential statistical analyses two-sided significance tests will be applied with a local significance level alpha=0.05. All point estimates of parameters of interest will be supplemented by 95% confidence intervals unless otherwise stated.

Secondary outcomes will be evaluated and compared between the randomized treatment arms with stratification by center using the following methods. In case of normally or non-normally distributed metric outcomes, a two-way ANOVA or the van Elteren test will be applied, respectively. Categorical outcomes will be analyzed using (generalized) Cochrane Mantel Haenszel tests. Survival data will be analysed using the Kaplan-Meier method and the stratified log-rank test. Hazard ratios will be estimated based on stratified Cox proportional hazards models after checking the proportional hazards assumption based on the Schoenfeld residuals.

The duration of RRT will be evaluated in two ways.
1. If a patient dies or reaches his/her end of follow-up, the duration of RRT is finished.
2. In a competing risk analysis, besides the event of interest end of RRT, death (while RRT otherwise would be ongoing) is considered a competing event. If a patient reaches his/her end of follow-up, the duration of RRT is censored.

The duration of hospital stay will be evaluated in two ways.
1. If a patient dies or reaches his/her end of follow-up, the hospital stay is finished.
2. In a competing risk analysis, besides the event of interest end of hospital stay, death (while hospital stay otherwise would be ongoing) is considered a competing event. If a patient reaches his/her end of follow-up, the duration of hospital stay is censored.

The duration of ICU stay will be evaluated in two ways.
1. If a patient dies or reaches his/her end of follow-up, the ICU stay is finished.

[2] A one-sided version of the Cochrane Mantel Haenszel test (CMH) is constructed, using the fact that the CMH test statistic – if it is not squared – follows a standard normal distribution under the null hypothesis H_0: $AKI_1=AKI_0$ (in all strata).

2. In a competing risk analysis, besides the event of interest end of ICU stay, death (while ICU stay otherwise would be ongoing) is considered a competing event. If a patient reaches his/her end of follow-up, the duration of ICU stay is censored.

5.5 Sample size calculation

Sample size and power calculations are based on a group sequential adaptive plan with one interim analysis and the primary endpoint occurrence of moderate or severe AKI (KDIGO stage 2 or 3) within 72 hours after surgery. The expected AKI rate (≥ stage 2), substantiated by published data of the BigpAK trial, in the control group and the intervention group is 20% and 14%, respectively. The interim analysis will be performed when 309 evaluable patients have been recruited in each treatment group (2x309=618 patients in total). Using the results of the interim analysis, the sample size of the final analysis will be recalculated (see below). If no sample size adjustment is found to be necessary, the final analysis will be performed when additional 309 evaluable patients have recruited in each treatment group after the interim analysis (2x309+2x309=1236 evaluable patients in total). A number of up to 5% of recruited patients are expected to be lost to follow up and in the worst case have completely non-evaluable data. Therefore, a total number of 1302 patients will be recruited in order to provide a number of 1236 evaluable patients.

The characteristics of the group sequential adaptive plan for the one-sided tests are shown below.

Information rate	bounds accept H_0	bounds reject H_0	sign.level one-sided	α spent	β spent	power achieved	stage n_1	sizes n_2
0.5	-	2.963	0.0015	0.0015	-	0.1647	309.0	309.0
1.0	1.969	1.969	0.0245	0.0250	-	0.8010	309.0	309.0

Calculations were performed using the ADDPLAN software.

5.6 Primary statistical analysis: Details of the interim analysis, sample size adaptation and final analysis

The interim analysis will be performed using data of n_1 := 2x309 = 618 evaluable patients (309 patients per treatment group).
If the p-value of the one-sided null hypothesis H_{0a}: $AKI_1 \geq AKI_0$ is $p_a^{interim} \leq 0.0015$, the trial is finished. It will be concluded that the primary endpoint rate under the interventional treatment AKI_1 is lower than under standard care (AKI_0), and that the interventional treatment is superior to standard care.
If the p-value of the one-sided null hypothesis H_{0b}: $AKI_1 \leq AKI_0$ is $p_b^{interim} \leq 0.0015$, the trial is finished. It will be concluded that the primary endpoint rate under the interventional treatment AKI_1 is higher than under standard care (AKI_0), and that the interventional treatment is inferior to standard care.
If either $p_a^{interim} \leq 0.0015$ or $p_b^{interim} \leq 0.0015$, recruitment will stop as soon as possible. The primary final analysis will be performed with all patients who have been included in the interim analysis. Additional sensitivity analyses will be performed, additionally including “interim patients” who have been recruited between the data base lock of the interim analysis and the final recruitment stop.
If both p-values of the one-sided null hypotheses H_{0a}: $AKI_1 \geq AKI_0$ and H_{0b}: $AKI_1 \leq AKI_0$ are >0.0015, the trial will be continued and the following procedure will be applied.

Let $p_a^{interim}$ denote the p-value of the one-sided null hypothesis H_{0a}: $AKI_1 \geq AKI_0$ in the interim analysis. Based on $p_a^{interim}$ the number of additional evaluable patients n_2 (recruited after the interim analysis) will be determined as follows.

- If $p_a^{interim} \leq 0.0015$, the trial is finished after the interim analysis (see above).
- If $0.0015 < p_a^{interim} \leq \alpha_{interim}$,
  an additional number of n_2=2x250=500 evaluable patients will be recruited.
  The final statistical analysis will be performed using data of n_2=2x250=500 additional patients, who were recruited after the interim analysis.
  The total number of evaluable patients is n_1+n_2 = 618+500 = 1118 (559 patients per treatment group).
- P-values in the range $p_a^{interim} > \alpha_{interim}$ include those with $p_a^{interim} > 0.5$. P-values $p_a^{interim} > 0.5$ occur when in the interim analysis the observed primary endpoint rate under the interventional treatment AKI_1 is higher than under standard care (AKI_0), indicating that the interventional treatment is inferior to standard care.
  If $p_a^{interim} > \alpha_{interim}$ (or even >0.5),
  an additional number of n_2=2x400=800 evaluable patients will be recruited.
  The final statistical analysis will be performed using data of n_2=2x400=800 additional patients, who were recruited after the interim analysis.
  The total number of evaluable patients is n_1+n_2 = 618+800 = 1418 (709 patients per treatment group).

The parameter $\alpha_{interim}$ was specified at the time when the present Statistical Analysis Plan (SAP) was written (12 May 2023). Since then, its value is finally determined and will not be changed. The specified value of $\alpha_{interim}$ however is not included in the present SAP. It must be kept confidential until the end of the trial, in order to preserve the integrity of the trial. An appendix to the SAP was written at the same time as the present SAP (12 May 2023). The appendix to the SAP contains the specified value of $\alpha_{interim}$. The appendix to the SAP is sent to the Data Safety Monitoring Board (DSMB). The DSMB deposits the appendix to the SAP until the end of the trial. It guarantees that the appendix to the SAP is not assessed by the Principal Coordinating Investigator or any other study personell. After the end of the trial, the DSMB will check and confirm that the trial was conducted applying the previously specified value of $\alpha_{interim}$.

If the trial is continued after the interim analysis, the information about the further recruitment will be reported to the Principal Coordinating Investigator and all other study personell in a stepwise way. That means after the interim analysis in a first step the Principal Coordinating Investigator will be informed only, if the trial will be finished after the interim analysis (n_2=0), or if it will be continued. If the trial is continued, in the first step the Principal Coordinating Investigator will not get to know the sample size n_2 yet. If the sample size n_2 is n_2=500 or n_2=800, will be reported to the Principal Coordinating Investigator only in a second step, at the time when additional 500 evaluable patients have been recruited. At that time, the Principal Coordinating Investigator will be informed, if the final sample size is reached (n_2=500), or if recruitment will be continued up to additional 300 evaluable patients (so that n_2=500+300= 800).

In the final statistical analysis, data of n_2 additional patients will be evaluated, who were recruited after the interim analysis (n_2=500 or n_2=800, see above).
The final overall p-value p_a^{final} of the one-sided null hypothesis H_{0a}: $AKI_1 \geq AKI_0$ is defined as the minimal possible one-sided significance level, that could have been specified, so that the final statistical analysis would result in rejection of the null hypothesis H_{0a}: $AKI_1 \geq AKI_0$ against the one-sided alternative H_{1a}: $AKI_1 < AKI_0$. If the final overall p-value is $p_a^{final} \leq 0.025$, it will be concluded that the primary endpoint rate under the interventional treatment AKI_1 is lower than under standard care (AKI_0), and that the interventional treatment is superior to standard care.
The final overall p-value p_b^{final} of the one-sided null hypothesis H_{0b}: $AKI_1 \leq AKI_0$ is defined as the minimal possible one-sided significance level, that could have been specified, so that the final statistical analysis would result in rejection of the null hypothesis H_{0b}: $AKI_1 \leq AKI_0$ against the one-sided alternative H_{1b}: $AKI_1 > AKI_0$. If the final overall p-value is $p_b^{final} \leq 0.025$, it will be

concluded that the primary endpoint rate under the interventional treatment AKI_1 is higher than under standard care (AKI_0), and that the interventional treatment is inferior to standard care. The final overall two-sided p-value p^{final} is defined as the minimum of the doubled one-sided p-values p_a^{final} and p_b^{final}, i.e. $p^{final} = \min(2 \cdot p_a^{final}, 2 \cdot p_b^{final})$.

All reported numbers of patients refer to patients with evaluable primary endpoint data. The corresponding numbers of recruited patients may be increased in order to compensate for missing data.

5.7 Operating characteristics

The operating characteristics of the applied design and sample size adaptation are shown below, including the total numbers of patients, percent stops after stage 1, and overall power. The operating characteristics were determined with the R software (RPACT package) in 1000000 simulation runs. Simulated data were generated assuming equal primary endpoint rates across all strata under standard care (AKI_0=20% in all strata), as well as under the interventional treatment (AKI_1=10%/14%/15% in all strata). The statistical analysis of simulated data was performed with Pearson's chi-squared tests.

Expected primary endpoint rate under standard care AKI_0	20%	20%	20%	20%
Expected primary endpoint rate under the interventional treatment AKI_1	10%	14%	15%	20%
Total number of patients: 1. Minimum 2. Average 3. Maximum	 618 772.8 1418	 618 1145.1 1418	 618 1222.8 1418	 618 1400.8 1418
Percent stops after stage 1	70.8%	16.3%	9.1%	0.3%
Overall power (rejection of H_{0a})	99.9%	82.1%	66.0%	2.5%

6 Treatment of missing values

Missing values that may arise in effectiveness or safety parameters will not be replaced by any kind of statistical imputation.

7 Software

Statistical analyses will be performed using the statistical software SAS.