Target
Statistical analysis plan

Final 1.0 18.05.2020

Dr. Stefan Hagel, Principal investigator

Date, signature

Prof. Dr. Peter Schlattmann /
Dr. T. Lehmann, Study Biometrics

Date, signature

1 Study background

1.1 Study Goal

The primary objective of the study is to determine whether optimization of antimicrobial therapy by means of patient-specific dose adjustment of the test substance piperacillin has a beneficial effect on organ function in severe sepsis or septic shock and is superior to dosing according to the technical information. This will be investigated using a global morbidity measure (mean total SOFA score).

1.2 Study design

This is a multicenter, single-blinded, randomized controlled trial.

1.3 Protocol version and amendments

The statistical analysis plan is based on the final study protocol as of 12/03/2019 (version 3.0).

2 Evaluation collectives (definition and scope)

2.1 Intention-to-Treat (ITT)

The intention-to-treat collective (ITT) includes all patients who were enrolled and randomized into the study. The analysis of the endpoints (primary and secondary) is performed in the intention-to-treat population.

2.2 Safety

Safety analysis will be performed for the intention-to-treat population.

2.3 Foreseeable serious protocol violations

Serious protocol violations (major deviations):

For inclusion:

- No sepsis or septic shock
- Piperacillin measurement not possible within 24 h after randomization
- No planned or started therapy with piperacillin.
- Age < 18 years
- Pregnancy/Lactation
- Life expectancy <28 days

Non-serious protocol violations (minor deviations):

For inclusion:

- Onset of severe sepsis or septic shock > 24 h before randomization.
- Hypersensitivity to ß-lactam antibiotics or to any of the other components of the test substance.
- Pretreatment with piperacillin (in combination with tazobactam) >24h before randomization.

- Impaired liver function (Child-Pugh C)

3 Study Centers

The plan was to include 276 patients recruited in 10-15 centers. With an expected drop-out rate of 15%, this would result in 234 evaluable patients (informed consent available and not revoked, documented primary outcome).

4 Evaluation variables

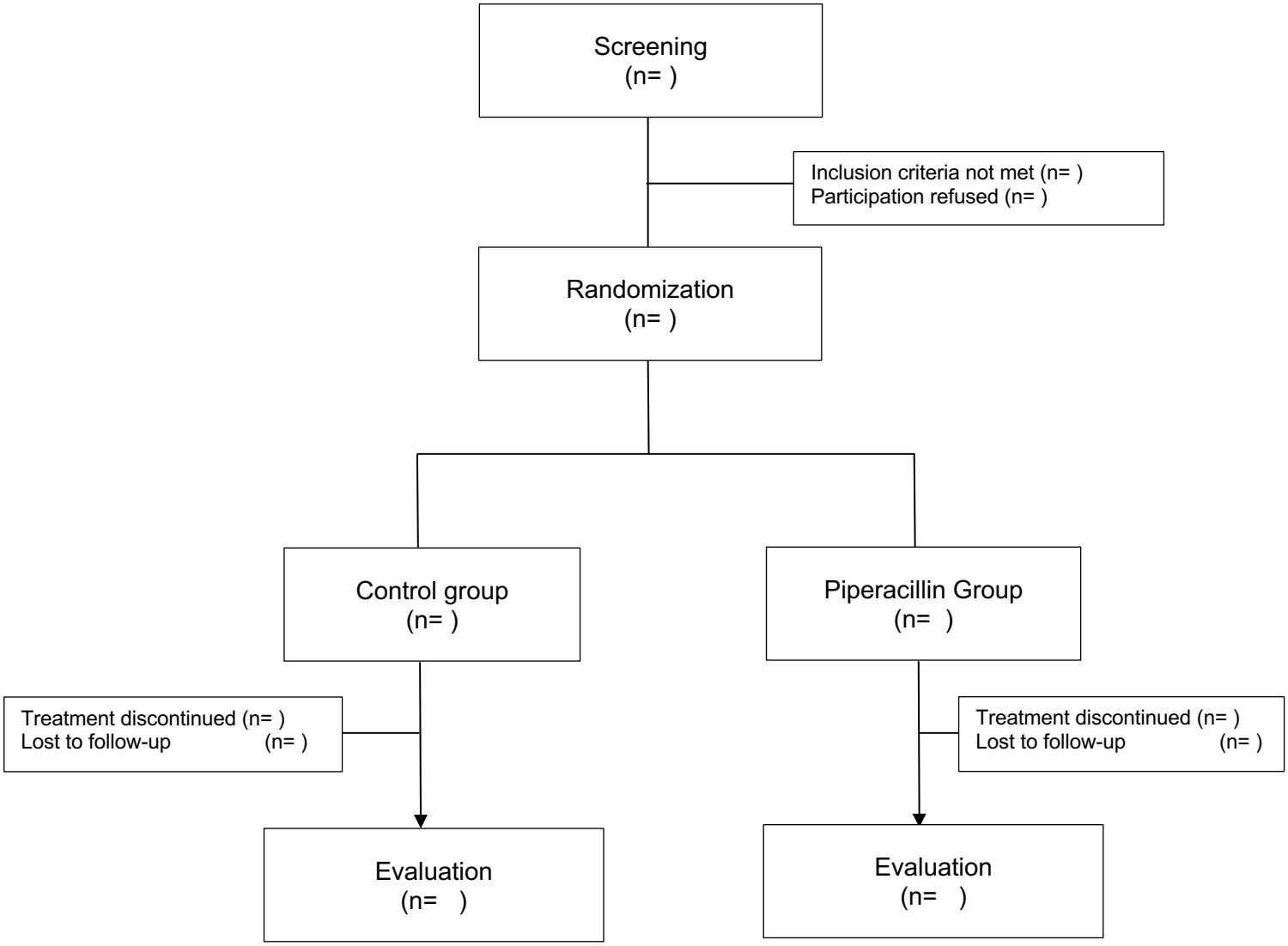
4.1 Population characteristics and baseline data

Baseline data of patients at the time of screening are used to describe the population

- Age
- Gender
- Size
- Weight
- BMI
- Physiological data (APACHE II score, Charlson comorbidity index, SAPS II, SOFA after randomization).
- Focus infection
- Origin of the infection
- Degree of diagnostic confirmation (clinical vs microbiological)
- Laboratory values (baseline lactate, baseline procalcitonin)
- Frequency septic shock
- Frequency of individual organ dysfunctions
- Time of onset of septic shock until randomization
- Frequency ventilation

compared between both groups (piperacillin, control group).

The patient flow is represented by the following figure:



4.2 Primary variable

The primary outcome measure is the SOFA total score. It is included in the analysis as an individual mean value over the course from day 1 after randomization until discharge from the ITS or until death, but at most until day 10. If the SOFA score cannot be determined on one day, the average of the remaining ITS days is calculated.
The SOFA total score is determined as the sum of the following distribution of points for each patient and day:

Respiratory

PaO2/FiO2 (mmHg)	Points
≥ 400	0
< 400	1
< 300	2
< 200 and ventilation	3
< 100 and ventilation	4

Central nervous system

Glasgow Coma Score	Points
15	0
13-14	1
10-12	2
6-9	3
<6	4

Cardiovascular

Mean arterial blood pressure (MAP) or use of vasopressors.	Points
MAP ≥ 70 without catecholamines	0
MAP < 70 mmHg	1
Dopamine ≤ 5 µg/kg/min or dobutamine (any dosage).	2
Dopamine > 5 µg/kg/min or epinephrine ≤ 0.1 µg/kg/min or norepinephrine ≤ 0.1 µg/kg/min.	3
dopamine > 15 µg/kg/min or epinephrine > 0.1 µg/kg/min or norepinephrine > 0.1 µg/kg/min	4

Liver function

Bilirubin µmol/l	Points
<20 µmol/l	0
20-32 µmol/l	1
33-101 µmol/l	2
102-204 µmol/l	3
> 204 µmol/l	4

Coagulation

Platelets × $^{103/µl}$	Points
≥ 150	0
< 150	1
< 100	2
< 50	3
< 20	4

Kidney function

Creatinine µmol/l, urine l/d	Points
<110 µmol/l	0
110-170 µmol/l	1
171 - 299 µmol/l	2
300 - 440 µmol/l or urine < 0.5 l/d	3
> 440 µmol/l or urine < 0.2 l/d	4

4.3 Secondary variables

For the duration without organ-supporting procedures, the days during which no artificial ventilation, no renal replacement therapy (in each case until day 28) or vasopressor therapy (until day 14) took place are added up for each patient.
For SOFA subscores, the mean is calculated over the course from day 1 after randomization until discharge from ITS or death, but not exceeding day 10. The other secondary endpoints are 28-day mortality, duration and cumulative dose of antibiotic therapy, number of dose adjustments/therapy cycle, antibiotic-free days to a maximum of day 14, ICU length of stay to a maximum of day 28, hospital length of stay to a maximum of day 28.

4.4 Safety variables

4.4.1 Adverse events

Adverse events (AEs and SAEs) are recorded according to the classification in the CRF and reported for both groups. A distinction is made according to severity and association with the investigational product.

4.4.2 Laboratory parameters

Laboratory parameters listed in the CRF are reported on a group-specific basis for the time points 24 h before randomization, immediately after randomization, and day 1-14.

5 Treatment of missing values and outliers

5.1 Missing values

In all analyses performed, the number of missing values is reported. For the primary endpoint, in case of a relevant number of missing values, appropriate replacement procedures (multiple imputation) are applied in the context of a sensitivity analysis.

5.2 Outlier

In the case of outliers in the secondary continuous endpoints, the non-parametric Mann-Whitney U test is used to test for group differences. The median, which is robust to outliers, is given as the positional measure on a group-specific basis.

6 Statistical methodology

6.1 Sample size calculation

The basis for the calculation of the sample size is the primary outcome measure SOFA score. Superiority of the study intervention is shown when the null hypothesis of equal expected values is rejected in favor of the alternative hypothesis. Data from the VISEP and MAXSEP trials of the SepNET Study Group (N Engl J Med 2008; 358:125-39, JAMA 2012; 307:2390-2399) have shown that a 1.4-point lower SOFA score in the intervention group would be clinically relevant. Assuming a standard deviation of 3.8 points (SepNet data), this results in an effect size of 0.368 standard deviations. To show this effect with a power of 80% using a 2-sample t-test at a two-sided significance level of 5%, 117 patients per study arm are required (software:

nQuery Advisor 7.0 Statistical Solutions). Based on the experience of the SepNet studies, a dropout rate of 15% is expected, so that a total of 276 (2x138) patients need to be randomized to reach the number of cases required for the analysis.

6.2 General principles

All variables are described by adequate non-confirmatory statistics, which includes at least: Number of subjects (values), number of missing values, mean, standard deviation, minimum, median, maximum for metric variables, and frequency tables for categorical variables.

6.3 Population characteristics and baseline data

For baseline metric data, the following location parameters and measures of dispersion per group (control, BDG) are reported:

- Mean and standard deviation for normally distributed data, otherwise median and interquartile range (25th and 75th percentile).
- Minimum
- Maximum

as well as

- Number of patients N
- Number of missing values (missings)

For categorical baseline data, report absolute and percent frequencies by group.

6.4 Pre-existing/concomitant diseases and medications

Absolute and relative frequencies per group are given for the pre-existing conditions.

6.5 Treatment-specific data

Data during the course of treatment (e.g., new infections, physiological data) will be reported for each study arm using the characteristics listed in 0

6.6 Primary Analysis

The difference between the intervention arms is evaluated confirmatorily using a mixed linear model. Fixed factors are intervention and SOFA score at the time of randomization (baseline). The change in exclusion criteria (1st amendment) allows inclusion of patients with renal insufficiency (acute or chronic) and renal replacement therapy or expected renal replacement therapy within the following 6 hours after randomization. This affects the SOFA score (kidney subscore 4 each for patients with renal replacement therapy) and may bias the intervention effect if unevenly distributed. Therefore, in the primary model, the presence of this circumstance (yes/no) at baseline is considered as an additional cofactor. As a random factor, the study center is modeled as a random intercept. The main effect of the intervention is tested at the two-sided significance level of 5%.

6.7 Secondary analyses

All secondary and further target variables are exploratively analyzed according to their scaling:

Chi2 test or Fisher's exact test for binary data:

- 28 days mortality
- Incidence of antibiotic-resistant bacteria by day 28.
- Re-infection with the same pathogen by day 28.
- Superinfection with another pathogen by day 28.

2-sample t-test or Mann-Whitney U-test for metric data:

- SOFA subscores of each organ system averaged over day 1 to day 10 after randomization.
- Duration and cumulative dose of antibiotic therapy.
- Number of dose adjustments/therapy cycle
- Antibiotic-free days until day 14
- Days without vasopressors until day 14
- Days without mechanical ventilation Day 28
- Days without renal replacement therapy until day 28
- Antibiotic therapy costs

Kaplan-Meier curves / logrank test for right-censored time-to-event data:

- Intensive care unit length of stay up to a maximum of day 28
- Hospital stay up to a maximum of day 28

Generalized linear mixed effects models for binary history data:

- Cure of infection ("Clinical Cure") on day 3, 5, 7, 10, 14 after randomization and at end of therapy with piperacillin.
- Remediation of microbiological findings ("Microbiological Cure") on day 3, 5, 7, 10, 14 after randomization and at the end of therapy with piperacillin.

Generalized linear mixed effects models for ordinal history data:

- Neurological outcome (ICDSC) days 7, 10, and 14 after randomization and at the end of therapy.

Indices resulting from analysis of pharmacokinetics (PK) and pharmacodynamics (PD) (including antibiotic concentration >4x MIC 100% (fT>4MIC)) are compared using appropriate methods mentioned above.
The frequency and incidence of adverse events are reported on a group-specific basis.
Appropriate effect measures with 95% confidence intervals are provided for the target variables.

6.8 Safety/compatibility analyses

6.8.1 Adverse events

For each of the two groups, it is stated which adverse events occurred how often (absolute and percentage frequencies). A distinction is made according to severity and relationship to treatment.

6.8.2 Laboratory parameters

For laboratory parameters, the following position and dispersion measures are reported on a group-specific basis:

- Mean value
- Standard deviation
- Median
- 25th percentile
- 75th percentile
- Minimum
- Maximum

as well as

- Number of patients N
- Number of missing values (missings)

6.9 Subgroup analyses

Exploratively, for the primary outcome measure and selected secondary outcome measures, appropriate statistical models will be used to examine whether there are differential intervention effects in patient subgroups. The following subgroups are considered:

- Patients with renal replacement therapy vs. other patients
- Patients with infections with germs that have a high minimum inhibitory concentration (MIC) vs. other patients.
- Patients with increased renal clearance (defined by creatinine clearance >130 ml/min per 173m2) vs. other patients.
- Patients with different infection focus (pulmonary, bloodstream, intra-abdominal, bone/soft tissue infections).

Differences will be statistically assessed using the Intervention*Subgroup Characteristic interaction term.
In addition, the primary outcome measure and all secondary outcome measures will be analyzed in the population of all patients surviving at least 48 hours.

6.10 Interim evaluations

No interim evaluations were performed.

6.11 Deviations from protocol and amendments

Serious deviations from the protocol are reported on a group-specific basis; these do not result in exclusion of the patient from the intention-to-treat population.

7 Software, program validation

The statistical program SAS version 9.4 (SAS Institute, Inc., Cary, North Carolina) is used for data analysis. The program validation is performed according to the validation plan BVL of the SOP BI06 of the ZKS Jena.