Statistical analysis plan for clinical outcomes in the AIM-study

Statistical analysis plan (SAP) for clinical outcomes in:

A randomized trial of antibiotic treatment in patients with chronic low back pain and Modic Changes (the AIM-study)

Publication date: 02.11.2018

Document version: 1.0

EudraCT Number: 2013-004505-14

This document is a supplement to the AIM-(Antibiotic In Modic changes-) study protocol[1] and comprises a statistical analysis plan for the article "A randomized trial of antibiotic treatment in patients with chronic low back pain and Modic Changes (the AIM-study)". This statistical analysis plan is prepared in accordance with guidelines for Statistical analysis plans in clinical trials[2].

Scientific board:

John-Anker Zwart, professor dr.med
Kjersti Storheim, professor, PhD
Ansgar Espeland, professor dr.med
Jens Ivar Brox, professor dr.med
Olav Lutro, MD

Christian Hellum, MD, PhD
Margreth Grotle, professor, PhD
Benedicte A Lie, professor, PhD
Anne Froholdt, MD, PhD
Lars Grøvle /Anne Julsrud Haugen, MD, PhD
Jan Sture Skouen, professor, dr.med
Audny Anke, professor, dr.med
Øystein P Nygaard, professor, dr.med
Karianne W Gammelsrud, MD, PhD
Thor Einar Holmgaard (patient representative)

Statistical advisor:

Jörg Aßmus, PhD

Sponsor:

Oslo University Hospital

Coordinating Investigator:

John-Anker Zwart, professor dr.med
FORMI, OUS Ullevål, Bygg 37B, Postboks 4956 Nydalen, 0424 Oslo
Tel : 22 11 86 21
E-mail: j.a.zwart@medisin.uio.no

Contributors to current SAP:

Lars Christian Bråten, MD, PhD-student (writing the SAP)
Mads Rolfsen, MD, PhD-student
Ansgar Espeland, professor, dr.med
Kjersti Storheim, professor, PhD
John-Anker Zwart, professor, dr.med
Christian Hellum, MD, PhD
Jörg Aßmus, PhD

Signature page

Title	**A randomized trial of antibiotic treatment in patients with chronic low back pain and Modic Changes (the AIM-study)**
Trial registration:	ClinicalTrials.gov, ID: NCT02323412
EudraCT no:	2013-004505-14

I hereby declare that I have reviewed and approved the Statistical Analysis Plan:
To be signed by *persons writing the SAP*, Senior statistician responsible, Project managing and Coordinating Investigator.

Name	Title	Role	Signature	Date
Lars Christian H Bråten	MD, PhD-student	Main author of the SAP		26/10-18
Jörg Aßmus	PhD	Senior statistician responsible		29/10-18
Ansgar Espeland	Professor, dr.med	Contributor to current SAP		29/10-18
Mads P Rolfsen	MD, PhD-student	Contributor to current SAP		26.10.18
Christian Hellum	MD, PhD	Contributor to current SAP		26/10-18
Kjersti Storheim	Professor, PhD	Project manager		25/10-18
John-Anker Zwart	Professor, dr.med	Coordinating Investigator		24/10-18

Table of contents

List of Abbreviations and definitions

Abbreviation or special term	Explanation
AE	Adverse Event
AIM	Antibiotics In Modic changes
CI	Confidence interval
CTCAE	Common Terminology Criteria for Adverse Event
ITT	Intention to treat
LBP	Low Back Pain
LME	Linear mixed-effects
MCs	Modic changes
MedDRA	Medical Dictionary for Regulatory Activities
NRS	Numerical Rating Scale
ODI	Oswestry Disability Index
RMDQ	Roland Morris Disability Questionnaire
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
Safety population	All patients who took at least one tablet of the study medication
Study medication	Medication given in the study context containing either amoxicillin (test treatment) or placebo (the comparator)

Study objectives and outcomes

Main objective

- To evaluate the effect of Amoxicillin versus placebo on disease-specific disability evaluated by the Roland Morris Disability Questionnaire (RMDQ) at one year (12 months) follow-up in patients with chronic LBP and MCs type I or II at the level of a previously herniated disc (Hypothesis A).

Secondary objective (SO 1)

- To evaluate the effect of amoxicillin versus placebo on RMDQ at 1-year follow-up separately in patients with type I and type II MCs, respectively (Hypothesis B and C).

Key clinical supportive objectives (KSOs) and exploratory objectives

- To evaluate the effect of amoxicillin versus placebo on Oswestry Disability Index (ODI) at 1-year follow-up in the whole cohort of included patients (KSO 2, Hypothesis D)
- To evaluate the effect of amoxicillin versus placebo on LBP intensity at 1-year follow-up in the whole cohort of included patients (KSO 3, Hypothesis E)
- To evaluate the effect of amoxicillin versus placebo on health-related quality of life (the EQ-5D) at 1-year follow-up in the whole cohort of included patients (KSO 6, Hypothesis H)
- To evaluate the difference in incidence of AEs and SAEs between the two intervention groups from inclusion to 1-year follow-up in the safety population. The decision to evaluate this difference in the safety population rather than in the whole cohort of included patients was made after the publication of the study protocol[1] (but prior to analysis).

Trial methods

The trial is a six center, randomised, parallel-group, placebo-controlled trial. Treatment allocation is stratified on MC type group (I/II) and previous disc surgery with a 1:1:1:1 allocation and random block sizes of 4 and 6. Patients are randomised to either amoxicillin or placebo control.

The sample size was calculated to assess the treatment effect separately in each MC type group (I/II). In each MC type group, the study is designed to detect ($\beta = 0.1$, two-sided $\alpha = 0.05$) a mean difference of 4 (SD 5) in the RMDQ score between the two treatment groups (amoxicillin or placebo) at one-year follow-up. See reference for further details of trial methods [1].

All outcomes are analysed for superiority of one intervention compared to the other.

One pre-specified interim analysis was performed by a statistician in an independent Data Monitoring Committee (DMC), blinded to treatment arm, using the primary outcome measure (RMDQ) at one-year follow-up in the first 80 included patients. The intention of this interim analysis was safety and secure that continuation of trial, giving patients placebo or antibiotics, was ethical. The DMC was instructed to consider stopping the trial only if they detected a mean difference of >7.0 in one-year RMDQ score between the two treatment arms, adjusted for baseline RMDQ score. The trial was not stopped after the interim analysis.

All analyses described in this document will take place after database locking, which will occur after all patients have finished their last visit and monitoring has been completed in all study centers (anticipated November 2018).

Statistical principles

All analyses described in this plan are considered *a priori* analyses in that they have been defined in the protocol and/or this SAP. All *post hoc* analyses will be identified as such in the article.

All relevant statistical tests will be 2-sided and the nominal p value will be reported. All confidence intervals presented will be 95% and 2-sided. The assumption of normal distribution will be checked by visual inspection of a QQ-plot. For skewed data the interquartile range will be reported.

All primary and secondary analyses (Hypothesis A, B and C) will be carried out independently by a senior statistician using software package R version 3.5.0, and Ph.D-student using software package Stata version 15, both blinded to treatment group.

Hypotheses

This study is designed to assess the superiority of 3 months of oral Amoxicillin 750 mg three times daily compared to 3 months of oral placebo tablets three times daily with regard to the the RMDQ score at one-year follow-up in patients with chronic LBP and MCs type I or II at the level of a previously herniated disc.

For hypothesis A, the null hypothesis is that patients with MCs type I or II at baseline in the antibiotic group report the same RMDQ score at 1-year as patients in the placebo group.

The alternative hypothesis is that patients with MCs type I or II at baseline in the antibiotic group report a significantly lower RMDQ score at 1-year follow-up than patients in the placebo group.

Statistical significance is claimed if the null hypothesis is rejected on the significance level (alpha) of 0.05 (two-sided). That is, if the p-value of the null hypothesis test is less than or equal to 0.05. We will not adjust the significance level due to the interim analysis (conducted for safety and ethical reasons) as there was not a group-sequential design of the trial. As described in the protocol article[1], we use the same significance level for hypothesis B and C. For hypotheses D, E and H, the significance level will be adjusted to 0.0167 to avoid type 1 errors due to multiple testing.

Statistically significant differences of < 4 RMDQ points are however not clinically relevant and will not be used as a basis for recommending antibiotic treatment in the studied patient groups.

Analysis populations

In the following definitions of terms, the study medication refers to the medication given as part of the study, and includes both amoxicillin and placebo. We used the following definitions of analysis populations:

- Intention-to-treat population is defined as all patients randomized to the given treatment.
- Per-protocol population is defined as all patients who completed the trial without any major protocol deviations.
- Safety population is defined as all patients who took at least one tablet of the study medication.

We will decide which analysis populations each patient belongs to in advance of database locking.

Trial population

The following summaries will be presented in a flow diagram:

The number of days recruiting, the number of patients screened, the number of patients included and randomised, the number of screened patients not included, and the reason for non-inclusion. The number of, if any, ineligible patients who were randomised will be reported, with reasons for ineligibility. The flow diagram will also show separately lost to follow-up, withdrawal from follow-up and discontinuation of the intervention, all reported for each treatment arm and with timing and respective reasons.

Baseline patient characteristics

Patients will be described with respect to age, gender, BMI, smoking, educational level, comorbidity (Functional Comorbidity Index[3]), presence of leg pain, NRS-leg pain (0-10), subjective health complaints, emotional distress (Hopkins Symptom Check List-25), Fear-avoidance beliefs questionnaire (FABQ), symptom-specific well-being, duration of back pain, physically heavy work, compensated work injury or sick leave, level(s) with both Modic Change and previously herniated disc, and concomitant medication use, separately for the two treatment groups (see Table 1). Continuous variables will be summarized by mean and SD in case of normal distribution and median and interquartile range (difference between 75^{th} and 25^{th} quantiles) in case of skewed distribution. Categorical variables will be summarized by numbers and percentages. We will not perform any test of statistical significance, but rather note the potential clinical importance for outcome of any baseline imbalance between the treatment groups.

Compliance and Protocol Deviations

Compliance is defined as the percent of pills that the patient has taken out of the planned number of pills:

% compliance = (number of pills taken / 300)*100%. The number and % of patients taking less than 80% of the prescribed treatment will be presented in table 2. Noncompliance was defined after the publication of the study protocol (but prior to analysis) as taking less than 80%[4] of the prescribed pills.

Compliance less than 95% but more than 80% is defined as minor protocol deviation.

Compliance less than 80% is defined as major protocol deviation.

Pause of study medication (both treatment groups) or intake of antibiotics (in the placebo group) will also be considered as minor or major protocol deviation, as reported in Table 5 – Protocol deviations.

Any back surgery performed at any time point between baseline and 1-year follow up is defined as major protocol deviation. Any other treatment for back pain, which was initiated at any time point between baseline and 1-year follow up, is defined as minor protocol deviation.

The number and % of patients with one or more protocol deviations will be reported separately for minor and major protocol deviations in table 2. No formal statistical testing will be performed on differences between the treatment groups in any calculations on compliance or protocol deviations.

Analysis

Primary analysis

(See Main objective) In the primary analysis we will assess the effect of amoxicillin versus placebo on RMDQ at 1-year follow-up in the ITT population using ANCOVA with RMDQ score at 1 year as dependent variable adjusted for baseline RMDQ score and the stratification variables in the randomization (modic study group and former surgery for disc herniation). This analysis deviates from the initial protocol, as we find this analysis more correct and in accordance with guidelines[5]. Missing RMDQ values will be imputed. We will first perform a mean imputation of RMDQ at any timepoint in those cases where less than 30% of the questions are missing and secondly a multiple imputation in those cases where more than 30% of the questions are missing at any timepoint. The multiple imputation model will use predictive mean matching. We will impute both RMDQ and LBP intensity at baseline, 3 months, 6 months, 9 months and 12 months, ODI and index EQ5D-5L score at baseline, 3 months and 12 months, using age, leg pain (NRS 0-10), comorbidity (Functional Comorbidity Index), fear avoidance, emotional distress, physical work load, former surgery for disc herniation, study center, Modic type group and treatment group as predictors in the imputation model. The imputation model can be reduced if the imputation does not converge. The mean imputation implies a proportional recalculation of the item. If for instance a patient has scored 12 at RMDQ at baseline, but only answered 18 (out of 24) questions, the mean imputation will impute the value 16 =(12/18) *24. Fifty imputations will be performed[6] and the imputation model will include the same patients who are included in the analysis. The imputation will only be performed once and then used for all further analysis on the RMDQ-score.

For sample size calculations, we considered a between-group difference of 4 points on the RMDQ as the smallest clinically relevant difference. However, since our study is designed to detect the smallest clinically relevant difference separately in each Modic type group, it may detect a smaller and clinically irrelevant difference in the total sample. Hence, statistically significant differences of < 4 RMDQ points will not be used as a basis for recommending antibiotic treatment in the studied patient groups.

Secondary analysis

Modic types

(See Secondary objective (SO 1)) We will analyze the effect of amoxicillin versus placebo on RMDQ at 1-year follow-up in the ITT population separately in patients with type I and patients with type II MCs, using ANCOVA with RMDQ at 1 year as dependent variable adjusted for baseline value of

RMDQ and the stratification variable (except modic type). We will impute missing values of RMDQ using the same imputed values as for the primary analysis (main objective).

Analysis of Key supportive objectives

Oswestry Disability Index

(See Key clinical supportive objectives (KSOs) and exploratory objectives) The effect of amoxicillin versus placebo on ODI at 1-year follow-up will be analyzed in the ITT population, using ANCOVA with ODI at 1 year as dependent variable adjusted for baseline value of ODI. ODI has an inbuilt mechanism of handling missing values as the ODI score is calculated using the proportion of items answered. However, if more than 30% of items are not answered, the ODI score will be viewed as missing. It will then be imputed using the same imputed data as for the primary analysis. The imputation will only be performed once and then used for all further analysis on the ODI-score.

LBP intensity

(See Key clinical supportive objectives (KSOs) and exploratory objectives) The effect of amoxicillin versus placebo on LBP intensity at 1-year follow-up will be analyzed in the ITT population, using ANCOVA with LBP intensity at 1 year as dependent variable adjusted for baseline value of LBP intensity. LBP intensity is defined as a mean of three 0–10 NRS scores; current LBP, the worst LBP within the last 2 weeks, and the usual/mean LBP within the last 2 weeks. Missing values of LBP intensity will be imputed using the same imputed data as for the primary analysis. The imputation will only be performed once and then used for all further analysis on the LBP intensity.

Health-related quality of life

The effect of amoxicillin versus placebo on health-related quality of life at 1-year follow-up will be evaluated in the ITT population, using ANCOVA with EQ5D-5L at 1 year as dependent variable adjusted for baseline value of EQ5D-5L. In this analysis, we will use index values for EQ5D-5L, which are calculated using a crosswalk value set from the UK population. Missing values of EQ5D-5L will be imputed using the same imputed data as for the primary analysis. We intend to impute index values, but if the pattern of missingness follows an item non-response, we will consider imputing on item score level [7]. The imputation will only be performed once and then used for all further analysis on the EQ5D index values.

Analysis of exploratory objectives

Adverse Events

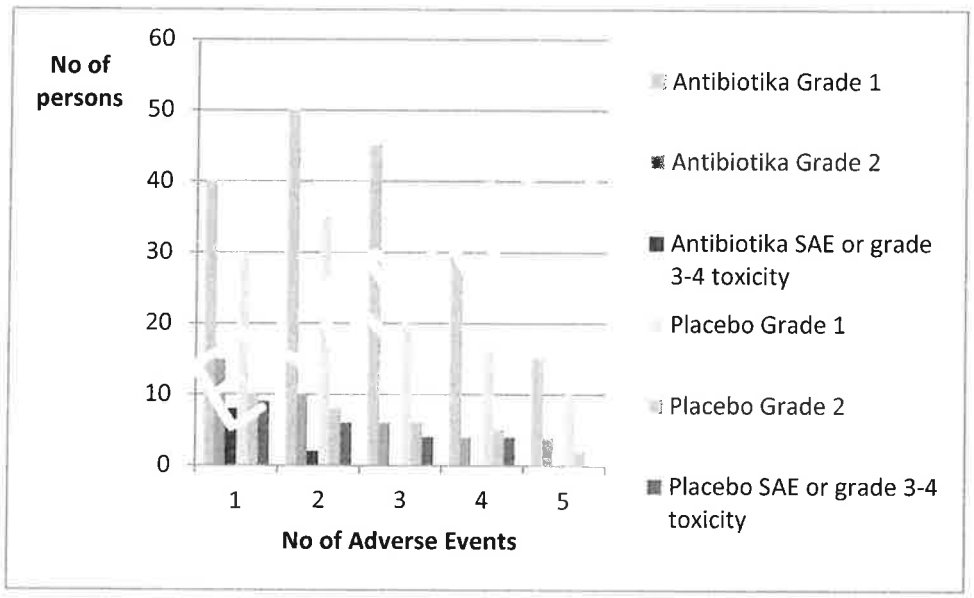
(See Key clinical supportive (KSOs) and exploratory objectives) We will report the incidence of adverse events (AEs) and serious adverse events (SAEs) from inclusion to 1-year follow-up in the two intervention groups in the safety population, according to Consort guidelines on reporting of harms in randomized trials [8]. An SAE is defined as: "Any untoward medical occurrence that at any dose results in; death, is immediately life-threatening, requires in-patient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability or incapacity, is a congenital abnormality or birth defect, or is an important medical event that may jeopardize the subject or may require medical intervention to prevent one of the outcomes listed above". Recurrences of the same event in the same individual will be reported as only one event. Adverse events are recorded from

baseline to 1-year follow-up and will be coded at all follow-up times (0,1,2,3 and 12 months) using the Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 in accordance with medDRA-coding.

We will calculate, separately for each treatment group, the number and percentage of patients with

1) one or more AEs, 2) one or more grade 2 AEs, 3) SAE or grade 3-4 toxicity and 4) symptoms of diarrhea, abdominal pain, rash, candida infection (see Table 3 - Adverse Events). For each of these four categories of events/symptoms, we will also report the number with possible/probable/definite (versus unrelated/unlikely) relationship to treatment. For all these calculations we will only present events/symptoms descriptively without any formal statistical testing.

We will also display in a histogram, for each treatment group, the number of AEs per person categorized by severity grade. Example of a histogram:



Further analyses

Responder analyses

We will compare proportions of patients with > 75%, > 50% and > 30% reduction in RMDQ score from baseline to 1-year follow-up between treatment groups using chi-square tests and reporting number needed to treat (NNT) with 95% CI. We will also present Cumulative distribution function of absolute change in RMDQ from baseline to 1-year with 95% CI for each treatment group, as recommended by NIH Task Force on Research Standards for Chronic Low Back Pain[9] (See this reference for details [10]).

Linear mixed-effects (LME) models

We will estimate the following linear mixed-effects models:

We will use linear mixed-effects models for the ITT population with RMDQ as dependent variable, an interaction term between time and intervention group, treating time as a categorical variable. We will analyse LME models for the ITT population with LBP intensity as a dependent variable (with a total of 17 time points), with an interaction term between time and intervention group, treating time as a categorical variable. All LME model specifications will be selected using the Aikaike's Information Criterion.

We will consider doing, but not necessarily report, a similar LME analysis of the other key supportive objectives (ODI and EQ5D) in a similar manner as the LME models of RMDQ.

Further analysis to be considered, but not necessarily reported
We will consider doing, but not necessarily report, responder analysis of repeated measures and analysis of time to effect with percentiles:

Responder analysis of repeated measures
Based on the LME models described above, we will consider doing, but not necessarily report, responder analysis of repeated measures of both RMDQ (at baseline, 3,6,9 and 12 months) and LBP intensity (at baseline, weekly during the treatment period, and at 3, 6, 9 and 12 months) with three different definitions of treatment responders: >30% improvement of outcome at each time point compared to baseline; >50% improvement of outcome at each time point compared to baseline; >75% improvement of outcome at each time point compared to baseline.

We will use a generalized linear mixed models (GLMM) with treatment responder as dependent variable, an interaction term between time and intervention group, treating time as a categorical variable adjusted for baseline values and with a random intercept.

Time to effect with percentiles
Based on the LME models described above, we will consider doing, but not necessarily report, analyses to estimate time to effect with CI. We will use the same definition of treatment responder as in the responder analysis of repeated measures above, for RMDQ and LBP intensity. We will define a 'Time to treatment response' (T_{30}, T_{50} and T_{75}) for each category of treatment response as the first time (measured in days) a subject is a treatment responder and the subject stays a treatment responder for the rest of the follow-up time in the study (i.e. there are no point in time after the 'time to treatment response' where the subject is not a treatment responder). We will calculate for each treatment group the mean values for T_{30}, T_{50} and T_{75} with 25^{th} and 75^{th} percentiles, and compare each value between the treatment groups with a Mann-Whitney U test.

Blinding:
To assess patients' blinding to treatment allocation, patients are asked post treatment (100 days after start of treatment) and at 1-year follow-up to report which study medicine they think that they received (antibiotics/placebo/unsure). If they respond 'unsure', the patients are forced to reply whichever of antibiotics or placebo they find most likely they received. We will calculate the number

of patients reporting 'antibiotics', 'placebo' or 'unsure' for each treatment group (see Table 4 – Response to blinding question) and Bang's blinding index[11].

Sensitivity analyses

The following analyses will be performed to assess the impact of key assumptions or variations of the primary outcome.

Problem/scenario	Analysis	Motivation
Protocol violation	Per-protocol analysis Analysis of the effect of the intervention on RMDQ at 1 year using ANCOVA adjusted for baseline RMDQ in the per protocol population	A lack of effect of the intervention on the primary outcome could be due to too many participants in the amoxicillin group who did not complete treatment, or that patients in the placebo group had co-interventions outside the study protocol (e.g. exercise, multidisciplinary rehabilitation, antibiotics, spinal manipulation, etc).
Baseline differences in variables strongly correlated to outcome	Analysis of the effect of the intervention on RMDQ at 1 year for the ITT population using ANCOVA adjusted for baseline RMDQ and variables with clinically relevant baseline differences (known predictors for LBP [12,13]): • Age • LBP intensity • NRS leg pain (0-10) • Comorbidity • Fear-avoidance • Emotional distress • Physical work load • Compensated work injury or sick leave	Baseline differences in key variables that are known to predict outcome could occur by chance despite the randomized setup, and could possibly impact on effect estimates.
Covariate due to multicenter trial	Analysis of the effect of the intervention on RMDQ at 1 year for the ITT population using ANCOVA adjusted for baseline RMDQ and the stratification variables in the randomization (modic study group and former surgery for disc herniation), and study center.	There might be aspects in treatment given to participants that are particular for one or more study center, which could influence the outcome.

Analyses in case of effect of treatment on the primary outcome and evidence of unsuccessful blinding

In case of suspected unsuccessful blinding and effect of treatment on the primary outcome, defined as a statistically significant difference of more than 4 points of RMDQ between the estimates in the primary analysis, we consider no statistical test able to estimate what the treatment effect would have been if there had been more successful blinding. Furthermore, the perceived treatment allocation could be caused by both the treatment group and the outcome, and hence be a collider on the causal pathway between the treatment and the outcome. Hence, adding the patients' responses for perceived treatment allocation as a covariate in the analysis could induce bias. For this reason, no such analysis will be performed. In case of evidence of treatment effect in the primary analysis and evidence of possible unsuccessful blinding, we will discuss how side effects and treatment response may have influenced the patients' responses about perceived treatment group, and how the results of the study may be interpreted. A logistic regression analysis may be performed using patients' responses about treatment group at 1-year (placebo/uncertain vs antibiotic) as a dependent variable, and the following independent variables:

1. Treatment group (dichotomous variable)
2. Number of adverse reactions which are attributed to a possible, probable or definite relation to treatment effect
3. Improved symptoms, defined as more than 30% drop in RMDQ from baseline to 3 months follow-up and from baseline to 12 months follow-up.

We will also consider doing, but not necessarily present, probability density graphs of RMDQ-score at 1-year follow up across patients in the three patient responses about treatment group and allocated treatment sub-groups (see example in figure 3 in this reference[14]).

Tables

Table 1 – Baseline Characteristics

	Amoxicillin (n =)	Placebo (n =)
Age		
Gender		
BMI		
Smoking- no. (%)		
Educational level		
Comorbidity		
Presence of leg pain		
NRS-leg pain, 0-10, mean (SD)		
Subjective health complaints ├		
Emotional distress, 1-4, mean (SD) •		
FABQ physical activity, 0-24, mean (SD) ►		
FABQ work, 0-42, mean (SD) ►		
Symptom specific well-being, 1-5, mean (SD) ¶		
Duration of back pain		
Physically heavy work (%)		
Compensated work injury or sick leave (%)		
Level of Modic Change and previous disc herniation - no. (%)		
L1/L2		
L2/L3		
L3/L4		
L4/L5		
L5/S1		
Concomitant medication use		
Analgesics for back pain – no.		
Opioids for back pain – no.		

• Emotional distress (Hopkins Symptom Checklist–25)
► Fear-avoidance beliefs Questionnaire
¶ Symptom specific well-being (5-point Likert scale)

Table 2 - Primary and Secondary Outcomes

Variable	Amoxicillin (n =)			Placebo (n =)			Estimated differences between Intervention groups				
	No of partici-pants	Overall mean (CI 95%)	Mean change from baseline	No of partici-pants	Overall mean (CI 95%)	Mean change from baseline	No of partici-pants	Crude difference (95% CI)	P-Value	Adjusted difference* (95% CI)	P-Value
RMDQ†											
baseline								-			
3 months											
6 months											
9 months											
12 months											
ODI‡											
baseline											
3 months											
12 months											
LBP ¶ intensity											
baseline											
3 months											
6 months											
9 months											
12 months											
EQ-5D |											
baseline											
3 months											
12 months											
Days with sick leave											
baseline											
3 months											
12 months											
Global perceived effect		-			-						
Patient satisfaction		-			-						
Non-compliance		-			-						
Minor protocol deviation		-			-						
Major protocol deviation		-			-						

† Roland Morris Disability Questionnaire. Pain and disability measure, ranges from 0 to 24, with a lower score indicating less severe pain and disability.

‡ Oswestry Disability Index. Measure of functional capacity that range from 0 to 100. It is based on a scale of 0-5 on 10 topics. These scores are summed and then multiplied by two to obtain the index score.

¶ NRS back pain. A mean of three Numeric Rating Scales; current low back pain, worst low back pain within the last 2 weeks, and usual/mean low back pain within the last 2 weeks

| Health related quality of life scores (EuroQoL -5D). Measured on 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, with a score 1-5 on each dimension. These values are converted to a single summary index by applying a standardized formula that attaches weights to each of the levels in each dimension, giving a score from -0.59 to 1.0 (higher scores indicate a higher quality of life).

*Adjusted for baseline value

Table 3 - Adverse Events

Adverse Event	Amoxicillin (n =)	Placebo (n =)
≥1 adverse event – no. of patients (%)		
Unrelated/unlikely only		
Possible /probable/definite, any		
≥1 grade 2 (or higher) adverse event – no. of patients (%)		
Unrelated/unlikely only		
Possible /probable/definite, any		
≥1 serious adverse event – no. of patients (%)		
Unrelated/unlikely only		
Possible /probable/definite, any		
Reported symptoms or events – no. of patients (%)		
Abdominal pain		
Diarrhea		
Rash		
Vaginal Candida infection		
Oral Candida infection		

Any patient with both 'Unrelated/unlikely' and 'Possible/probable/definite' adverse events will be counted in the group 'Possible/probable/definite'. 'Unrelated/unlikely' and 'Possible/probable/definite' is summed up in their respective title lines above.

Table 4 – Response to blinding question

	Amoxicillin (n=)	Placebo (n=)
Antibiotics		
Placebo		
Unsure		

Table 5 – Protocol deviations due to pause of study medication or other antibiotic treatment

	Antibiotic group	Placebo group
A pause of study medication for more than 2 continuous days but less than 2 weeks WITHOUT other relevant* antibiotic treatment in that period	Minor protocol deviation	
A pause of study medication for 2 continuous weeks or more WITHOUT other relevant* antibiotic treatment in that period	Major protocol deviation	
Relevant* antibiotic treatment for less than 4 continuous weeks (at any time from baseline to 1-year)		Minor protocol deviation
Relevant* antibiotic treatment for 4 continuous weeks or more (at any time from baseline to 1-year)		Major protocol deviation

*Antibiotic treatment with pharmacokinetic and pharmacodynamics properties considered to make the treatment likely to affect a P.acne discitis

Literature:

1. Storheim K, Espeland A, Grøvle L, et al. Antibiotic treatment In patients with chronic low back pain and Modic changes (the AIM study): study protocol for a randomised controlled trial. 2017;18.
2. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA 2017;318:2337-43.
3. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. Journal of Clinical Epidemiology 2005;58:595-602.
4. WR KPB. Compliance in anti-infective medicine. Adv Stud Med 2006;6:652-8.
5. Group IEW. ICH Harmonised tripartite guideline – Statistical principles for clinical trials E9. International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH); 1998.
6. Buuren Sv. Flexible imputation of missing data. Boca Raton, Fla: Taylor and Francis; 2012 page 38.
7. Simons CL, Rivero-Arias O, Yu LM, Simon J. Multiple imputation to deal with missing EQ-5D-3L data: Should we impute individual domains or the actual index? Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2015;24:805-15.
8. Ioannidis JPA, Evans SJW, Gøtzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Annals of internal medicine 2004;141:781.
9. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on Research Standards for Chronic Low Back Pain. Journal of Pain 2014;15:569-85.
10. McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. Expert Review of Pharmacoeconomics & Outcomes Research 2011;11:163-9.
11. Bang H, Ni L, Davis CE. Assessment of blinding in clinical trials. Controlled Clinical Trials 2004;25:143-56.
12. Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results—guidance for future prognosis reviews. Journal of Clinical Epidemiology 2009;62:781-96.e1.
13. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. Lancet (London, England) 2018;391:2356-67.
14. Arandjelovic O. A new framework for interpreting the outcomes of imperfectly blinded controlled clinical trials. PLoS One 2012;7:e48984.