Statistical analysis plan for analysis of cytokines as prognostic factors and predictors of antibiotic treatment effect in the AIM trial

Statistical analysis plan for:

Cytokines as prognostic factors and predictors of antibiotic treatment effect in patients with chronic low back pain and Modic changes (the AIM trial)

Publication date: 21.SEPT 2021

Document version: 1.0

EudraCT Number: 2013-004505-14

This document is a supplement to the AIM trial protocol (1) and comprises a statistical analysis plan (SAP) for two additional objectives regarding cytokines as prognostic factors and treatment effect predictors of cytokines. The results of these analyses will be presented in one or two articles. Separate SAPs exist for other articles based on the AIM trial (2). The current SAP was prepared based on the Prognosis research strategy (PROGRESS) framework (3) and guidelines for Statistical analysis plans in clinical trials (4), and after results for treatment efficacy in the main trial (5) and in clinical and radiological subgroups (6, 7) were available.

Statistical Analysis Plan for analysis of cytokines as prognostic factors and predictors of antibiotic treatment effect in patients with chronic low back pain and Modic changes (the AIM trial) - - version 1.0

Scientific board

John-Anker Zwart, professor dr.med
Kjersti Storheim, professor, PT, PhD
Jens Ivar Brox, professor dr.med
Ansgar Espeland, professor dr.med
Margreth Grotle, professor, PT, PhD
Olav Lutro, MD
Anne Froholdt, MD, PhD
Lars Grøvle /Anne Julsrud Haugen, MD, PhD
Audny Anke, professor dr.med
Øystein P Nygaard, professor dr.med
Jan Sture Skouen, professor dr.med
Christian Hellum, MD, PhD
Benedicte A Lie, PhD
Karianne W Gammelsrud, MD, PhD
Thor Einar Holmgaard (patient representative)

Statistical advisor

Are Hugo Pripp, 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.swarz@medisin.uio.no

Contributors to current SAP

Lars Christian Bråten, MD, PhD
Elisabeth Gjefsen, MD
Kristina Gervin, PhD
Are Hugo Pripp, PhD, statistician
Jan Sture Skouen, prof
Elina Schistad, PhD
Linda Margareth Pedersen
Monica Wigemyr
Kaja Kristine Selmer
Hans Christian Aass
Guro Løvik Goll
Jens Ivar Brox, professor dr.med
Ansgar Espeland, professor dr.med

Statistical Analysis Plan for analysis of cytokines as prognostic factors and predictors of antibiotic treatment effect in patients with chronic low back pain and Modic changes (the AIM trial) - - version 1.0

Lars Grøvle, MD, PhD
John-Anker Zwart, professor dr.med
Kjersti Storheim, professor, PT

Signature page

Title	Cytokines as prognostic factors and predictors of antibiotic treatment effect in patients with chronic low back pain and Modic changes (the AIM trial)
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 person writing the SAP, project manager, and coordinating investigator.

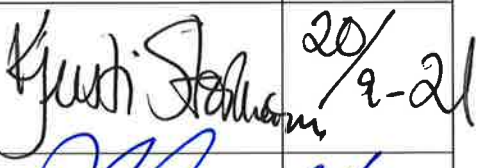
Name	Title	Role	Signature	Date
Lars Christian Bråten	PhD, dr.med	Main author of current SAP		20/9-21
Kjersti Storheim	Professor, PT	Project manager		20/9-21
John-Anker Zwart	Professor, dr.med	Coordinating investigator		29/9-21

Table of contents

Abbreviations and definitions

Abbreviation or term	Explanation
AIM	Antibiotics in Modic changes
CI	Confidence interval
ITT	Intention to treat
MCs	Modic changes
MRI	Magnetic resonance imaging
NRS	Numerical rating scale
ODI	Oswestry disability Index
RMDQ	Roland Morris disability questionnaire
SAP	Statistical analysis plan
STIR	Short tau inversion recovery
CACE	Complier average causal effect
C.acnes	Cutibacterium acnes
SD	Standard deviation
TNF-α	Tumour necrosis factor alpha
IL	Interleukin
IFN-γ	Interferon gamma
CCL	Chemokine (C-C motif) ligand
MIP-1α	Macrophage inflammatory protein 1-alpha
MCP-1	Monocyte chemoattractant protein 1

Background and rationale

A subgroup of patients with chronic low back pain has vertebral bone marrow changes extending from the endplates visible on MRI (also called Modic changes (MCs)). They are classified as type 1 (oedema type), 2 (fatty type) and 3 (sclerotic type, which is less common) based on standard T1- and T2 weighted sequences (8).

The etiology of MCs is uncertain. One theory of the underlying etiology implicates an infection with *C.acnes* in the intervertebral disc (9-11). Several studies have assessed the presence of these bacteria in intervertebral discs adjacent to MCs, with varying results (9). Contamination or colonisation might explain the findings of *C.acnes*, and some argue that microbiologically inadequate techniques might explain why C.acnes is not found in other studies (12). In studies attempting to reduce the risk of contamination, *C.acnes* has been found in a smaller subset of patients(13). Hence,

an effect of antibiotic treatment, if any, would only be expected to occur in a subset of patients with chronic low back pain and MCs.

Other possible etiologies of MCs include biomechanical stress, degeneration and auto-immune/auto-inflammatory mechanisms (13, 14). Evidence for these theories is mainly based on basic science and epidemiological evidence, and there is currently no available clinical test or investigation to establish any cause of MCs in individual patients.

The effect of antibiotic treatment in patients with chronic low back pain and MCs has been tested in two separate RCTs (15, 16). The 2013 trial reported a large treatment effect, while the AIM trial reported no clinically relevant (but statistically significant) differences in clinical outcome between the treatment groups. Since amoxicillin (little or no anti-inflammatory properties) would only be efficacious in patients with an infectious etiology, a possible explanation for the divergent results between the two trials is that the two populations comprised different proportions of subjects with a disc/endplate infection.

Although the overall results of the AIM trial were negative, some aspects might nevertheless suggest the possibility of effect in subgroups. The evidence for different etiologies of MCs in different studies (and patients) suggests an infection might at best only be occurring in subgroups. In the AIM trial, there was a statistically significant effect, and the variability, as expressed by the SD of the primary outcome, was higher in the amoxicillin group than in the placebo group (unpublished results)(16). Ideally, identification of individuals who could benefit from amoxicillin treatment would be by documenting evidence of *C.acnes* infection of the intervertebral disc or endplate. However, identifying patients with local *C.acnes* infections is known to be challenging for several reasons, eg. difficulties with collecting good quality and relevant sample material of sufficient amount in a safe manner, methodological challenges in microbiological labs, and securing sterility during the entire process from collection of sample to completed culture period(17). An alternative approach could therefore be to identify patients using indirect evidence of *C.acnes* infection.

Cytokines are key signal molecules of the immune system produced by various cells and tissues(18), including disc cells (19) involved in the pathophysiology of disc degeneration and MCs (20, 21). The regulation of each cytokine is intricate and involves the orchestration of a complex network of cells and other cytokines. The cytokines play a key role in the host response to infection and trauma as well as in immune-mediated diseases (22). Distinct cytokine patterns may distinguish between infection, and other inflammatory etiologies in conditions like encephalitis (23-25), systemic lupus erythematosus (26) and diabetes (27). Local infection with *C.acnes* can induce increased cytokine levels in the intervertebral disc (28-31) and serum (32-35) (table 1). Cytokine levels in peripheral blood mononuclear cells may distinguish pathogenic from commensal *C.acnes* phylotypes in skin (36). A possible infectious etiology of MCs could therefore potentially be associated with increased levels of certain cytokines and/or a distinct serum cytokine pattern.

Serum cytokine levels could also be associated with prognosis in patients with low back pain. Low serum interleukin(IL)-6 and high TNF-α was associated with worse outcome after 6 months in patients with acute low back pain (37). A systematic review found no other studies that looked at how cytokine levels was related to prognosis in non-specific low back pain (excluding e.g. disc herniation, search ended July 2019) (38). In patients with lumbar radicular pain, increased levels of IL-6 (39) and IL-8 was associated with worse outcome after 12 months follow-up (40).

Table 1- Summary table of evidence of increase in cytokine levels (in disc and serum), relevant for objective 2 (cytokines as predictors of treatment effect)

Cytokine	*Measured in intervertebral disc*	*Measured in Serum*
IL-1β	↑ in C.acnes stimulated human disc cells(30, 31)	↔ inflammatory back pain (ankylosing spondylitis)(41)
IL-8	↑ in histologically and PCR confirmed C.acnes discitis(28) ↑ in C.acnes inoculated rat discs(29) ↑ in C.acnes stimulated human disc cells(30, 31)	↑ in acne(33) ↔ in C.acnes prostatitis(35) ↑ in occult infections (Onchocerca volvulus infection(42))
TNF-α	↑ in histologically and PCR confirmed C.acnes discitis(28)	↑ in acne(33), ↑ in C.acnes hepatitis(32) ↑ inflammatory back pain (axial SpA)(43)
IL-6	↑ in C.acnes stimulated human disc cells(30, 31)	↑ in C.acnes hepatitis(32), ↔ in C.acnes prostatitis(35) ↑ periprostethic (knee and hip) infection generally(44) and with C.acnes(34) ↑ inflammatory back pain (axial SpA(43) and seronegative SpA(45))
IFN-γ		↑ in C.acnes hepatitis(32) and presence of serum bacterial DNA or cell wall products (lipopolysaccharides)(46-49)
IL-12		↑ in C.acnes hepatitis(32)
IL-10		Associated with an increased susceptibility to and persistence of various bacterial agents (50)
CCL3 (MIP-1α)	↑ in C.acnes stimulated human disc cells(31) ↑ in histologically and PCR confirmed C.acnes discitis(28)	↑ in occult infections (Onchocerca volvulus infection(42))
CCL2 (MCP-1)	↑ in histologically and PCR confirmed C.acnes discitis(28)	↑ in occult infections (Onchocerca volvulus infection(42))

The AIM-trial

The AIM trial was a six centre, randomised, parallel-group, placebo-controlled trial on the effects of three months of treatment with Amoxicillin in chronic low back pain patients with type 1 or 2 MCs at a level with previous disc herniation (defined as index level). In primary type 1 MCs, type 1 is the most extensive type at the evaluated endplate. In secondary type 1 MCs, type 1 is present, but another type is more extensive. Patients were classified in the MC type 2 group if they had type 2 MCs, but not any type 1 MCs (primary or secondary), at a previously herniated disc level.

Further details of the inclusion and exclusion criteria can be found in (1).

The primary outcome in the AIM trial was pain-related disability measured by the Roland-Morris Disease Questionnaire (RMDQ, score range 0 to 24) at one-year follow-up. Secondary outcomes included pain-related disability assessed by the ODI and low back pain intensity on a 0-10 NRS (1). The minimal clinically important between-group difference in mean RMDQ score was defined as 4 points.

The sample size for the main trial results was calculated to assess the treatment effect separately in each MC type group (1/2). In each MC type group, the study was 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. In total 180 patients were actually included (intention to treat population), 118 in the MC type 1 group and 62 in the MC type 2 group (16).

The trial protocol gives further details regarding sample size calculation and trial methods (1).

The AIM trial showed a statistically significant – but not clinically important – effect of Amoxicillin in the type 1 MC group, but no effect in the in the type 2 MC group (16). Subgroup analyses indicated a clinically relevant effect in patients with MC related oedema on MRI STIR sequences/series (7) and possibly also a borderline clinically relevant effect in younger patients (aged < 40 years)(6).

Patients

Patients included in this present study were a selected group with as 'pure' MC types as possible, who gave consent to participate in the AIM trial, and who provided blood samples and underwent MRI prior to enrolment. Of those selected 78 subjects were eventually randomized and took part in the AIM trial, 6 subjects did not satisfy the AIM eligibility criteria and were not included in the AIM trial (figure 1).

The selection criteria for this present study were, firstly, that two study radiologists agreed that the MRI inclusion criteria of the AIM trial (52) were fulfilled (i.e., type 1 and/or type 2 MCs with height ≥ 10% of vertebral body height and diameter > 5 millimeters existed at the level of a prior disc herniation (index level)). Secondly, that type 1 was the most extensive MC type at two or more index level endplates (categorized as MC1, n=47) or that type 2, but not type 1, existed at the index level (categorized as MC2, n=48). From this selected group we included 84 subjects (46 MC1, and 37 MC2) in the present study.

Controls

50 healthy individuals without back pain last 3 months provided blood samples at the blood donor centre (Blodbanken), Oslo University Hospital. All patients and controls provided a written informed consent, and the trial as well as the added project with controls was approved by the Regional Ethics Committee in South East Norway (project 2014/158/REK sør-øst C)

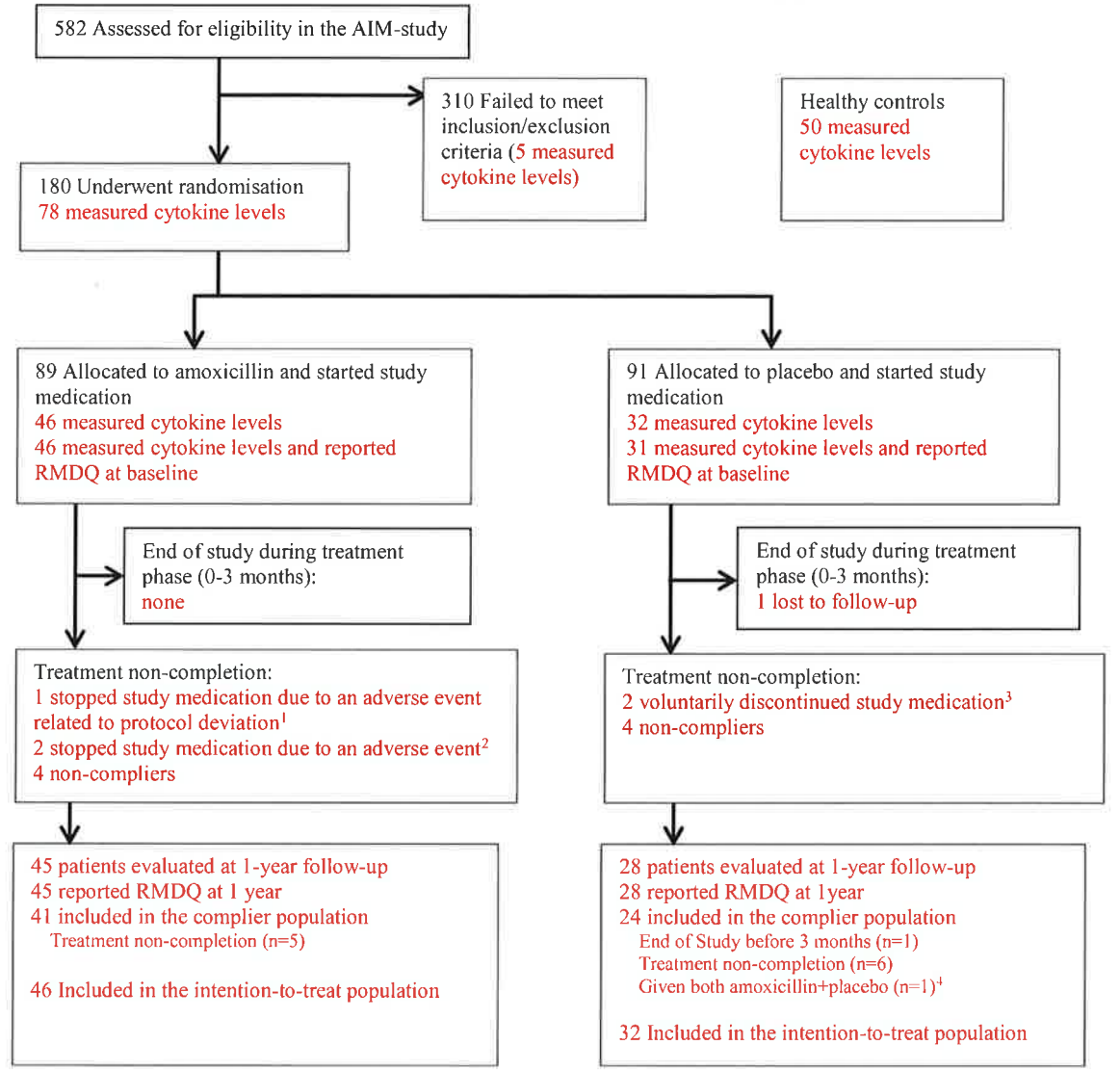
Cytokine measurements

We measured a standard panel of cytokines in serum using a Pro Human Chemokine multi-bead assay (Bio-Rad, Norway). Data was recorded with a Luminex IS 100 instrument (Bio-Rad, Hercules, CA, USA) and protein concentrations were finalized using recombinant standard curves. One cytokine (CXCL5) was not detected in the measurements of most patients and was excluded from the statistical analyses.

Cytokine profiling

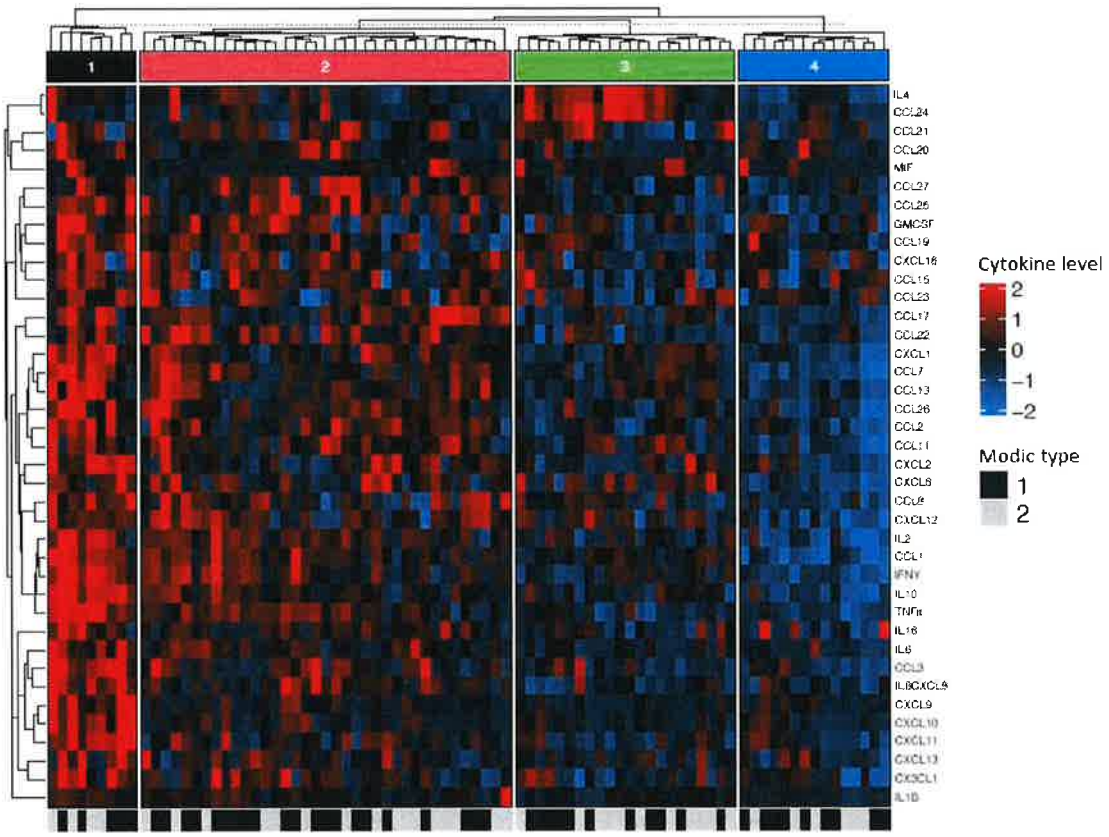
In order to identify subsets of subjects based on cytokine profiles, we performed cluster analyses using unsupervised hierarchical clustering and k-means partitioning. We split the heatmap using k-means partitioning of patients and hierarchical clustering of cytokines using the complex heatmap R package (53). Individuals were clustered by using complete linkage of Euclidean distance, and cytokines were clustered using Pearson correlation. The analysis was based on between-subject distances in log-transformed cytokine values standardised to z-scores, thus individuals with shorter distance would more likely be clustered together. The cluster analysis was performed without any post-randomization variable and done prior to any analyses described in this SAP. The results were a classification of patients with MCs into 4 categories (named as category 1 to 4 below) based on their cytokine profile (figure 2) that will be used in the analyses for this SAP.

Figure 1 – Flowchart (with numbers relevant for cytokine measurements in red)



1 One patient in the amoxicillin group became pregnant (protocol deviation as all patients were instructed to use contraception), not included in the complier population
2 Two patients in the amoxicillin group stopped study medication due to adverse events and were included in the complier population
3 Two patients in the placebo group discontinued because they started three month treatment with amoxicillinc/clavulanic acid
4 Due to a mistake at pharmacy, the patient was given a mix of bottles containing amoxicillin and placebo

Figure 2 -Heatmap of cytokine levels with classification of patients based on unsupervised hierarchical clustering using k-means partitioning



The heatmap X-axis shows patients (ordered by clustering based on complete linkage with Euclidean distance measure), and the Y-axis shows cytokines (ordered by clustering based on Pearson correlation as distance measure). The top colour row highlights the 4 categories of patients which will be used in analyses

Objectives

The objectives covered in the present Statistical Analysis Plan are:

1. To evaluate whether certain profiles of serum cytokine levels predict any treatment effect of 3 months oral amoxicillin in patients with chronic low back pain and type 1 or 2 MCs at the level of a previous lumbar disc herniation (cytokines as predictors of antibiotic treatment effect)
2. To evaluate whether certain profiles of systemic serum cytokine levels provide added prognostic value over known prognostic factors in patients with chronic low back pain and type 1 or 2 MCs at the level of a previous lumbar disc herniation (cytokines as prognostic factors independent of treatment)

Hypotheses

For each objective (1 and 2 above), we made respective hypotheses that

1. Patients with a certain pattern of serum cytokine levels at baseline in the antibiotic group report a significantly lower RMDQ score at one-year follow-up than patients with the similar cytokine pattern in the placebo group
2. Serum proinflammatory cytokine levels are associated with change in RMDQ score from baseline to one-year follow-up, given known prognostic factors.

Statistical principles

All analyses described in this plan are considered *exploratory* in the sense that the present objectives were not described in the original protocol, and are performed after results for treatment efficacy in the main trial (5) and in clinical and radiological subgroups (6, 7) were available. Any additional analyses not described in this SAP will be identified as *post hoc* in the article. Some analyses are based on *pre-defined subgroups*, (e.g. defined in advance of analyses), others on *post-defined subgroups*, meaning that subgroups will be defined by a data-driven approach.

All outcomes are analysed for superiority of one intervention compared to the other (hypothesis 1) or of having compared to not having a prognostic factor (hypothesis 2). 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.

Statistical significance is claimed on the significance level (alpha) of 0.05 (two-sided). We will not adjust the significance level due to multiple testing as these analyses are hypothesis generating (exploratory) in the sense that any evidence of predictors of treatment effect will need to be confirmed in future studies (54).

Statistically significant differences of < 4 RMDQ points will not be considered clinically relevant (52). We calculate power / sample size for the top ranked analysis of pre-defined subgroups for each hypothesis, as recommended by the PROGRESS framework (3) (details given below under the respective analyses).

All analyses will be carried out using software packages Stata and R, blinded to treatment group.

Analysis populations

We will perform all analyses of potential cytokines as prognostic factors and as predictors of treatment effect in the ITT population (n=78).

Baseline cytokine profiles

According to the cluster analysis described above subjects will be categorized by one of four cytokine profiles (figure 2).

Description of baseline characteristics

For analyses of predictors of treatment effect

In addition to the demographic and baseline characteristics given in the main report of the AIM trial (16), we will report for each treatment group number and percentages of patients with arthritis, osteoarthritis, diabetes, cardiac/pulmonary disease, psychiatric disease, self-reported sleep disturbance and STIR composite variables (groups 2 and 3 (7)). For each treatment group we will also report the distribution of the cytokine profile categories at baseline (Table 5). We will summarise categorical variables by numbers and percentages, and continuous variables by mean and SD (if normal distribution) or median and interquartile range (if skewed distribution). We will note the clinical importance of any imbalance in baseline data between the treatment groups without statistical testing.

For analyses of prognostic factors

The report on baseline characteristics for the paper on prognostic factors will describe for each prognostic factor variable of interest (i. – ii. In table 3) the same baseline characteristic as in the report on predictors of treatment effect, in each category of the prognostic factor.

Methods to handle missing data

Analyses of predictors of treatment effect (hypothesis 1) and of prognostic factors (hypothesis 2) will be based on imputed values unless stated otherwise. Missing RMDQ, ODI, LBP intensity and fear avoidance values will be imputed. For RMDQ we will first perform a mean imputation 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 score at baseline, 3 months

and 12 months, and fear avoidance at baseline, using index EQ5D-5L score (at baseline, 3 and 12 months), and the following baseline variables: age, leg pain (NRS 0-10), comorbidity (Functional Comorbidity Index), emotional distress, physical work load, psychiatric comorbidity (yes/no), other specific comorbidity (numerical, one point for each of rheumatoid arthritis, osteoarthritis, Diabetic, Cardiac/pulmonary), sick leave, NSAIDs use (at baseline and one-year follow-up), sleep disturbance, BMI, smoking, former surgery for disc herniation, study center, Modic type group, STIR composite group, cytokine classification group, principal component 1 (from analyses 3 A and B under analyses of hypothesis 1), treatment group, and the interaction terms cytokine classification group x treatment, principal component 1 (for 3A) x treatment and principal component 1 (for 3B) x treatment as predictors in the imputation model. The imputation model can be reduced if the imputation does not converge. Fifty imputations will be performed(55) and the imputation model will include the same patients who are included in the analysis (n=78).

Analyses of cytokines as predictors of antibiotic treatment effect (hypothesis 1)

Descriptive statistics

We will present the heatmap and dendrogram in figure 2, with the same order of patients and cytokines (based on analysis performed for figure 2), but with an additional color-coded marker row for treatment group, including a color-coded bar classifying patients into 1 (≤30% improvement of RMDQ from baseline to 1 year), 2 (>30% and ≤50% improvement of RMDQ from baseline to 1 year), 3 (>50% and ≤75% improvement of RMDQ from baseline to 1 year, 4 (>75% improvement of RMDQ from baseline to 1 year).

Inferential statistics

To explore potential predictors of treatment effect we will perform six main analyses (1A, 1B, 1C, 2A, 3A, 3B) of predictors of antibiotic treatment effect (Table 2). All six analyses will be performed using multiple regression with RMDQ score at one year as dependent variable and the interaction term (predictor*treatment group) as independent variable adjusted for baseline RMDQ score, the potential predictor, the treatment group, the randomization stratification variables (MC type (1/2) and for former disc herniation surgery (yes/no)).

We will only use the baseline value for any potential predictor of treatment effect and for any covariate. Within each category (subgroup) of categorical variables, we will report the treatment effect predicted in the interaction term analysis by marginalizing (using the margins command in Stata) over the other covariates (and not perform a separate stratified analysis).

We will present results of analyses 1A, 1B, 1C, 2A, 3A and 3B in a forest plot of effect estimates for each potential predictor of treatment effect with 95 % CIs and the p value for the interaction (predictor *treatment group). The overall treatment effect will be marked on the plot (56). The plot will also show the treatment effect within subgroups.

Table 2 –Main analyses of cytokine-defined predictors of treatment effect

Hypothesis and rationale	Variables and subgroups	Analysis
Hypothesis 1: Patients with a certain pattern of serum cytokine levels at baseline in the antibiotic group report a significantly lower RMDQ score at one-year follow-up than patients in the placebo group. Rationale: An infection of the intervertebral disc (discitis) could be associated with a distinct serum cytokine pattern that is different from the pattern in a patient without such infection. We have ranked the *post-defined* (data-driven) subgroup analysis first as there is a limited knowledge about C.acnes intervertebral discitis and serum cytokine levels, and relevant classifications and cut-offs might go unnoticed in analysis of *pre-defined* subgroups (57). In the selection of the *pre-defined* subgroups, we pay more attention to the pattern rather than the number and amount of increased cytokine levels. In analyses 2, we hypothesise category 1 and 2 is the most likely category of patients with treatment effect, due to increased levels of TNF-α, IFN-γ and IL-10 (and correlated cytokines).	**Post-defined subgroups** Subgroups are not defined yet but will be defined based on a predefined data-mining statistical analyses with recursive partitioning (stage 1 of analysis 1). The three recursive partitioning models differ only in terms of variables included: i. Cytokines with a biological rationale for a possible association with C.acnes infection in disc (IL-1β, IL-8, TNF-α, IL-6, IFN-γ, IL-12, IL-10, CCL3, CCL2) ii. Cytokines with a biological rationale for a possible association with C.acnes infection in disc + age and STIR composite variable iii. All 39 cytokines **Pre-defined subgroups** Prespecified classification based on cluster analysis (figure 2), giving four categories of patients. iv. Classification group dichotomized into category 1+2 (hypothesized group with treatment effect) vs category 3+4. Prespecified classification based on principal component 1 of principal component analyses of: v. Cytokines with levels that were significantly different from the healthy controls (CCL11, CCL22, CCL21, IL4, CCL26, IL6, CXCL13, CX3CL1, CCL27, CXCL6, CCL20, IFNγ, CCL19, IL2, CCL17, IL16, CCL25, CCL7, CCL13, MIF, CCL3, CCL15, CCL23, CXCL16) vi. All 39 cytokines	**1 A** **1 B** **1 C** **2 A** **3 A** **3 B**

Main analyses

The following main analyses are ranked in order, from what we consider most to least likely predictors of antibiotic treatment effect (based on prior evidence and knowledge of cytokines).

Analyses 1 (A, B and C), based on post-defined subgroups

For analysis 1, we will perform a two-stage analysis. The first stage will be based on a Recursive partitioning analysis in order to classify patients into categories, while the second stage will be a regression analysis with the categories defined in stage 1 as potential predictors of treatment effect. The recursive partitioning analyses has the advantage that it can identify subgroups defined based on more than one variable.

In stage 1, we will use Recursive partitioning analyses (using the R package SIDEScreen), with RMDQ at 1 year follow up as dependent variable, each branch assessed by the effect estimate of the interaction term between the subgroup and the treatment group. Stopping conditions will be L=3 and n_{min}=5, and a complexity criterion 0.5 (58). An adaptive SIDEScreen procedures approach will be used(58), selecting the biomarkers with variable importance scores above a certain threshold, defined as 15th percentile of the null distribution of the largest variable importance score (that is, if there is completely noninformative cytokines, no cytokine will be selected for the second step 85% of the time). These stage 1 analyses will only include 73 patients with nonmissing outcomes at one-year follow-up. The three following stage 1 models will be analysed (differs only in terms of variables included), in prioritized order:

A. Variables will include cytokines with a biological rationale for a possible association with C.acnes infection in disc (IL-1β, IL-8, TNF-α, IL-6, IFN-γ, IL-12, IL-10, CCL3, CCL2).

B. Variables will include cytokines with a biological rationale for a possible association with C.acnes infection in disc (IL-1β, IL-8, TNF-α, IL-6, IFN-γ, IL-12, IL-10, CCL3, CCL2) and also include age and STIR composite variable.

C. All 39 cytokines are included as variables.

In all analyses (A, B and C) we will use cytokine variables as categorical with cutoffs at -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5 and 2 SD above the mean in the healthy control groups, STIR composite variable as categorical (1, 2 or 3) and age as continuous variable.

In stage 2, a regression analysis (as described above) will be performed (n=78), where the potential predictor of treatment effect of interest will be defined based on the results of the stage 1 analyses. Each of the three stage 1 analyses (A, B and C) will result in one regression analysis that we will perform in prioritized order (1A, 1B and 1C, respectively).

Analyses 2, based on prespecified subgroups

We will use the same regression statistical model for the analyses 2 (n=78), with the cytokine classification group (4 categories of patients, figure 2) as the potential predictor of treatment effect. The classification group variable will be dichotomized into category 1 and 2 vs category 3 and 4. We hypothesize a treatment effect of antibiotics in the combined classification category 1 and 2.

Analysis 2 will be underpowered to detect a difference of 4 RMDQ points. The number of participants (n=24 vs n=13) will provide the power to detect a difference in mean RMDQ scores of 5.0 (α=0.05, β=0.2, SD=5, (using 'power twomean' command in Stata)) in the stratified analysis (answering what the effect of antibiotics in the combined group of cytokine classification category 1+2 is). Although 5.0 RMDQ points is higher than what we assumed patients consider as minimally clinically important difference, it is still less than 50% of the mean baseline value, and we find this an acceptable effect size given the nature of the intervention and that the former trial reported a difference in means of 8 RMDQ-points between the antibiotic and the placebo group. Inserting n=13 (placebo) and n=24 (amoxicillin) for the subgroup of interest, and n=19 (placebo) and n=22 (amoxicillin) for the comparable subgroup (see table 5A), a twoway ANOVA power calculation (using 'power twoway' command in Stata, given α=0.05, SD=5, and effect size=6 RMDQ-points, provide a power of 0.72 for the interaction test (answering what the difference between the effect of antibiotics in the combined cytokine classification category 1+2 and the effect of antibiotics in combined cytokine classification category 3+4 is). The same numbers (n=13,24,19 and 22; α=0.05; SD=5) with β=0.2 provide an effect size of 6.6 RMDQ-points. The real power of our analysis is likely improved by the inclusion of the other covariates. However, as our analyses are likely underpowered to detect a minimally clinically important difference of 4 RMDQ-points, we will focus on the effect size rather than p-values when interpreting our results.

Analysis 3 (A and B)

Analyses 3 (A and B) are considered an overall test of prediction of treatment effect of cytokines. We will perform a two-stage analysis. The first stage will be principal component analyses performed in order to create new variables (principal components) based on cytokine levels, which we will test as a potential predictor of treatment effect in the second stage by a regression analysis. We will perform two stage 1 analyses (see models A and B below), each will form basis for a stage 2 analysis.

In stage 1, we will use Principal component analysis (by using the pca command in Stata) in two following models (differs only in terms of variables included), in prioritized order:

A. Variables will include cytokines with levels that were significantly different from the healthy controls (CCL11, CCL22, CCL21, IL4, CCL26, IL6, CXCL13, CX3CL1, CCL27, CXCL6, CCL20, IFNγ, CCL19, IL2, CCL17, IL16, CCL25, CCL7, CCL13, MIF, CCL3, CCL15, CCL23, CXCL16).
B. All 39 cytokines are included as variables.

These stage 1 analyses are already performed prior to publication of this SAP, giving a proportion of variance explained by principal component 1 of 34.9% in analysis A and 31.4% in analysis B.

In stage 2, a regression analysis (as described above) will be performed, where the potential predictor of treatment effect of interest will be defined as the principal component 1 from the stage 1 analyses A and B. Both stage 1 analyses (A and B) will result in one regression analysis that we will perform in prioritized order (3A and 3B, respectively).

Supportive and sensitivity analyses

As supportive analyses for 1 (A, B and C), we will perform:

I. Analyses with LBP intensity and ODI score at one year as dependent variables to assess if any subgroup effect is consistent across related outcomes, as recommended (59).

If clinically important differences between the treatment groups in analysis 1 (A, B and C), we will also perform the following supportive analyses:

II. Analyses with RMDQ score at 3 months follow-up as dependent variable, otherwise similar as above.

III. Responder analyses with responders defined as >30%, >50% and >75% improvement in RMDQ from baseline to one year, including calculation of numbers needed to treat. For these analyses we will consider >50% and >75% improvements as more likely in any underlying infection (discitis).

IV. Linear mixed effects (LME) models with RMDQ as dependent variable, and 3-part interaction term time x treatment group x cytokine classification group (treating time as a categorical variable), adjusted for MC type (1/2) and for former disc herniation surgery (yes/no). All LME model specifications will be selected based on Aikaike's Information Criterion. Based on the selected model, we will do a pairwise comparison of marginals to estimate the treatment effect for the cytokine classification groups of interest.

If analysis 2 (A and B) or 3 (A, B) result in clinically relevant treatment effect, we will consider the same supportive analyses (I-IV) as listed above.

Sensitivity analyses for 1 (A, B, C), 2 (A) and 3 (A, B):

-Independence of effect: If the treatment effect (difference in RMDQ score at 1 year follow-up between antibiotic and placebo groups) are more than 4 RMDQ points in any of the main analyses 1(A, B, or C), 2(A), or 3(A or B) we will repeat the same analysis including interaction terms of all such relevant variables x treatment group, age (dichotomized as above or below 40 years) x treatment group and STIR composite variable x treatment group (without the individual covariates age and STIR composite variable) in the same regression analysis (59). Both age and STIR composite variable modified the treatment effect in our prior analyses.

Analyses of cytokines as prognostic factors (hypothesis 2)

For hypothesis 2, we define eleven main analyses (1, 2A, 2B, 3, 4A, 4B, 5A-E) of prognostic factors. These main analyses are ranked in order, from what we consider most to least likely to be prognostic factors (based on prior evidence and knowledge of cytokines).

Main analyses

We will perform linear regression models as main analyses that all will have the following in common:

-Relative change in RMDQ from baseline to one-year follow-up will be dependent variable. This differs from other analyses in the AIM trial (RMDQ at one-year follow-up was dependent variable) because we here consider it problematic to adjust for baseline RMDQ-score as it could induce bias (due to regression to the mean or horse-riding / collider problems) (60). Such adjustment could on the other hand also be highly desirable because baseline score is thought to be one of the most important predictors of future follow-up scores (61, 62). Our choice of dependent variable attempts to take the baseline value into account, while avoiding the possible problem related to adjusting for the baseline value.

-Due to power limitations we will restrict the number of covariates/independent variables to eight (10 patients/independent variable)(63). Based on suspected prognostic factors for LBP disability (61, 62) and suspected factors associated with serum cytokine levels (64, 65), we will add the following as covariates into the model due to possible confounding: age, leg pain (0-10 NRS), fear avoidance, emotional distress (dichotomized into HSCL-25 ≥1,75 and <1,75), psychiatric comorbidity (yes/no), other specific comorbidity (numerical, one point for each of rheumatoid arthritis, osteoarthritis, Diabetic, Cardiac/pulmonary) and sick leave (categorized into 1. working full time, 2. partial sick leave and 3. complete sick leave / disability pension / unemployed / student/other/unknown). Age, leg pain and fear avoidance will be continuous variables.

For each model we will evaluate:

1. The predictiveness by comparing the multiple correlation coefficient R^2 of the complete model with the R^2 in the model without the independent variable of interest (66).
2. The effect size described in terms of relative change in RMDQ. We will consider effect size >30% as clinically relevant, which equates approximately to a difference of 4 out of 12.8 (at baseline) RMDQ points.

The following models (1, 2A, 2B, 3, 4A, 4B, 5A, 5B, 5C, 5D, 5E) will differ in choice of independent variable of interest, which will be:

1. Cytokine category group based on cytokine classification group (4 categories of patients, figure 2). Each category of patients will be assessed versus the three other categories combined.
2. Classification variable defined by a Recursive Partitioning analysis using RPART command in R. Relative change in RMDQ from baseline to one-year follow-up will be dependent variable. Only patients nonmissing values in both baseline and one-year follow-up RMDQ score will be included. Stopping conditions will be maxdepth=3 and n_{min}(minimal number of observations in terminal/leaf node)=5. We will use tenfold cross-validation and a complexity parameter of 0,01. We will perform two Recursive Partitioning models that differs only in what variables that are included:

   A. all 39 cytokines
   B. all 39 cytokines, age, leg pain (0-10 NRS), fear avoidance, emotional distress, psychiatric comorbidity (yes or no), physical workload, and sick leave.

   In models 2A and 2B we will use cytokine variables as categorical with cutoffs at -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5 and 2 SD above the mean in the healthy control groups, and other variables as defined for the linear regression analyses above. Physical workload will be categorized into 1. mostly sitting, 2. job requires a lot of walking and 3. job requires a lot of walking and lifting / job requires physically heavy work.
3. Abnormal cytokine values, based on PCA analysis of healthy controls and patients with MCs (figure 3), dichotomized into normal (inside yellow circle) and abnormal (outside yellow circle) cytokine levels.
4. Principal component 1 (continuous variable) from
   A. PCA analysis 3A of cytokines as predictors of treatment effect
   B. PCA analysis 3B of cytokines as predictors of treatment effect
5. Individual cytokines categorized with cutoffs at 0, 1 and 2 SD above the mean in the healthy control groups
   A. MIF levels
   B. TNFα levels (67)
   C. IL-6 levels
   D. IL-8 levels
   E. IL-1β levels

   For model 2B we will avoid adding any covariate that is also used to define the independent variable of interest. E.g. if the recursive partitioning model define the independent variable based on age and leg pain, we will exclude age and leg pain as covariates in the final linear regression model.

For all models a sample size (n=78) will, assuming an R^2 of 30% in the incomplete model including seven known prognostic factors but without the one independent variable of interest (estimates of R^2 in models excluding baseline values are 20-30% (68, 69)), give a power (α=0.05, β=0.2) to detect increases of 7% (absolute percentage points) in R^2 in the complete model with one variable of interest and seven covariates (using the 'power rsquared' function in Stata). For model 1, detecting a difference of 30% in change from baseline RMDQ (SD 5) between two groups of n=28 and n=44 will give a power (α=0.05) of 90% (using the 'power oneway' function in Stata).

Table 3 – Main analyses of cytokine-defined prognostic factors

Hypothesis and rationale	Variables and subgroups	Analysis
Hypothesis 2: Serum proinflammatory cytokine levels are associated with change in RMDQ score from baseline to one-year follow-up. Rationale: Serum cytokine levels could be associated with prognosis in patients with low back pain. Low serum IL-6 and high TNF in patients with acute low back pain was associated with worse outcome after 6 months (37). In patients with lumbar radicular pain, increased levels of IL-6 and IL-8 was associated with worse outcome after 12 months follow-up (40). Subgroups ii. and iii. are motivated by their ability to assess different cut-off values of cytokines, which could be important given the possibility that physiological reactivity may require threshold concentrations (67).	**i.** Prespecified classification based on cluster analysis (figure 2), giving four categories of patients (pre-defined) **ii.** Subgroups defined based on data-mining statistical analyses with recursive partitioning of all 39 cytokines (post-defined) **iii.** Subgroups defined based on data-mining statistical analyses with recursive partitioning of all 39 cytokines and known prognostic factors for disability: age, leg pain (0-10 NRS), fear avoidance, emotional distress, comorbidity, physical workload, and sick leave (post-defined) **iv.** Abnormal cytokine values, based on PCA analysis of healthy controls and patients with MCs (figure 3), dichotomized into normal (inside yellow circle) and abnormal (outside yellow circle) cytokine levels (pre-defined) **v.** Principal component 1 from PCA analysis 3A of cytokines as predictors of treatment effect (pre-defined) **vi.** Principal component 1 from PCA analysis 3B of cytokines as predictors of treatment effect (pre-defined) Individual cytokines, categorized with cutoffs at 0, 1 and 2 SD above the mean in the healthy control groups (pre-defined) **vii.** MIF levels **viii.** TNFα levels **ix.** IL-6 levels **x.** IL-8 levels **xi.** IL-1β levels	**1** **2 A** **2 B** **3** **4 A** **4 B** **5 A** **5 B** **5 C** **5 D** **5 E**

Supportive and sensitivity analyses

Supportive analyses:

Calculate Area under ROC curve (AUC), cutoff value, positive predictive value, negative predictive value, sensitivity and specificity (internal validation/assessing performance (70, 71))
With 30% improvement of RMDQ from baseline to one-year follow-up as the gold standard status (diagnostic variable), we will calculate Area under ROC curve (AUC), cutoff value, positive predictive value, negative predictive value, sensitivity and specificity for each independent variable of main analyses (1, 2A and 2B) as a prognostic variable of interest (test variable). We will consider the same calculations for each independent variable of the additional analyses (3 to 5).

Assess the direct effect of cytokines on relative change in RMDQ score (baseline to 1-year) in the possible case that baseline value of RMDQ mediates any overall/total effect:
For each linear regression model of the main analyses (1, 2A and 2B) we will perform the same analysis including baseline value of RMDQ as a covariate. These models could also be interpreted to provide estimates of the prognostic effect (without causal inference) of cytokines given a specific baseline value, and we can therefore safely ignore the possible problems with bias mentioned above (under Main analyses). We will consider the same sensitivity analysis for each linear regression model of the main analyses (3 to 5).

Assess the effect of cytokines on relative change in RMDQ score (baseline to 1-year) in the placebo group only, in the possible case that that the subgroup predicts treatment effect:
We will consider (depending on results and power), for each linear regression model of the analyses 1, 4 and 5, performing the same analysis including only patients in the placebo group.

Internal validation
We will consider internal validation on any analyses 1 to 5 (depending on results) by using bootstrapping.

Sensitivity analyses:

The following sensitivity analyses will be performed for the main analyses (1, 2A and 2B), and will be considered performed for the main analyses 3 to 5:

Assess the possibility that covariates included in main analyses are non-confounders/mediators:
For each linear regression model of the main analyses 1, 2A and 2B, we will perform unadjusted analysis with relative change in RMDQ from baseline to one-year follow-up as dependent variable, and the same independent variable as used in each respective model (1, 2A and 2B). These unadjusted analyses are recommended as we cannot rule out the possibility that the covariates included in the main analyses could be non-confounders or mediators (72). Each model will be compared to their respective main analysis by assessing R^2.

Assess other possible confounders (not included in main analyses due to power limitations):
We will perform a sensitivity analysis for the main analysis (1, 2A and 2B), with the same statistical model but only including covariates with statistically significant beta coefficient in the corresponding main analysis, and depending on available number of further covariates (given maximum eight), the

following covariates will be considered: treatment group (amoxicillin/placebo), comorbidity, NSAIDs use (at baseline and one-year follow-up), physical workload, sleep disturbance(64), BMI, smoking, LBP intensity at baseline (0-10 NRS), and prior surgery for disc herniation. Each sensitivity analysis will be compared to their respective main analysis by assessing R^2.

The following sensitivity analyses will be considered performed:

Assess the possibility of non-linear effects of the cytokine

For each main analyses 5 (A-E), we will consider using a correspondent multivariable fractional polynomial (MFP) model or other polynomial regression model (73). All covariates of the original regression analyses will be forced included (by setting nominal p-values for variable selection to one in MFP model), only the independent variable of interest will be tested for a non-linear relationship (degrees of freedom for other continuous variables will be set to one) with (for MFP) standard powers (-2, -1, -0.5, 0, 0.5, 1, 2, and 3) and a fractional polynomial election significance level of 5% (using the mfp command in Stata). Each MFP model will be compared to its respective main analysis by assessing R^2.

Independence of effect:

In the case of statistically significant association between the relative change in RMDQ from baseline to one-year follow-up and the independent variables in analyses 1 to 5, we will assess the independence of any effect by including all variables significantly associated with outcome into the same model. We will add the following as covariates into the model due to possible confounding (same as above) but will consider leaving out some covariates if many cytokines are included: age, leg pain (0-10 NRS), fear avoidance, emotional distress, comorbidity, physical workload, and sick leave. We will also consider selecting out any cytokines in the model that are highly correlated.

Further tables and figures

Table 4 - Criteria for evaluating the results of the present analyses of potential predictors of treatment effect

When interpreting the plausibility of any potential predictors of treatment effect, we will use the following criteria for evaluating subgroup analyses (59) (currently known answers for present analysis in parentheses):

Design

Is the subgroup variable a characteristic measured at baseline or after randomisation?* (At baseline)
Is the effect suggested by comparisons within rather than between studies?

Was the hypothesis specified a priori? (Partly. The hypotheses were specified before any analyses described in the current SAP were performed, but after results for treatment efficacy in the main trial (5) and in clinical and radiological subgroups (6, 7) were available)
Was the direction of the subgroup effect specified a priori?* (Yes)
Was the subgroup effect one of a small number of hypothesised effects tested?

Analysis

Does the interaction test suggest a low likelihood that chance explains the apparent subgroup effect?
Is the significant subgroup effect independent?*

Context

Is the size of the subgroup effect large?
Is the interaction consistent across studies?
Is the interaction consistent across closely related outcomes within the study?*
Is there indirect evidence that supports the hypothesised interaction (biological rationale)?

*New criteria.

Table 5A – Distribution of baseline cytokine categories by treatment group

Variable	Amoxicillin group (N = 46)		Placebo group (N = 32)	
	n	%	n	%
Cytokine category, used for analyses 2 (A and B)				
category **1**	5	11	3	9
category **2**	19	41	10	31
category **3**	13	28	7	22
category **4**	9	20	12	38
Sum	46	100	32	100

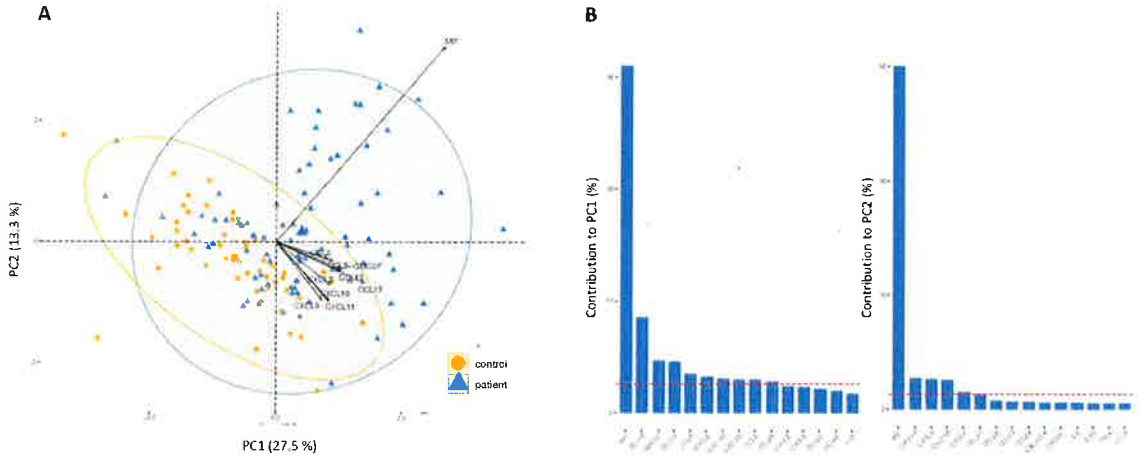
Table 5B – Distribution of baseline cytokine categories by treatment group in those with nonmissing values for RMDQ at baseline and 1 year follow-up

Variable	Amoxicillin group (N = 45)		Placebo group (N = 27)	
	n	%	n	%
Cytokine category, used for analyses 2 (A and B)				
category **1**	5	11	3	11
category **2**	19	42	9	33
category **3**	13	29	6	22
category **4**	8	18	9	33
Sum	45	100	27	100

Table 5C – Distribution of baseline cytokine categories by treatment group in compliers

Variable	Amoxicillin group (N = 41)		Placebo group (N = 24)	
	n	%	n	%
Cytokine category, used for analyses 2 (A and B)				
category **1**	5	12	3	13
category **2**	18	44	8	33
category **3**	12	29	5	21
category **4**	6	15	8	33
Sum	41	100	24	100

Figure 3 - PCA biplot of serum cytokine levels in healthy controls and patients with MCs



References

1. Storheim K, Espeland A, Grovle L, Skouen JS, Assmus J, Anke A, et al. Antibiotic treatment In patients with chronic low back pain and Modic changes (the AIM study): study protocol for a randomised controlled trial. Trials. 2017;18(1):596.
2. Bråten LC RM, Espeland A, Storheim K, Zwart JA, Hellum C, Aßmus J. Statistical analysis plan (SAP) for clinical outcomes in the AIM-study: A randomized trial of antibiotic treatment in patients with chronic low back pain and Modic Changes (the AIM-study). 2018.
3. Riley R. PROGRESS (PROGnosis RESearch Strategy) 2020 [Available from: https://www.prognosisresearch.com/.
4. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318(23):2337-43.
5. Bråten LCH, Rolfsen MP, Espeland A, Wigemyr M, Aßmus J, Froholdt A, et al. Efficacy of antibiotic treatment in patients with chronic low back pain and Modic changes (the AIM study): double blind, randomised, placebo controlled, multicentre trial. bmj. 2019;367.
6. Bråten LCH, Grøvle L, Espeland A, Pripp AH, Grotle M, Helllum C, et al. Clinical effect modifiers of antibiotic treatment in patients with chronic low back pain and Modic changes-secondary analyses of a randomised, placebo-controlled trial (the AIM study). BMC musculoskeletal disorders. 2020;21(1):1-11.
7. Kristoffersen PM, Bråten LC, Vetti N, Grøvle L, Hellum C, Storheim K, et al. Oedema on STIR modified the effect of amoxicillin as treatment for chronic low back pain with Modic changes—subgroup analysis of a randomized trial. European Radiology. 2020:1-13.

8. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology. 1988;166(1 Pt 1):193.
9. Urquhart DM, Zheng Y, Cheng AC, Rosenfeld JV, Chan P, Liew S, et al. Could low grade bacterial infection contribute to low back pain? A systematic review. BMC Medicine. 2015;13:13.
10. Chen Z, Zheng Y, Yuan Y, Jiao Y, Xiao J, Zhou Z, et al. Modic Changes and Disc Degeneration Caused by Inoculation of Propionibacterium acnes inside Intervertebral Discs of Rabbits: A Pilot Study. BioMed research international. 2016;2016:9612437.
11. Li B, Dong Z, Wu Y, Zeng J, Zheng Q, Xiao B, et al. Association Between Lumbar Disc Degeneration and Propionibacterium acnes Infection: Clinical Research and Preliminary Exploration of Animal Experiment. Spine. 2016;41(13):E764.
12. Benoist M. The Michel Benoist and Robert Mulholland Yearly European Spine Journal Review: A survey of the "medical" articles in the European Spine Journal, 2018. European Spine Journal. 2019;28(1):10-20.
13. Capoor M, Ruzicka F, Machackova T, Jancalek R, Smrcka M, Schmitz J, et al. Prevalence of Propionibacterium acnes in Intervertebral Discs of Patients Undergoing Lumbar Microdiscectomy: A Prospective Cross-Sectional Study. PLoS One. 2016;11(8).
14. Geiss A, Larsson K, Rydevik B, Takahashi I, Olmarker K. Autoimmune properties of nucleus pulposus: an experimental study in pigs. Spine (Phila Pa 1976). 2007;32(2):168-73.
15. Albert HB, Sorensen JS, Christensen BS, Manniche C. Antibiotic treatment in patients with chronic low back pain and vertebral bone edema (Modic type 1 changes): a double-blind randomized clinical controlled trial of efficacy. European Spine Journal. 2013;22(4):697-707.
16. Bråten LCH, Rolfsen MP, Espeland A, Wigemyr M, Aßmus J, Froholdt A, et al. Efficacy of antibiotic treatment in patients with chronic low back pain and Modic changes (the AIM study): double blind, randomised, placebo controlled, multicentre trial. bmj. 2019;367:l5654.
17. McNamara AL, Dickerson EC, Gomez-Hassan DM, Cinti SK, Srinivasan A. Yield of Image-Guided Needle Biopsy for Infectious Discitis: A Systematic Review and Meta-Analysis. AJNR Am J Neuroradiol. 2017;38(10):2021-7.
18. Schirmer M, Kumar V, Netea MG, Xavier RJ. The causes and consequences of variation in human cytokine production in health. Current Opinion in Immunology. 2018;54:50-8.
19. Risbud MV, Shapiro IM. Role of cytokines in intervertebral disc degeneration: pain and disc content. Nat Rev Rheumatol. 2014;10(1):44-56.
20. Dudli S, Fields AJ, Samartzis D, Karppinen J, Lotz JC. Pathobiology of Modic changes. Eur Spine J. 2016;25(11):3723-34.
21. Dudli S, Miller S, Demir-Deviren S, Lotz JC. Inflammatory response of disc cells against Propionibacterium acnes depends on the presence of lumbar Modic changes. Eur Spine J. 2018;27(5):1013-20.
22. Turner MD, Nedjai B, Hurst T, Pennington DJ. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2014;1843(11):2563-82.
23. Kothur K, Wienholt L, Mohammad SS, Tantsis EM, Pillai S, Britton PN, et al. Utility of CSF cytokine/chemokines as markers of active intrathecal inflammation: comparison of demyelinating, anti-NMDAR and enteroviral encephalitis. PLoS One. 2016;11(8):e0161656.
24. Michael BD, Griffiths MJ, Granerod J, Brown D, Davies NW, Borrow R, et al. Characteristic cytokine and chemokine profiles in encephalitis of infectious, immune-mediated, and unknown aetiology. PLoS One. 2016;11(1):e0146288.
25. Kothur K, Bandodkar S, Wienholt L, Chu S, Pope A, Gill D, et al. Etiology is the key determinant of neuroinflammation in epilepsy: Elevation of cerebrospinal fluid cytokines and chemokines in febrile infection-related epilepsy syndrome and febrile status epilepticus. Epilepsia. 2019;60(8):1678-88.

26. Kyogoku C, Smiljanovic B, Grün JR, Biesen R, Schulte-Wrede U, Häupl T, et al. Cell-specific type I IFN signatures in autoimmunity and viral infection: what makes the difference? PloS one. 2013;8(12):e83776.
27. Yeung W-CG, Al-Shabeeb A, Pang CNI, Wilkins MR, Catteau J, Howard NJ, et al. Children with islet autoimmunity and enterovirus infection demonstrate a distinct cytokine profile. Diabetes. 2012;61(6):1500-8.
28. Yuan Y, Chen Y, Zhou Z, Jiao Y, Li C, Zheng Y, et al. Association between chronic inflammation and latent infection of Propionibacterium acnes in non-pyogenic degenerated intervertebral discs: a pilot study. European Spine Journal. 2018;27(10):2506-17.
29. Jiao Y, Yuan Y, Lin Y, Zhou Z, Zheng Y, Wu W, et al. Propionibacterium acnes induces discogenic low back pain via stimulating nucleus pulposus cells to secrete pro-algesic factor of IL-8/CINC-1 through TLR2–NF-κB p65 pathway. Journal of Molecular Medicine. 2019;97(1):25-35.
30. Schmid B, Hausmann O, Hitzl W, Achermann Y, Wuertz-Kozak K. The Role of Cutibacterium acnes in Intervertebral Disc Inflammation. Biomedicines. 2020;8(7):186.
31. Capoor MN, Konieczna A, McDowell A, Ruzicka F, Smrcka M, Jancalek R, et al. Pro-Inflammatory and Neurotrophic Factor Responses of Cells Derived from Degenerative Human Intervertebral Discs to the Opportunistic Pathogen Cutibacterium acnes. International Journal of Molecular Sciences. 2021;22(5):2347.
32. Romics Jr L, Dolganiuc A, Velayudham A, Kodys K, Mandrekar P, Golenbock D, et al. Toll-like receptor 2 mediates inflammatory cytokine induction but not sensitization for liver injury by Propioni-bacterium acnes. Journal of leukocyte biology. 2005;78(6):1255-64.
33. Basal E, Jain A, Kaushal G. Antibody response to crude cell lysate of propionibacterium acnes and induction of pro-inflammatory cytokines in patients with acne and normal healthy subjects. The Journal of Microbiology. 2004;42(2):117-25.
34. Di Cesare PE, Chang E, Preston CF, Liu C-j. Serum interleukin-6 as a marker of periprosthetic infection following total hip and knee arthroplasty. JBJS. 2005;87(9):1921-7.
35. Ugge H, Carlsson J, Söderquist B, Fall K, Andén O, Davidsson S. The influence of prostatic Cutibacterium acnes infection on serum levels of IL6 and CXCL8 in prostate cancer patients. Infectious agents and cancer. 2018;13(1):34.
36. Yu Y, Champer J, Agak GW, Kao S, Modlin RL, Kim J. Different Propionibacterium acnes phylotypes induce distinct immune responses and express unique surface and secreted proteomes. Journal of Investigative Dermatology. 2016;136(11):2221-8.
37. Klyne DM, Barbe MF, van den Hoorn W, Hodges PW. ISSLS Prize in clinical science 2018: longitudinal analysis of inflammatory, psychological, and sleep-related factors following an acute low back pain episode—the good, the bad, and the ugly. European Spine Journal. 2018;27(4):763-77.
38. Morris P, Ali K, Merritt M, Pelletier J, Macedo LG. A systematic review of the role of inflammatory biomarkers in acute, subacute and chronic non-specific low back pain. BMC musculoskeletal disorders. 2020;21(1):1-12.
39. Schistad E, Espeland A, Pedersen L, Sandvik L, Gjerstad J, Røe C. Association between baseline IL-6 and 1-year recovery in lumbar radicular pain. European Journal of Pain. 2014;18(10):1394-401.
40. Pedersen LM, Schistad E, Jacobsen LM, Røe C, Gjerstad J. Serum levels of the pro-inflammatory interleukins 6 (IL-6) and-8 (IL-8) in patients with lumbar radicular pain due to disc herniation: a 12-month prospective study. Brain, behavior, and immunity. 2015;46:132-6.
41. Hreggvidsdottir HS, Noordenbos T, Baeten DL. Inflammatory pathways in spondyloarthritis. Molecular immunology. 2014;57(1):28-37.
42. Lechner CJ, Gantin RG, Seeger T, Sarnecka A, Portillo J, Schulz-Key H, et al. Chemokines and cytokines in patients with an occult Onchocerca volvulus infection. Microbes and infection. 2012;14(5):438-46.

43. Lorenzin M, Ometto F, Ortolan A, Felicetti M, Favero M, Doria A, et al. An update on serum biomarkers to assess axial spondyloarthritis and to guide treatment decision. Therapeutic Advances in Musculoskeletal Disease. 2020;12:1759720X20934277.
44. Berbari E, Mabry T, Tsaras G, Spangehl M, Erwin PJ, Murad MH, et al. Inflammatory blood laboratory levels as markers of prosthetic joint infection: a systematic review and meta-analysis. JBJS. 2010;92(11):2102-9.
45. Tuttolomondo A, Pecoraro R, Buttà C, Di Raimondo D, Ferrante A, Della Corte V, et al. Arterial stiffness indexes and serum cytokine levels in seronegative spondyloarthritis: relationships between stiffness markers and metabolic and immunoinflammatory variables. Scandinavian journal of rheumatology. 2015;44(6):474-9.
46. Cowdery JS, Chace JH, Yi A-K, Krieg AM. Bacterial DNA induces NK cells to produce IFN-gamma in vivo and increases the toxicity of lipopolysaccharides. The Journal of Immunology. 1996;156(12):4570-5.
47. Rubins JB, Pomeroy C. Role of gamma interferon in the pathogenesis of bacteremic pneumococcal pneumonia. Infection and immunity. 1997;65(7):2975-7.
48. Heinzel FP. The role of IFN-gamma in the pathology of experimental endotoxemia. The Journal of Immunology. 1990;145(9):2920-4.
49. Nathan CF, Murray HW, Wiebe ME, Rubin BY. Identification of interferon-gamma as the lymphokine that activates human macrophage oxidative metabolism and antimicrobial activity. The Journal of experimental medicine. 1983;158(3):670-89.
50. Mege J-L, Meghari S, Honstettre A, Capo C, Raoult D. The two faces of interleukin 10 in human infectious diseases. The Lancet infectious diseases. 2006;6(9):557-69.
51. Gjefsen E, Gervin K, Goll G, Bråten LCH, Wigemyr M, Aass HCD, et al. Macrophage migration inhibitory factor: a potential biomarker for chronic low back pain in patients with Modic changes. RMD open. 2021;7(2):e001726.
52. Storheim K, Espeland A, Grøvle L, Skouen JS, Aßmus J, Anke A, et al. Antibiotic treatment In patients with chronic low back pain and Modic changes (the AIM study): study protocol for a randomised controlled trial. Trials. 2017;18(1):1-11.
53. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847-9.
54. Kent P, Keating JL, Leboeuf-Yde C. Research methods for subgrouping low back pain. BMC medical research methodology. 2010;10(1):1-14.
55. Van Buuren S. Multiple imputation. Flexible Imputation of Missing Data: Chapman and Hall/CRC; 2012. p. 51-78.
56. Cuzick J. Forest plots and the interpretation of subgroups. Lancet. 2005;365(9467):1308.
57. Lipkovich I, Dmitrienko A, B D'Agostino Sr R. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. Stat Med. 2017;36(1):136-96.
58. Lipkovich I, Dmitrienko A. Strategies for identifying predictive biomarkers and subgroups with enhanced treatment effect in clinical trials using SIDES. Journal of biopharmaceutical statistics. 2014;24(1):130-53.
59. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ. 2010;340:c117.
60. Glymour MM, Weuve J, Berkman LF, Kawachi I, Robins JM. When Is Baseline Adjustment Useful in Analyses of Change? An Example with Education and Cognitive Change. American Journal of Epidemiology. 2005;162(3):267-78.
61. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. The Lancet. 2018;391(10137):2356-67.
62. Hayden J, 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(8):781-96. e1.

63. Yao X, Vella E, Brouwers M. How to conduct a high-quality original study on a prognostic research topic. Surgical Oncology. 2018;27(4):A9-A13.
64. Heffner KL, France CR, Trost Z, Ng HM, Pigeon WR. Chronic low back pain, sleep disturbance, and interleukin-6. The Clinical journal of pain. 2011;27(1):35.
65. IJ LS, Scholten-Peeters G, Bontkes H, Coppieters M. Multiple confounders influence the association between low-grade systemic inflammation and musculoskeletal pain. A call for a prudent interpretation of the literature. The spine journal: official journal of the North American Spine Society. 2018;18(11):2162-3.
66. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. British journal of cancer. 1994;69(6):979.
67. Klyne DM, Barbe MF, Hodges PW. Systemic inflammatory profiles and their relationships with demographic, behavioural and clinical features in acute low back pain. Brain, behavior, and immunity. 2017;60:84-92.
68. Chan HL, Chiu TT. The correlations among pain, disability, lumbar muscle endurance and fear-avoidance behaviour in patients with chronic low back pain. Journal of back and musculoskeletal rehabilitation. 2008;21(1):35-42.
69. Woby SR, Watson PJ, Roach NK, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? European journal of pain. 2004;8(3):201-10.
70. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605.
71. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ. 2009;338:b604.
72. Schipholt IJL, Scholten-Peeters GG, Bontkes HJ, Coppieters MW. Authors' Reply: Confounding and mediation to reveal the true association between systemic inflammation and musculoskeletal pain. The Spine Journal. 2019;19(11):1901.
73. Auget J-L, Balakrishnan N, Mesbah M, Molenberghs G. Advances in statistical methods for the health sciences: applications to cancer and AIDS studies, genome sequence analysis, and survival analysis: Springer Science & Business Media; 2007.