Statistical Analysis Plan

1 Overview

TRIAL FULL TITLE	PICASSO
SAP VERSION	1.0
SAP VERSION DATE	30.06.2020
TRIAL STATISTICIAN	Dr. Andrew Atkinson
TRIAL CHIEF INVESTIGATOR	Prof. Dr. Andreas Widmer
SAP AUTHOR	Dr. Andrew Atkinson

2 SAP Signatures

I give my approval for the attached SAP entitled PICASSO dated 30.6.20.

Chief Investigator

Name: Prof. Dr. Andreas Widmer

3 Table of Contents

4 Introduction

4.1 Preface

Surgical site infections (SSIs) are the most common nosocomial infections in surgical patients causing significant increases in morbidity, mortality, and health care costs. As they are usually caused by components of the normal skin flora, disinfection of the surgical site with an antiseptic skin preparation is standard practice prior to any surgical intervention. The most commonly used disinfectants are either chlorhexidine in alcoholic solution (CHX) or PVP iodine in alcoholic solution (PI). Currently, the best compound for prevention of surgical site infections is unknown.

4.2 Purpose of the analyses

The objective of the study is to prove non-inferiority of PI compared to CHX in preoperative skin antisepsis on the main outcome surgical site infection as defined by Swissnoso criteria and adapted to the National Nosocomial Infection Surveillance System (NNIS)-score.

5 Study Objectives and Endpoints

5.1 Study Objectives

(ICH E3; 8)

Non-inferiority of PI to CHX for preoperative skin disinfection in terms of surgical site infections in patients undergoing cardiac or abdominal surgery.

5.2 Endpoints

(ICH E9; 2.2.2)

<u>Primary endpoint</u>

The primary endpoint is presence of surgical site infection for abdominal and cardiac surgeries as defined by Swissnoso according to the National Healthcare Safety Network (NHSN) criteria. The occurrence of surgical site infections is evaluated at three time points:

1. At time of dismissal from the hospital

2. 30 days after abdominal and cardiac surgery

3. Between 90 days and 1 year for cardiac surgery

The analysis will be performed using the composite SSI rate endpoint, that is, for abdominal SSI at 30 days, and for cardiac surgery SSI at 30 days, 90 days or 1 year. The infection ratio will be calculated as a crude ratio as well as adapted to the NNIS score. The method of aggregation for the combined SSI rate for both types of surgery together will be a weighted average based on the proportion of SSIs from each surgery type.

Secondary endpoints

Subgroup analyses stratified by:

1. Surgical factors

    a. Type of SSI: superficial, deep, organ-space

    b. Type of Surgery

    c. Timing of antimicrobial prophylaxis

    d. Duration of surgery,

    e. Wound contamination –class according to CDC

2. Patient factors

    a. Gender

b. BMI

c. presence of implant (for cardiac surgery only)

d. ASA-score

e. Presence of colorectal cancer for colorectal surgery

f. Presence of diabetes mellitus, arterial hypertension, active smoking

3. Laboratory parameters such as:

hemoglobin, creatinine, leukocytes, CRPNNIS score (combination of ASA

score; wound contamination; duration of surgery)

4. Mortality (in-hospital and 30 day for abdominal surgery; in-hospital, 30 day and 365 day for cardiac surgery), length of stay

6 Study Methods

6.1 General Study Design and Plan

(ICH E3; 9)

Identify the study design, including the following

- Study configuration and experimental design: The data for this study will be based on routinely collected data from Swissnoso, as part of further use research of this data. The intervention will be conducted as an open-label cluster-randomized cross-over multicentre trial with an allocation ratio of 1:1.
- Type of control(s): We compare CHX (control) to PI (new).

- Level and method of blinding: Open-label allocation. The responsible statisticians are blinded to the allocation, personnel in the OR responsible for performing skin disinfection will not be blinded as both disinfectants are easily distinguishable by their color. This is also relevant for safety reasons in clinical routine. In addition, Infection Prevention (IP-) nurses responsible for follow-up and assessment of SSI may not be completely blinded to allocation of patients.
- Method of treatment assignment: A computer-generated randomization list will be generated.
  Each surgical study site will be randomized in clusters by month to use CHX or PI and will be switched over according to randomization.
- At what point in time subjects are randomised relative to treatments, events and study periods: The study nurse will provide the information which product must be used in the OR according to surgical subspecialty and hospital during the study month. A new randomization table will be generated in case the products are not available for > 1 week on the market, and the study will be halted during this phase. The allocation sequence will be generated by the main study nurse based on the computer-generated list. All respective surgical patients are eligible unless exclusion criteria are present or they refuse general IC. Patients are allocated by month, department, and subspecialty in this cluster-randomized trial.
- Sequence and duration of all study periods: Eligibility screening, general informed consent, allocation, skin antisepsis and collection of baseline variables are all performed at time 0 (day of surgery). All patients undergoing cardiac or any of the aforementioned abdominal surgeries at any of the three study sites will be eligible for this study unless they meet exclusion criteria. Day of surgery, allocation to intervention group according to cluster randomization. Inclusion into study as soon as surgery performed. Entry of baseline data into database (within 5 days after surgery). The departments of surgery of the three study sites will be randomized in clusters by month of surgical subspecialty to use CHX or PI. Assessment of SSIs is following Swissnoso standard procedures and comprises a routine surveillance for means of quality control in all participating centers, irrespective of participation in this study. Usually for Swissnoso, patients are followed-up by IP nurses during their hospital stay and post-discharge after 30 (-50) days and additionally after 365 days (56 weeks) in case of cardiac surgery. For this study, the follow-up time for cardiac surgery deviates from the Swissnoso standard of 365 days, with a follow-up time defined tp be 90 to 365 days (12-56 weeks).

6.2 Non-Inferiority Study

(ICH E3; 9.2, 9.7.1, 11.4.2.7. ICH E9; 3.3.2)

The outcome variable is crude SSI rate, and the study is designed to prove non-inferiority of PI to CHX in preoperative skin disinfection to prevent SSI. We define noninferiority as overall rate of SSI (cardiac and abdominal surgery not exceeding plus/minus 2.5% (i.e. weighted average crude SSI rate of 7.5 ± 2.5) based on surveillance data of Swissnoso.

6.3 Inclusion-Exclusion Criteria and General Study Population

(ICH E3; 9.3. ICH E9; 2.2.1)

Inclusion criteria:

- male and female patients

- aged ≥ 18 years

- undergoing cardiac or certain types of abdominal surgery (colorectal surgery, cholecystectomy, herniotomy, appendectomy, and bariatric surgery) which are routinely followed up by the study centers during the study period

Exclusion criteria:

- Contraindications to the use of either one of the compounds

  - CHX: intolerance to any of the compounds of the preparation, application on cornea, wounds or mucosal membranes
  - PI: Hyperthyroid disease, intolerance to any of the compounds, iodine allergy 2 weeks prior to radio-iodine treatment, dermatitis herpetiformis during, application on cornea, wounds or mucosal membranes.

- Emergency surgical intervention

- Patients refusing general consent for use of personal data

6.4 Randomisation and Blinding

(ICH E3; 9.4.3, 9.4.6. ICH E9; 2.3.1, 2.3.2)

Simple randomization at cluster level (no blinding).

6.5 Study Variables

(ICH E3; 9.5.1. ICH E9; 2.2.2)

	Baseline	30-50 day follow-up	90-365 day follow-up
SSI Abdominal Cardiac		 x x	 x
Depth of SSI (superficial, deep, organ space)		x	x
Time of follow-up (x full days from surgery date)		x	x
Pathogen (of tested)		x	x
Date of Discharge		x	x
Hospital	x		
Date of SSI		x	x
Date of surgery (time zero)	x		
Type of disinfectant (PI or CHX)			
Date of birth	x		

Sex	x		
BMI Cardiac Abdominal	 x -		
ASA score I, II, III, IV, V	x		
T-score (0 or 1, if in upper quartile of surgery duration for this type of surgery)	x		
Endoscopic (if relevant for surgery type)	(x)		
Colorectal cancer Colorectal	 x		
Diabetes mellitis	x		
Hypertension	x		
Smoker	x		
Type of surgery	x		
Duration of surgery	x		

Time of surgery	x		
Antimicrobial prophylaxis - antibiotic	x		
Antimicrobial prophylaxis - dose	x		
Antimicrobial prophylaxis – time of application	x		
Additional interventions (e.g. re-operation).	x		
Wound class Type 1, 2, 3, 4 (CDC definition)	x		
Exclusion criteria			
Side effects during application			
Haemoglobin			
Leukocytes			
Creatinine			
CRP			

Deceased		x	x
Reoperation	x		
Reason for re-operation	(x)		
Re-hospitilisation	x		
Reason for re-hospitilisation	x		
Discharge destination		x	x

(x) not relevant for all surgery types.

Exact definitions of each variable (unless defined above) are available at Swissnoso.

Data post-processing (new variables and transformations)

- Levels of categorical variables with less than 30 patients will be combined in a clinically relevant manner.
- Continuous variables will be transformed to be normally distributed using a suitable transformation.
- Antibiotic prophylaxis:
    - Continuous: negative is prior to incision.
    - New two level categorical variable: 60 mins prior to incision = 1, otherwise = 0.

Multiple imputation of missing covariate data

Baseline variables will be multiply imputed if the missingness is more than 10% in the full analysis set. Multiple imputation will be performed assuming missing at random (MAR) using the MICE packaged in R, i.e. using multiple imputation using chained equations.

Multiple imputation of missing endpoint (SSI) data

Missing SSI (yes/no) information will be multiply imputed using the MICE package in R assuming MAR if the missingness is above 10% for the full analysis set.

7 Sample Size

(ICH E3; 9.7.2. ICH E9; 3.5)

We hypothesize that neither compound is inferior to the other in terms of primary outcome SSIs. We define non-inferiority as crude rate of SSIs not exceeding plus/minus 2.5% in cardiac surgery and abdominal surgery combined (7.5% ± 2.5%) based on surveillance data of Swissnoso.

Based on a review of Swissnoso SSI rates, we estimate the crude SSI rate in cardiac surgery to be 5% and in abdominal surgery to be 10%, adding up to an overall crude SSI rate of 7.5%, assuming equal proportions of the respective surgery types. We dimensioned the study to have 80% power at a significance level of 0.05 with an absolute non-inferiority margin of 2.5%. According to the samplesize calculation ("sealed envelope" website confirmed using R (package "TrialSize", function "TwoSampleProportion.NIS")) 1,374 patients are needed in each group. Assuming a maximum of 10% dropout/lost to follow-up from the trial, this would lead to a total sample size of 1,527 in each group, and 3,054 patients overall.

The protocol was amended on 15.10.19 (Version 0.3) so that the follow-up time was reduced from 365 days to 90 to 365 days for cardiac curgery, to ensure that achievement of the total sample size is logistically feasible.

Since a shorter follow-up time for cardiac surgeries will result in fewer SSIs being detected, the original sample size calculation was no longer appropriate. It was decided that the sample size should not be amended. For cardiac surgeries, we estimate that we will lose approximately *up to* 17% of cardiac surgery SSIs (see table below).

Question 1: How many infections do we miss if we only did 90 day follow-up?					
time from iv to diagnosis of infection					
	365 days	180 days	120 days	90 days	
Number of infections	1676	1493	1430	1385	
		183	246	291	
		11%	15%	17%	

For the analysis, we assume that there is no bias from the shortening of follow-up time for the SSIs in terms of the CHX/PI groups. We will present summary statistics to compare the differences between the shortening for each group, and perform statistical tests to determine if the differences are systematic.

8 Interim analysis

9 General Considerations

9.1 Timing of Analyses

- An interim analysis was originally planned, but this was not performed due to missing data in the trial (refer to protocol amendment).
- The final analysis will be performed when all patients have been recruited according the sample size calculation, and SSI follow-up information is available for 30 day (abdominal) or 90-365 days (cardiac).

9.2 Analysis Populations

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 5.2)

This section is designed to identify the characteristics needed for inclusion in particular populations used in the analyses. Clearly define all the populations with a formal title (e.g. Full Analysis, Per Protocol, Safety) and give criteria to determine if a subject or observational unit belongs to that population. The criteria typically relate to adherence to protocol and the taking of observations, which relate to missing data (section 9.4).

Note that “intention to treat” refers to how subjects are assigned to a treatment group for the purposes of analysis (i.e. the treatment they are randomised to but not necessarily the one received); it can be used within *any* analysis population and thus is not a suitable description for a population itself.

It is not enough just to use a standard label for population. Such labels are vague and need further precise definitions within each trial; ***examples*** are given below.

9.2.1 Full Analysis Population

All subjects who were randomised to either treatment. This includes those patients with possibly missing baseline information and/or those with missing SSI information. This is the study population used to investigate the potential effects of missing covariate and endpoint data as part of a supplementary missing data analysis.

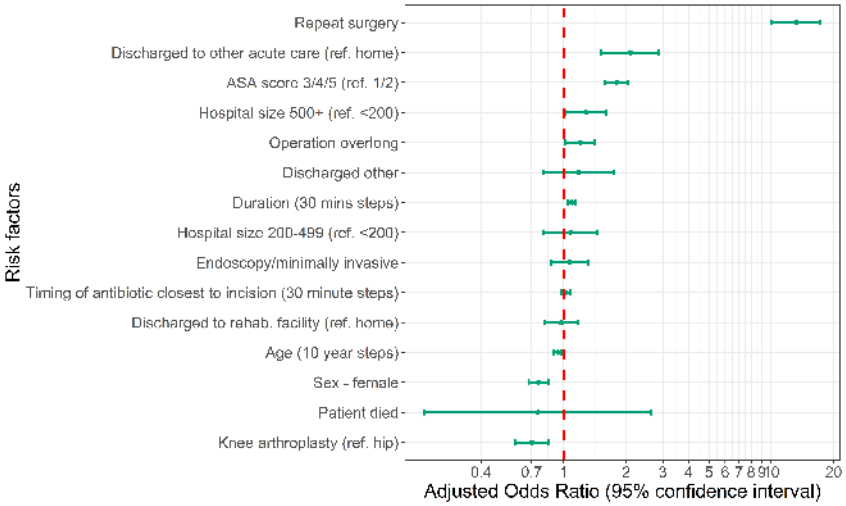
9.2.2 Per Protocol Population

All subjects who were randomised to either treatment, and have SSI information (yes/no, if yes, date). This is the study population to determine the primary endpoint.

9.3 Covariates and Subgroups

(ICH E3; 9.7.1, 11.4.2.1. ICH E9; 5.7)

Previous analyses of knee/hip surgeries have identified the following covariates as being associated with risk of SSI for the Swissnoso data: Repeat surgery, destination of discharge, ASA score, Hospital size, T-score, duration, sex.



Model selection for risk factor analysis (secondary endpoints)

Those covariates significant at the 10% level ($p \leq 0.1$) in univariable analyses will be considered in multivariable models. In multivariable models forwards then backwards selection will be used to included covariates using the Bayesian Information Criteria (BIC) as criteria.

The following variables will be considered in univariable analyses: age, sex, wound class, T-score, duration, ASA score, surgery duration, hospital, endoscopic, discharge location, re-operation, patient died, timing of antibiotic prophylaxis (continuous or 2-level categorical variable). Hospital will be considered as a random effect in models.

Of note, the study was powered for the primary endpoint only. Therefore, it is possible that the risk factor analyses may be underpowered for specific strata in univariable and multivariable analyses.

9.4 Missing Data

(ICH E3; 9.7.1, 11.4.2.2. ICH E9; 5.3. EMA Guideline on Missing Data in Confirmatory Clinical Trials)

If the missingness of any one variable (covariate or endpoint) in the full analysis set is more than 10% a missing data analysis will be performed.

Patients with missing records are identified in the data set, and compared to those without missing records to determine if there are systematic patterns of missingness in the data. To this end, a logistic regression model will be fitted with the missingness indicator as dependent variable and the relevant covariates as independent variables (i.e. those included in Table 1 of the analysis). Graphical methods will be used to visualise the missingess.

Methods for handling missing covariate and endpoint data are defined in section 7.5. A total of 50 completed data sets will be multiply imputed, and then the appropriate analysis model will be fitted to these complete data sets, with Rubin's rules used to calculate point estimates and their standard errors.

Once this process is complete the results from the complete case analysis using the per protocol analysis set will be compared to the results from fitting the appropriate analysis model to the multiply imputed data sets. If the point estimates of the covariates in the analysis model from the analysis with the multiply imputed data sets are outside of the 95% confidence intervals of the estimates from the complete case analysis with the analysis model, the it will be concluded that further sensitivity analyses will be required to investigate plausible departures from the missing at random assumption.

Such sensitivity analyses will be defined in detail *post hoc* depending on the variable, but will be based on either delta, or reference based, multiple imputation methods, whichever seems most appropriate.

9.5 Interim Analyses and Data Monitoring

(ICH E3; 9.7.1, 11.4.2.3. ICH E9; 4.1, FDA Feb 2010 “Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics”)

An interim analysis was planned to be performed after one year of study. At this point in time, if clear superiority of one of the compounds was determined, the study would be stopped, and the institutions would return to their regularly used disinfection compound.

9.6 Multi-centre Studies

(ICH E3; 9.7.1, 11.4.2.4. ICH E9; 3.2)

The study includes three centres.

Relevant covariates will be compared between the different centres to determine if there were significant differences in patient characteristics between the centres.

For the primary analysis data from all three centres will be pooled.

For multivariable analysis models of secondary endpoints, centre will be included as a random effect in appropriate mixed effects models. If such models are unstable, or fail to converge, then cluster terms will be defined in fixed effect only models, and sandwich type standard errors calculated for point estimates. These are inflated to compensate for potential correlation effects from the centres. Centre by treatment interactions will be included in such models.

9.7 Multiple Testing

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 2.2.5)

In subset analyses and other analyses in which multiple testing occurs, a simple bonferonni correction will be used.

10 Summary of Study Data

The following tables will be produced to summarise the data.

10.1 Plausibility check (Table 0)

Data set: Full analysis data set

Goals:

- Check range of values for plausibility
- Levels of missing data for each covariate.
- Define levels of categorical variables to be combined. Post hoc discussion with clinicians required.
- Identify shape of continuous variables

Patient characteristics	**Overall**
	N = XX
Primary endpoint	
SSI, n (%)	

% missing
Depth of SSI, n (%)
Superficial
Deep
Organspace
% missing
Intervention, n (%)
CHX
PI
% missing
Centre, n (%)
Basel
Bern
Zürich
% missing
Age in years, median (range)
% missing

Sex, n (%)
Male
Female
% missing
Haemoglobin, median (range)
% missing
Leukocytes, median (range)
% missing
Creatinine, median (range)
% missing
CRP, median (range)
% missing
Diabetes mellitus, n (%)
No
Yes
% missing
Smoker, n (%)

No
Yes
% missing
Hypertension, n (%)
No
Yes
% missing
Colorectal
Colorectal cancer, n (%)
No
Yes
% missing
Cardiac surgery
BMI in kg/m^2, median (range)
% missing
Surgery Type, n (%)
Cardiac

Abdominal
% missing
ASA Score, n (%)
I
II
III
IV
V
% missing
Timing of antibiotic prophylaxis, median (range)
% missing
Endoscopic, n (%)
No
Yes
% missing
ScoreT, n (%)
0

1
% missing
Wound class
1
2
3
4
% missing
Re-operation, n (%)
No
Yes
% missing
Re-hospitilisation, n (%)
No
Yes
% missing
Discharge destination, n (%)

Home or old people's home
Other acute care hospital
Rehabilitation clinic
Patient died
% missing

Histograms will be used to determine shape (skew) of continuous variables. Skewed variables will be transformed appropriately. No formal test of normality will be used; visual checking and summary statistics will be used. Boxplots will be used to compare distributions of continuous variables stratified by intervention group.

10.2 Patient characteristics, stratified by primary endpoint (Table 1)

Data set: Per protocol data set

Goals:

- Compare levels of the primary endpoint, continuous and categorical variables following stratification by PI / CHX

Patient characteristics	**Overall** **N = XX**	**CHX** **N = XX**	**PI** **N = XX**	**Statistical test for group differences**
Primary endpoint				chi-square

SSI, n (%)	
Depth of SSI, n (%)	proportions test
Superficial	
Deep	
Organspace	
Intervention, n (%)	chi-square
CHX	
PI	
Centre, n (%)	proportions test
Basel	
Bern	
Zürich	
Age in years, median (IQR)	Wilcoxon test
Sex, n (%)	chi-square
Male	
Female	

Haemoglobin, median (IQR)	Wilcoxon test
Leukocytes, median (IQR)	Wilcoxon test
Creatinine, median (IQR)	Wilcoxon test
CRP, median (IQR)	Wilcoxon test
Diabetes mellitus, n (%)	chi-square
No	
Yes	
Smoker, n (%)	chi-square
No	
Yes	
Hypertension, n (%)	chi-square
No	
Yes	
Colorectal	
Colorectal cancer, n (%)	chi-square
No	
Yes	

Cardiac surgery	
BMI in kg/m^2, median (range)	Wilcoxon test
Surgery Type, n (%)	chi-square
Cardiac	
Abdominal	
ASA Score, n (%)	proportions test
I	
II	
III	
IV	
V	
Timing of antibiotic prophylaxis, median (range)	Wilcoxon test
Endoscopic, n (%)	chi-square
No	
Yes	
ScoreT, n (%)	chi-square

0	
1	
Wound class	proportions test
1	
2	
3	
4	
Re-operation, n (%)	chi-square
No	
Yes	
Re-hospitilisation, n (%)	chi-square
No	
Yes	
Discharge destination, n (%)	proportions test
Home or old people's home	
Other acute care hospital	

Rehabilitation clinic				
Patient died				
% missing				

Variants of the chi-square test such as Fisher's exact test will be used if the number of patients in any one cell is under 5. The "proportions test" is the chi-square test for more than 2 categories.

10.3 Patient characteristics, stratified by centre (Supplementary Table 1)

Data set: Per protocol data set

Goals:

- Compare levels of the primary endpoint, continuous and categorical variables following stratification by center

Patient characteristics	**Basel** **N = XX**	**Bern** **N = XX**	**Zürich** **N = XX**	**Statistical test for group differences**
Primary endpoint **SSI, n (%)**				proportions test
Depth of SSI, n (%)				proportions test
Superficial				

Deep	
Organspace	
Intervention, n (%)	proportions test
CHX	
PI	
Centre, n (%)	proportions test
Basel	
Bern	
Zürich	
Age in years, median (IQR)	Kruskal-Wallis
Sex, n (%)	proportions test
Male	
Female	
Haemoglobin, median (IQR)	Kruskal-Wallis
Leukocytes, median (IQR)	Kruskal-Wallis
Creatinine, median (IQR)	Kruskal-Wallis

CRP, median (IQR)	Kruskal-Wallis
Diabetes mellitus, n (%)	proportions test
No	
Yes	
Smoker, n (%)	proportions test
No	
Yes	
Hypertension, n (%)	proportions test
No	
Yes	
Colorectal	
Colorectal cancer, n (%)	proportions test
No	
Yes	
Cardiac surgery	Kruskal-Wallis
BMI in kg/m^2, median (range)	
Surgery Type, n (%)	proportions test

Cardiac	
Abdominal	
ASA Score, n (%)	proportions test
I	
II	
III	
IV	
V	
Timing of antibiotic prophylaxis, median (range)	Kruskal-Wallis
Endoscopic, n (%)	Proportions test
No	
Yes	
ScoreT, n (%)	proportions test
0	
1	
Wound class	proportions test

1		
2		
3		
4		
Re-operation, n (%)		proportions test
No		
Yes		
Re-hospitilisation, n (%)		proportions test
No		
Yes		
Discharge destination, n (%)		proportions test
Home or old people's home		
Other acute care hospital		
Rehabilitation clinic		
Patient died		
% missing		

Variants of the chi-square test such as Fisher's exact test will be used if the number of patients in any one cell is under 5. The "proportions test" is the chi-square test for more than 2 categories.

10.4 Patient characteristics, stratified by missing data indicator variable (Supplementary Table 2)

Data set: Full analysis data set

Definition

- R = 0: patients with no missing data
- R= 1: patients with missing data
- Complete case analysis: Uses only those patients with R=0.

Goals:

- Compare levels of the primary endpoint, continuous and categorical variables following stratification by the missingness indicator variable.

Patient characteristics	**Overall** **N = XX**	**R=0** **N = XX**	**R = 1** **N = XX**	**Statistical test for group differences**
Primary endpoint **SSI, n (%)**				chi-square
Depth of SSI, n (%)				proportions test
Superficial				
Deep				

Organspace	
Intervention, n (%)	chi-square
CHX	
PI	
Centre, n (%)	proportions test
Basel	
Bern	
Zürich	
Age in years, median (IQR)	Wilcoxon test
Sex, n (%)	chi-square
Male	
Female	
Haemoglobin, median (IQR)	Wilcoxon test
Leukocytes, median (IQR)	Wilcoxon test
Creatinine, median (IQR)	Wilcoxon test
CRP, median (IQR)	Wilcoxon test

Diabetes mellitus, n (%)	chi-square
No	
Yes	
Smoker, n (%)	chi-square
No	
Yes	
Hypertension, n (%)	chi-square
No	
Yes	
Colorectal	
Colorectal cancer, n (%)	chi-square
No	
Yes	
Cardiac surgery	
BMI in kg/m^2, median (range)	Wilcoxon test
Surgery Type, n (%)	chi-square
Cardiac	

Abdominal	
ASA Score, n (%)	proportions test
I	
II	
III	
IV	
V	
Timing of antibiotic prophylaxis, median (range)	Wilcoxon test
Endoscopic, n (%)	chi-square
No	
Yes	
ScoreT, n (%)	chi-square
0	
1	
Wound class	proportions test
1	

2	
3	
4	
Re-operation, n (%)	chi-square
No	
Yes	
Re-hospitilisation, n (%)	chi-square
No	
Yes	
Discharge destination, n (%)	proportions test
Home or old people's home	
Other acute care hospital	
Rehabilitation clinic	
Patient died	
% missing	

10.5 Derived variables

- Age = (date of intervention – date of birth)/365.25
- R = 1 for those patients with one or more missing endpoint or covariate values, otherwise R = 0.
- Categorical variables will be combined posthoc depending on numbers of patients in each class.

10.6 Protocol Deviations

The following protocol deviations are considered to be reasons for excluding data patients from the per protocol analysis:

- Incomplete follow-up for the endpoint.
- Incorrect administration of the intervention according to the cluster randomisation schedule.

10.7 Demographic and Baseline Variables

The summary statistics will be produced in accordance with section 10.2.

10.8 Concurrent Illnesses and Medical Conditions

The summary statistics will be produced in accordance with section 10.

10.9 Treatment Compliance

The summary statistics will be produced in accordance with section 10.

11 Efficacy Analyses

11.1 Primary Efficacy Analysis

The per protocol data set will be used for this analysis.

A weighted SSI rate for abdominal and cardiac surgeries will be calculated for the CHX (reference) and PI (new) groups, pooled across all centres.

The null hypothesis is that there is no difference between the two rates, taking into account the non-inferiority margin. The alternative hypothesis is that PI is inferior to CHX.

A chi-square test will be used to test the difference in proportions between the two groups. A p-value of less then 0.05 will denote a significant difference between the two groups.

Inferiority: If the upper 95% confidence interval of the SSI rate for the "new" treatment PI group, expressed as a percentage, is less than the SSI rate for the reference treatment CHX group, again expressed as a percentage, plus 2.5% (absolute), then we may conclude non-inferiority, and accept the null hypothesis. Wald-type confidence intervals will be calculated.

Example:

SSI rate of CHX = 7.5%

SSI rate of PI = 8% with 95% confidence interval [7%, 9%].

The non-inferiority margin is 2.5%, which is equivalent to an SSI rate of 7.5%+2.5% = 10%. Since the 95% confidence interval of the SSI rate for PI is [7%, 9%], and 9% is less than 10%, we may accept the null hypothesis of non-inferiority.

11.2 Secondary Efficacy Analyses

The per protocol data set will be used for these analyses. We compare the primary outcome (and other categorical outcomes) between groups, and calculate relative risks with 99% confidence intervals.

NNIS adjusted SSI rate

The primary analysis defined in section 12.1 will be repeated using the NNIS adjusted SSI rates in each of the intervention strata. NNIS includes the wound class, ASA score and ScoreT.

SSI rate by hospital

The primary analysis defined in section 12.1 will be repeated using both the crude SSI rate and the NNIS adjusted SSI rate, stratified by hospital.

To determine whether the results were consistent across the three participating hospitals, a Breslow–Day test for homogeneity will be performed using the numerators and denominators from the crude SSI rate calculation per hospital.

Risk factor analysis

Univariable and multivariable mixed effects logistic regression models will be fitted with dependent variable SSI (0/1), and independent variables, type of disinfectant (CHD/PI), type of surgery (cardiac, abdominal), sex, BMI (for cardiac only), presence of implant (for cardiac surgery only), endscopic, ASA-score, presence of colorectal cancer (for colorectal surgery), presence of diabetes mellitus, arterial hypertension and smoking, hemoglobin, creatinine, leukocytes, CRP, with a random intercept for each hospital. Interactions between type of disinfectant and the above variables will be investigated.

Variables significant at the 5% level from univariable analyses will be included in multivariable analyses. Both full models (with all covariates mentioned above) and parsimonious models will be presented. Goodness of fit will be investigated visually and by using the Hosmer-Lemeshow test (or a variant thereof).

Point estimates and confidence intervals will be presented in table form, an example of which is shown below.

Variable	**Level**	**Univariable OR**	**P-value**	**Multivariable OR**	**P-value**
Type of disinfectant					
CHX		1 (reference)			
PI					
Type of surgery	Cardiac	1 (reference)		1 (reference)	
	Abdominal				
Sex	Female	1 (reference)		1 (reference)	
	Male				

BMI	Per 5-kg/m^2 increase		
Presence of implant	No	1 (reference)	
	Yes		
Endoscopic	No	1 (reference)	
	Yes		
ASA Score	I	1 (reference)	
	II		
	III		
	IV		
	V		
Colorectal cancer	No	1 (reference)	
	Yes		
Diabetes mellitus[d]	No	1	1
	Yes		
Hypertension	No	1 (reference)	
	Yes		

Smoking	No	1 (reference)
	Yes	
Hemoglobin		
Creatinine		
Leukocytes		
CRP		

Supplementary analysis

An interrupted time series analysis will be performed to investigate the infection rate trajectory over time within each period, with a random intercept for cluster and period, and random slope for time within period. Time 0 will be considered as the first day of cluster randomisation for the specific period.

Subgroup analyses

The above risk factor analysis will be repeated for each of the following subgroups of infections (superficial, deep, organ-space), types of surgery (cardiac, abdominal), timing of antimicrobial prophylaxis (within 60 minutes of incision, not within 60 minutes of incision), ScoreT (0/1), and wound contamination class (types 1, 2, 3, 4).

The consistency of the effects of the study intervention on infections across different follow up periods (i.e. discharge, 30 day, 90 day, 1 year (for cardiac surgery only) will be examined using an interaction term with time in the model.

11.3 Secondary endpoints

The per protocol data set will be used for this analysis.

The following additional endpoints will also be investigated.

- Mortality
    - In-hospital
    - End of follow-up: 30 day for abdominal surgery; 30 day and 365 day for cardiac surgery)
- Length of stay

A subset of the models fitted in sections 12.2 and 12.3 (where appropriate) will be used for the analysis.

11.4 Missing data study

The per protocol and full analysis set are used for these analyses.

This missing data study repeats the analyses in section 12.1 to 12.3 with multiply imputed data sets, and compares the corresponding results with those with the complete cases. Refer to sections 10.4 and 11.4 for more information.

11.5 Adverse Events

Please refer to the protocol

11.6 Deaths, Serious Adverse Events and other Significant Adverse Events

Please refer to the protocol.

12 Figures

Histograms will be used to investigate distributions of continuous variables. Boxplot will be used to compare distributions of continuous variables, stratified by the endpoint. Forest plots will be used to present odds ratios.

13 Reporting Conventions

P-values will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001".

14 Technical Details

R version 3.6.1 (or later) will be used to perform the analysis.

15 Summary of Changes to the Protocol

The analysis presented here extends and provides more detail regarding the analysis proposed in Version 3.0 of the protocol.

The following changes have been made:

- Relative risks has been replaced with Odds Ratios: This is the standard output from logistic regression models
- 99% confidence intervals has been replaced with 95% confidence intervals: 99% confidence intervals do not allow simple comparisons with similar studies.
- "For continuous factors, we will use a single-variable logistic-regression model that involves generalized estimating equations (GEE) and ….": We fit mixed effects models, GEEs will not be required unless mixed effects models do not converge.

16 References

None