[redacted]

[redacted]

20th Dec 2024
Version No. 7.0

PERSICA Statistical Analysis Plan
Confidential

EudraCT Number:	2018-004488-30
Study Number:	Persica 002
Protocol Title:	A phase 1b study investigating the safety, tolerability and efficacy of PP353 in the treatment of patients with chronic low back pain associated with vertebral body endplate bone oedema (Modic 1)
Drug:	PP353
Sponsor:	Persica Pharmaceuticals Ltd. 7 Denne Hill Business Centre, Womenswold, Canterbury, Kent CT4 6HD
Prepared by (Statistician): [redacted] Statistician, Veramed Ltd	[redacted] [redacted]
Approved by: [redacted] Data Manager, Micron Research	[redacted] [redacted] Signature:

Approved by: [redacted] Medical Consultant, Weatherden	[redacted] Signature:
Approved by: [redacted] Statistician, Micron Research	[redacted] Signature:
Statistical Analysis Plan: Version No. and Date	Version 7.0, 20 Dec 2024

TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS

AE	Adverse Event
bpm	Beats per minute (heart rate)
CLBP	Chronic low back pain
CRO	Contract research organization
DEMRIQ	Dynamic contrast enhanced MRI quantification
ECG	Electrocardiogram
ES	Enrolled Set
eCRF	Electronic case report form
EDC	Electronic Data Capture
FAS	Full Analysis Set
IRT	Interactive response technology
kg	Kilogram
LBP	Low Back Pain
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
mL	Milliliter
ms	Milliseconds
MMRM	Mixed Model for Repeated Measures
MRI	Magnetic Resonance Imaging
NRS	Numeric rating scale
PCS	Pain Catastrophising Score
PK	Pharmacokinetics
PPS	Per Protocol Set
PT	Preferred term
RMDQ	Roland-Morris Disability Questionnaire
RS	Randomised Set
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SOC	System organ class
SRC	Safety Review Committee
SS	Safety Set
TEAE	Treatment Emergent Adverse Event
TTP	Time to Peak
WHO	World Health Organization

2. STUDY OBJECTIVES

This statistical analysis plan (SAP) is written with reference to and assumes familiarity with Protocol Amendment 6.0 dated 14th March 2024. as well as Protocol Amendment 4.0 dated 09 August 2021 and Protocol Amendment 5.0 dated 20th July 2022; the objectives below are from Protocol Amendment 6.0.

Objectives	Part A	Part B
Primary	- To evaluate the safety and tolerability of PP353 when administered by intradiscal injection - To characterise the PK profile of a single dose of PP353 in plasma to enable a choice of dosing regimen to be used in Part B to give approximately 8 days of intradiscal exposure to PP353 as inferred from plasma exposure	- To evaluate the safety and tolerability of PP353 when administered by intradiscal injection - To evaluate the efficacy of PP353 as measured by change from baseline in LBP NRS score at 12 months
Secondary		- To characterise the PK profile of PP353 in plasma for the dosing regimen selected in Part A to give approximately 8 days of intradiscal exposure to PP353 in the active arm as inferred from plasma exposure - To evaluate the efficacy of PP353 as measured by change from baseline in LBP score at 3, 6 and 9 months - To assess the efficacy of PP353 as measured by improvement in disability due to CLBP
Exploratory Part A and Part B	- To evaluate the relationships between drug exposure, safety, efficacy and subject reported outcomes - To explore imaging biomarkers related to quantitative assessment of inflammation and oedema in the spine - To characterise the PK profile of selected excipients from dose(s) of PP353 in plasma (analysis will depend on the stability of the	

	PK samples already collected and the consent of the subjects if consent has not already been obtained for such analysis) To investigate if systemic (blood) cytokine levels measured at baseline is related to response to treatment	
Exploratory Part A	- To evaluate the efficacy of PP353 as measured by change from baseline in LBP NRS score at 3, 6, 9 & 12 months - To assess the efficacy of PP353 as measured by improvement in disability due to CLBP	
Exploratory Part B		- To evaluate the efficacy of PP353 as measured by change from baseline in average LBP daily score over a 7-day period at 3, 6, 9 & 12 months

3. STUDY DESIGN

3.1 General Study Design and Plan

This is a two-part study, each of 12 months duration, in subjects with chronic low back pain (CLBP) associated with vertebral body endplate bone oedema as detected by MRI. In both parts of the study, the subject population will not have been previously treated with antibiotic agents for their back pain.

Part A of the study will be open label to evaluate the safety, tolerability and pharmacokinetics of PP353 in the above subjects. Part A of the study will be completed before Part B can begin and will be used to determine the optimum dosing regimen for Part B of the study.

Part B of the study is a randomised, placebo-controlled, investigator and subject-blinded, third-party unblinded study to evaluate the safety, tolerability and efficacy of two administrations of PP353 given approximately 4 days apart in the above subjects.

<u>Part A</u>

Part A will be open label including up to 6 subjects to study safety and tolerability and the pharmacokinetics (PK) of PP353 after a single intradiscal injection is administered. The first 3 subjects will receive a single injection of up to 3 mL of PP353 delivered directly into the intervertebral disc under contrast guidance. Systemic (blood) sampling will be performed over the next 2 weeks to establish the PK of PP353. The subjects will be dosed sequentially with a gap of at least 14 days between each subject being injected. Following a review of the safety and PK data for the first 3 subjects a decision will be made by the Safety Review

Committee (SRC) as to whether the PK data are consistent enough to determine a choice of injection regimen for Part B. If the PK data are not consistent (if the variability is too high to provide a reliable PK profile and there is not enough information to confirm the PK sample timepoints in Part B) then a further three subjects can be enrolled into Part A.

The second set of 3 subjects can be dosed sequentially or in parallel depending on the recruitment rate. The timing of the PK samples for the second set of three subjects may be altered following analysis of the first three subjects in order to optimise PK understanding. Drop-outs in Part A will be replaced if an adequate PK profile is not obtained.
Subjects will be followed for 12 months post dose to evaluate the safety, and tolerability of PP353. The clinical benefit of PP353 will also be evaluated as an exploratory endpoint in Part A of this study.

Part B

Part B is a randomised, investigator and subject blinded and third party (pharmacist and drug administrator) unblinded, placebo-controlled study of up to 40 subjects to study the safety, tolerability and efficacy of PP353 in treating subjects with CLBP associated with vertebral body endplate bone oedema detected on an MRI scan.

20 subjects on active
20 subjects on placebo

Part B consists of two arms, comparing active versus placebo for safety, tolerability and efficacy of PP353. Placebo will consist of sham injections.

The dosing regimen in Part B was determined by the PK data observed in Part A and will consist of two intradiscal injections providing approximately 8 days exposure to PP353 within the intervertebral disc, as inferred from plasma exposure. Subjects on the active arm will be treated with two intradiscal injections administered on Day 1 and then Day 5.

Subjects on placebo will receive two sham injections, with the skin pierced and the needle penetrating the deep fascia, stopping 2 to 3 cm short of the intervertebral disc.

Subjects will be followed for 12 months post dose to evaluate the safety, tolerability and efficacy of PP353.

As more than one intradiscal injection is required in Part B, a sentinel pair (one subject on active and one subject on placebo) will be dosed initially, followed by the rest of the cohort no sooner than 14 days after the last injection of the sentinel pair.

3.2 Randomisation and Blinding

Subjects selected for participation in Part A of the study must meet similar in- and exclusion criteria as subjects in Part B and therefore are representative of the population being studied.

Part A of the study will be open label and all subjects will receive active PP353.

Part B of the study is a placebo-controlled, investigator and subject-blinded, third-party (injector and pharmacy) unblinded because of the impracticality of blinding PP353 for injection due to differences between the active and placebo injections.

A sponsor blinding plan defines the roles and responsibilities of the study team and the procedure(s) to be followed by blinded and unblinded personnel throughout the Persica 002 study to ensure that the blind is maintained at all times. The sponsor blinding plan applies to Sponsor staff and contracted vendors. Each participating site will be required to draw up a site specific blinding plan prior to the study start to document the processes to be followed at their site to ensure that the blinded site staff members were not unblinded to treatment assignment.

The Contract Research Organisation (CRO), site study personnel and Sponsor study team personnel (with specified exceptions for safety reviews and early read out) will not be aware of which treatment (active study drug or placebo) the subjects will be given; full details will be provided in the Sponsor Blinding Plan. Subjects will not be aware of which treatment they had been administered.

Allocation of treatment will be performed by Interactive Response Technology (IRT) which will allocate a randomisation number and treatment allocation to each subject when they have passed the selection criteria.

Unblinding of the individual subject's treatment by the investigator is limited to medical emergencies or urgent clinical situations in which knowledge of the subject's study treatment was necessary for clinical management. Once a subject's treatment assignment had been unblinded, the medical monitor and study coordinator were to be notified within 24 hours. Information relating to unblinding (e.g. reason and date) will be clearly recorded in the subject's study file.

Persica, the SRC or designee will be permitted to unblind any SAE reports that were unexpected, and considered to be related to the study drug, in accordance with safety reporting guidance and regulations.

3.3 Sample Size Estimation

Part A: No formal power calculation was performed for Part A. Three subjects are likely to be sufficient to estimate the optimum dosage regime for the main trial but if the results are not clear then an additional three subjects could be recruited.

Part B: Power calculations were based on the results of the large CLBP randomised control trial (Albert *et al*, 2013b). That study's database was used to select the subset of subjects with baseline RMDQ score of ≥ 9 and baseline LBP NRS score of ≥ 4, and higher than the leg pain NRS score to fit the inclusion criteria. The subjects receiving antibiotics had a mean reduction from baseline in the LBP NRS score at 1 year of 3.28 units and placebo subjects had a reduction of 0.61 units. A conservative reduction of 1.2 units for the placebo group, as estimated from the AIM study (Bråten et al., 2019) was used rather than a reduction of 0.61 units in the final sample size calculation.

Sample size of 20 subjects per arm (taking into account 20% for withdrawals) is considered sufficient to detect a difference in CLBP reduction, compared with baseline, between the active and placebo groups of at least 2.08 units, assuming a common standard deviation of 2.3. A one-sided, two-group comparison with 5% type I error using a t-test provides power greater than 80%.

3.4 Study Schedules and Assessments

Screening and Enrolment/Randomisation: Subjects could be screened for eligibility up to 56 days prior to enrolment/randomisation (Day -56 to Day -1) under Protocol Amendment 5.0 and 6.0, and 28 days prior to enrolment/randomization under Protocol Amendment 4.0 (Day -28 to Day -1). For a full list of inclusion and exclusion criteria see the protocol. Day 1 will be the day of randomisation.

Intervention period:
Part A: On day 1, subjects will be admitted to hospital for the intradiscal injection and will remain in hospital overnight following the procedure. Subjects will return to the study site on Days 5, 7, 9 and 11 for safety assessments and collection of PK samples.
Part B: In Part B, subjects will be admitted to the hospital for the first intradiscal injection and will remain in hospital for at least 8 hours following the injection administration procedure. Subjects will return to the study site for one further injection on Day 5 and will remain in hospital for at least 8 hours following the second injection administration procedure. Under Protocol Amendment 5.0 and 6.0, the first 12 randomized subjects will return to the study site on Days 3, 5, 7, 9 and 11 for safety assessments and collection of PK samples. ;The remaining randomized subjects in Part B will have the same visit structure as the first 12 subjects randomized, but Days 3, 7 and 11 will be phone calls. Under Protocol Amendment 4.0 all randomized subjects will return to the study site on Days 3, 5, 7, 9 and 11 for safety assessments and collection of PK samples.

Follow up period:

In both parts A and B, all subjects will be followed until Month 12 (Day 365) to evaluate the safety, tolerability and clinical benefit of PP353 (evaluation of clinical benefit in Part A is exploratory only). Following the intervention period, subjects will have follow-up visits at Months 1, 3, 6, 9 and 12.

Subjects who discontinue the study early for whatever reason will be encouraged to complete all the day 365 (month 12) visit assessments as an Early Termination / End of Study visit.

Subjects who are unable to return to the study centre to attend follow-up visits are to be contacted by phone to: 1) determine health status, 2) record adverse event information, and 3) ask that the diary, LBP NRS scores, Oswestry and RMDQ-23 questionnaires be completed.

The end of the study is defined as the first day when all subjects have had a study completion visit, early termination visit or have otherwise been discontinued from the study.

The total duration for Part A and Part B is a maximum of 13 and 14 (if enrolled under Protocol Amendment 5.0 or Protocol Amendment 6.0, 13 if enrolled under Protocol Amendment 4.0) months respectively, including screening, enrollment, treatment, and follow-up.

4. GENERAL STATISTICAL CONSIDERATIONS

4.1 General Design of Report

All tables, listings and graphs for the final analysis will be produced by Micron Research Ltd. (Ely, UK) using SAS® Version 9.4 (SAS Institute, Cary, NC 27513). All tables, listings and graphs for the Early Readout will be produced by Veramed Ltd. (Twickenham, UK) using SAS® Version 9.4 (SAS Institute, Cary, NC 27513). Specifications for tables, listings and graphs can be found in the Tables Manual Version 6.0, 31st May 2024 for this study. All data collected on the eCRF will be listed by subject number. All subjects who gave informed consent will be included in the listings but for those who withdrew before treatment only the reason for withdrawal will be recorded.

The Part A subjects will not be summarised in the efficacy tables. The columns shown for Part A in the table shells will be omitted from the outputs.

All treated subjects will be included in the listings. The randomised subjects in Part B will have the relevant treatment group included in the listing while those from Part A will have "Part A" in place of the treatment group.

Unless otherwise stated, all summary tables will be presented by treatment group using the following column headings: PP353 (only Part B subjects are included in this treatment group), Placebo and Part A (non-effiacy tables only); a footnote will be added to detail that the PP353 column only includes Part B subjects. If possible, footnotes will only be placed at the end of tables which scroll over several pages; however, this may not be under program control. If there are too many footnotes to print on an output, a footnote cover page may be used instead.

Categorical variables will be summarized using frequencies and percentages. Unless otherwise stated, for efficacy outputs, the denominator for percentage calculations will be the total number of subjects with available data at each visit (where applicable) in the respective treatment group and for safety outputs, the denominator for percentage calculations will be the total number of subjects in the respective treatment group. Percentages will generally be presented to 1 decimal place but two decimal places may be used for very small numbers. A "Missing" category will only be included on the categorical summaries if the Investigator actually recorded "Missing" or "Not Done" as an outcome.

Continuous variables will be summarized using descriptive statistics, number of subjects with an observation (n), mean, standard deviation (SD), median, minimum (min) and maximum (max). Unless otherwise specified, all continuous variables will be given to one decimal place. SD will be presented to one more decimal place than the mean and median.

Apart from the Primary Analysis, which is based on a one-sided test, hypothesis tests will be two-tailed unless specified otherwise with a p-value of ≤0.05 taken as statistically significant and confidence intervals will be at the 95% level.

In general, where partial dates are recorded on the eCRF (missing day or missing day and month) and where these cannot be resolved by queries, dates will be estimated for the purpose of calculating durations. Where a date is partial, the first day of the month will be assumed if the day is missing and January will be assumed if the month is missing. If only the day is missing and the month and year is the same as the month and year of first medication administration, then the day will be assumed to be the medication start date. Partial concomitant medication and AE dates may need to be dealt with on a case by case basis.

4.2 Examination of Subgroups

No subgroup analyses are planned. Any post hoc subset analyses carried out to clarify the main results are to be regarded as purely exploratory.

4.3 Handling of Dropouts or Missing Data

Last Observation Carried Forward (LOCF) will be used for missing data points in the logistic regression analysis of Clinically Relevant Improvement in RMDQ-23, the binarized version of LBP NRS specified in Sections 6.5.1 and 6.5.3 and the binarized version of 7 day average LBP NRS daily score specified in Section 6.5.4 since missing data could skew this particular comparison, with Non-Responder Imputation included as a potential sensitivity analysis. Otherwise, unless stated otherwise, no imputation of data is planned.

4.4 Early Readouts and Data Monitoring

In Part B, there may be up to two Sponsor-unblinded early readouts to review summary results for internal decision making / corporate strategy development. It is likely that the first will be conducted after all subjects have completed the month 6 visit and the second, if required, after all subjects have completed the month 9 visit. This will not alter the conduct of the Persica 002 study. Early readout analyses will include data up until the timepoint that all subjects have completed. For example, for the early readout that will be conducted after all subjects have completed month 6, any data beyond month 6 will not be included in the analysis and reporting; output titles should be adjusted as appropriate if the output is truncated to an early readout timepoint but the title has further timepoints specified. Analyses will be conducted to plan for future studies and for confidential investor discussions important to the future of Persica Pharmaceuticals. Both analyses will be a subset of tables and listings sufficient to provide an early indication of the efficacy of PP353. The following tables and listings are planned to be prepared, by the unblinded programmer and statistician at Veramed, up to the respective early readout endpoints:

- LBP NRS Scores and change from baseline
- 7-day Average LBP NRS daily Scores and change from baseline
- RMDQ-23 Scores and change from baseline
- MMRM analysis may be performed on the change from baseline scores
- Global Perceived Effect response
    - LBP NRS stratified by global perceived effect response

    - LBP NRS Improvement, using 30% cut-off and cut-offs derived from LBP NRS stratified by global perceived effect response summaries
- Individual data, excluding subject identification numbers, may be looked at by parties independent of the project team should there be a corporate business requirement to do so.
- At sponsor discretion, MRI data which is available at the month 6 endpoint. The findings of the scans will be listed and may be summarised as follows: Shift tables comparing the grades of 0-1 versus 2-4 for baseline versus each post-baseline visit with a McNemar test p-value for change in the grades from Baseline to the visit. Change from Baseline Analysis for Dixon volume Sequences, DEMRIQ, DEMRIQ-ME and DEMRIQ-IRE.
- Adverse events – TEAE Overview and Summary, Serious TEAEs Summary, TEAEs regarded as Treatment Related, TEAEs and Serious TEAEs by treatment relation and maximum severity

The tables, figures and listings expected to be produced are as follows, though further tables, figures and listings may be produced at sponsor discretion, including repeats for the Per Protocol Set (PPS). MRI analyses are to be produced at sponsor discretion:

Tables

Table 3.1 Low Back Pain NRS Scores – Observed Results; Full Analysis Set

Table 3.2 Change from Baseline in LBP NRS Scores; Full Analysis Set

Table 3.4 Secondary Efficacy Analysis Results - Change in Low Back Pain NRS Score at Months 1, 3, 6 and 9; Full Analysis Set

Table 3.6 RMDQ-23 Total Score – Observed Results and Change from Baseline; Full Analysis Set

Table 3.7 Secondary Efficacy Analysis, Change from Baseline in RDMQ-23 Scores; Full Analysis Set

Table 3.11.2 Statistical Analysis of At Least 30% Improvement in LBP NRS; Full Analysis Set

Table 3.11.3 Statistical Analysis of Improvement in LBP NRS using the cut-off determined from Global Perceived Effect responders (Definition 1) at the timepoint; Full Analysis Set

Table 3.11.4 Statistical Analysis of Improvement in LBP NRS using the cut-off determined from Global Perceived Effect responders (Definition 2) at the timepoint; Full Analysis Set

Table 3.19 Global Perceived Effect of Treatment; Full Analysis Set

Table 3.20 Number of Subjects Using Analgesic Medication in the 7 Days Prior to Each Visit; Full Analysis Set

Table 3.26 Osteophytes, Endplate Defects and Modic Changes at Baseline; Full Analysis Set

Table 3.27 DCE MRI Performed at Baseline; Full Analysis Set

Table 3.30 Summary of the Nordic Modic Protocol Scores 0 and 1 v 2 to 4; Full Analysis Set

Table 3.32 Change Table of Nordic Modic Protocol Scores, 0 and 1 v 2 to 4; Full Analysis Set

Table 3.36 Results of MRI Scans, Total Volume of Lesions from Dixon Sequences; Full Analysis Set

Table 3.37.1 Results of MRI Scans, Total Volume of Inflammation (DEMRIQ Volume) by DCE-MRI; Full Analysis Set

Table 3.37.2 Results of MRI Scans, Total Volume of Inflammation (DEMRIQ-ME Volume) by DCE-MRI; Full Analysis Set

Table 3.37.3 Results of MRI Scans, Total Volume of Inflammation (DEMRIQ-IRE Volume) by DCE-MRI; Full Analysis Set

Table 3.38 7-Day Average Low Back Pain NRS Daily Scores – Observed Results; Full Analysis Set

Table 3.41 Change in 7-Day Average Low Back Pain NRS Daily Score at Months 3, 6 and 9; Full Analysis Set

Table 3.42.1 LBP NRS Observed Results stratified by Global Perceived Effect response (Definition 1); Full Analysis Set

4.5 Multi-center Studies

No test of homogeneity over Investigator centres is planned.

4.6 Multiple Comparisons / Multiplicity

No adjustment for multiple hypothesis tests is required.

4.7 Protocol Violations and Deviations

Subjects in the PPS must comply with the protocol in terms of adequate treatment, completion of the study and completion of efficacy assessments. Adequate treatment is defined as both injections being administered, with >2mls injected and <20% extravasation for each injection in the active group. Generally, subjects can be included or excluded from the PPS based on data recorded in the eCRF. Where the eCRF data are not adequate to define the subject, the final decision will be made during data evaluation meetings that are to be performed at sponsor discretion. The data evaluation meeting participants will be decided nearer to the time of the meeting but will not include any team members unblinded to individual treatment allocation. Inadequate treatment and major deviations for efficacy may exclude subjects from the PPS, though other protocol deviations may be deemed severe enough to lead to exclusion from PPS following data evaluation meeting.

Any eligibility criteria exceptions will be listed and protocol deviations will be listed and summarised by treatment group.

4.8 Definition of Baseline Value

For all assessments, the Baseline value is defined as the last available value prior to study drug administration.

4.9 Definition of Relative Day

Relative day for an event will be derived with the date of the first administration of study drug as reference.

Relative days for an event of measurement occurring before the date of first administration are calculated as follows:

Relative Day = [(Event Date – Date of First Administration)]

The relative day for an event or measurement occurring on or after the reference date is calculated as follows:

Relative Day = [(Event Date - Date of First Administration) + 1]

For events or measurements occurring before the date of the first administration, relative day will be prefixed with '-' in the data listings.

There is no relative Day 0. Relative day is not calculated for partial dates in cases where relative day is shown in a subject data listing. In such cases, relative day should be presented as '- -' in the subject data listings.

5. STUDY POPULATIONS

The Enrolled Set (ES) will consist of all subjects who give informed consent for Persica 002, including screen failures.

The Randomised Set (RS) is applicable to Part B only and will consist of all subjects who were randomised into the Persica 002 study.

The Safety Set (SS) will consist of all enrolled subjects who receive at least one dose of PP353 or placebo.

The Full Analysis Set (FAS) will consist of all enrolled subjects who receive at least one dose of PP353 or placebo and have a valid post-Baseline measurement for at least one efficacy variable.

The Per Protocol Set (PPS) is a subset of the FAS, consisting of those subjects who had no important protocol deviations that may affect the validity of the efficacy data. The following rules will be applied to identify subjects who may be excluded from the PPS:

- Incorrect administration of treatment
  - Incomplete dosing regimen
  - Adequate treatment as per SAP Section 4.7 not met
- Incorrectly entering the study
  - Failure to meet inclusion criteria / meeting exclusion criteria but still admitted
- Prohibited Medication Use
  - Any concomitant medication administered as described in Protocol Section 4.4.2.

Blinded data will be reviewed at data evaluation meetings (See section 4.7) prior to database lock to identify patients who should be excluded due to the above rules alongside other important protocol deviations. In some cases, patient data that occurs after the start date of the deviation may be excluded from analyses rather than excluding all the patient data. This will be considered on a case-by-case basis. The final list of patients to be excluded from the PP will be agreed at this meeting.

6. STUDY VARIABLES AND ANALYSIS

If warranted, selected efficacy outputs using the full analysis set (FAS), including exploratory and sensitivity analyses, may be repeated for the PPS at sponsor discretion.

6.1 Analysis and Reporting of Data from Part A of the Study

Data from Part A of the study are required for analysis before Part B of the study can begin. The PK analysis forms a separate report, parts of which are set out in section 8. Once the data from Part A have been cleaned they will be reported as simple listings of every subject's data. The first listings will be prepared for the Safety Review Committee once the initial three subjects have completed. If further subjects are recruited in Part A the listings will be repeated with the full number of subjects.

The listings will include (but not be limited to):

- Baseline and demographic data – age, gender, height, weight, BMI, LBP NRS score.
- Prior and concomitant medication.
- Medical history.

- Injection site reaction: Pain, tenderness, erythema and induration severity following the intradiscal injection
- Adverse events

All Part A data will be listed , and summarised (if applicable) at the end of Part B. Part A data will not be presented in the efficacy tables.

6.2 Primary Efficacy Variable and Analysis (Part B) / Exploratory Efficacy Variable and Analysis (Part A)

The Primary Efficacy Variable is the change from baseline in LBP NRS score at 12 months. The LBP NRS score is defined as the average score of the 3 questions:

1. Low back pain intensity now,
2. Worst low back pain intensity in the last 14 days,
3. Average low back pain intensity over the last 14 days.

Each question will be assessed by the subject on an 11-point scale with 0 = "no pain" and 10 = "the worst possible pain you can imagine." The average of the 3 scores will be used to summarise the subject's LBP NRS score.

The Primary Efficacy Analysis will be performed on the Full Analysis Set (FAS), which is the conservative approach in a placebo controlled study.

The statistical analysis will only be performed on the subjects in Part B of the trial. The primary efficacy endpoint will be analysed using a mixed model for repeated measures (MMRM) and will adjust for Baseline to test the difference in the two groups for the LBP NRS score change from baseline at 12 months.

For change from baseline for LBP NRS at 12 months, A 1-sided hypothesis test will be performed to test the primary hypothesis of superiority of PP353 versus placebo at 12 months, i.e. that the reduction of LBP NRS score at Month 12 relative to Baseline for PP353 is greater than the reduction of LBP NRS score at Month 12 relative to Baseline for placebo. The null hypothesis is that PP353 is not superior to Placebo at 12 months, i.e. that the reduction of LBP NRS score at Month 12 relative to Baseline for PP353 is not greater than the reduction of LBP NRS score at Month 12 relative to Baseline for placebo. A 2-sided hypothesis test will also be performed but this will be separate to the primary analysis as the study is powered on the basis of a 1-sided test.

The model will use all available post-Baseline LBP NRS score data at all visits up to and including 12 months and will be incorporated as repeated measures within each subject.

Treatment group (PP353 or placebo), visit, Baseline LBP NRS scores, and treatment group by visit interaction will be included as fixed effects. The multiple visits for each subject will be incorporated as repeated measures within each subject. An unstructured covariance matrix will be utilized unless the model does not converge, in which case alternative covariance structures will be investigated. Least Squares Means (LS Means) for changes from Baseline and 95% 1-sided confidence intervals (CI) for each treatment group will be presented. A 2-sided 95% CI will also be presented.

The Kenward-Roger method will be utilized for calculating denominator degrees of freedom, though alternative methods may be investigated and used instead as appropriate (e.g. if alternative covariance structures are used) – the approach will be footnoted.

Model checking output will also be produced for the LBP NRS MMRM model, as histograms and QQ-plots of the residuals for the overall model and also at Month 12 separately for the final analysis; non-parametric methods and transformations may be explored at sponsor discretion, following review of model appropriateness. If this analysis is repeated for other endpoints, then the model checking output as specified above will also be produced.

Missing data are not explicitly imputed in the primary MMRM analysis, although there is an underlying assumption that data are missing at random. All available scheduled post-Baseline assessments are utilized and, via modelling of the within subject correlation structure, the derived treatment differences at Month 12 are adjusted to take into account missing data. A mixed effects model for repeated measures analysis (using fixed effects and an unstructured variance covariance matrix) is considered appropriate as the primary method of analysis as it has been shown to give sensible answers to on-treatment questions in a range of practical situations (Siddiqui *et al.*, 2009). This analysis will be done with Proc Mixed SAS version 9.4 and the change from baseline in LBP NRS score at 12 months will be listed and summarised as continuous data.

The primary efficacy endpoint may also be analysed using Bayesian analysis with an informative prior for the placebo group, if appropriate. If Bayesian analysis is performed, details will be provided in a separate SAP.

6.3 Secondary Efficacy Variables (Part B) / Exploratory Efficacy Variables (Part A)

Note that Screening and Day 1 pre-dose will not be presented in summaries – for pre-dose assessments, only the baseline will be presented as defined in Section 4.8 and included as appropriate.

6.3.1. LBP NRS – Additional Timepoints

The observed values and change from Baseline values for the LBP NRS score will be summarised at every assessment point.

For Part A, subjects recorded LBP NRS scores on an electronic diary at screening, Day 1 pre-dose and on days 5, 30 (1 month), 60 (2 months), 90 (3 months), 120 (4 months), 150 (5 months), 180 (6 months), 210 (7 months), 240 (8 months), 300 (10 months), 330 (11 months). and 365 (12 Months) or Early Termination.

For Part B, follow-up visits are planned at days 30 (1 month), 90 (3 months), 180 (6 months), 270 (9 months), and 365 (12 months) or Early Termination, with the LBP NRS scores collected on either paper or electronic diaries. Assessment points which are follow-up visits and scheduled will be summarised with statistical analysis by repeated measures as for the Primary Efficacy Analysis. Statistical significance for the comparison between treatment groups and the 95% confidence intervals will be reported at months 1, 3, 6 and 9, using just a 2-sided hypothesis test at 95% level. These estimates will all come from the same MMRM,

which includes timepoints up to the relevant cut-off, for example up to Month 6 if looking at the early readout for Month 6 and Month 12 for the final analysis.
The sensitivity analyses specified in Section 6.5.1 may also be repeated for these additional timepoints; The further explanatory endpoint for 30% improvement will be repeated for these additional timepoints.

All LBP NRS observed values and changes from Baseline will be listed for the FAS.

6.3.2. Questionnaires

Three questionnaires will be recorded at baseline and at intervals during the study. These are the Roland-Morris Disability Questionnaire 23 (RMDQ-23), the RMDQ-24 and the Oswestry Disability Index (ODI) Questionnaire.

The RMDQ-23 is a self-administered disability measure. In addition to the 23 questions assessed in the RMDQ-23, a further 5 questions related to low back pain disability are used with a subset of the RMDQ-23 questions to derive the RMDQ-24 score, which is an exploratory endpoint. This score will be referred to in the clinical study report as the “derived RMDQ-24 score".

For Part A, the RMDQ-23 and Derived RMDQ-24 are completed at screening, day 1 pre-dose, days 30 (1 month), 60 (2 months), 90 (3 months), 180 (6 months), 270 (9 months), and 365 (12 months) or Early Termination. For Part B under Protocol Amendment 4.0 and later, these are completed at screening, day 1 pre-dose, days 30 (1 month), 90 (3 months), 180 (6 months), 270 (9 months), and 365 (12 months) or Early Termination.

The ODI is a subject-completed questionnaire which gives a subjective percentage score of level of function (disability) in activities of daily living in those rehabilitating from low back pain. The questionnaire examines the perceived level of disability in 10 everyday activities of daily living. The ODI is completed at Day 1 pre-dose, 3 months, 6 months and 12 months (or Early Termination).

Observed values and changes from baseline in all three questionnaires will be listed and summarised as continuous data at each of the respective assessment points. Statistical analysis will be by repeated measures as described for the analysis of LBP NRS scores. Statistical comparisons will be reported for months 3, 6, 9 and 12 for both the RMDQ-23 and the Derived RMDQ-24 and month 1 for RMDQ-23 only, and at months 3, 6, and 12 for the ODI, using just a 2-sided hypothesis test at 95% level. The FAS will be used for all three questionnaires. These estimates will all come from the same MMRM, which includes timepoints up to the relevant cut-off, for example up to Month 6 if looking at the early readout for Month 6 and Month 12 for the final analysis.
A clinically relevant improvement is defined as ≥30% improvement in the Roland Morris Disability Questionnaire-23 score. Subjects who record a score at 1, 3, 6, 9 and 12 months will be defined on this basis as clinically improved or not improved at each assessment point.

The number and percentage of subjects that achieve the clinically relevant improvement threshold for the RMDQ-23 in each treatment group will be summarised at each assessment point. In addition, the number and percentage of subjects that achieve ≥75% and/or ≥50% improvement from baseline in the RMDQ-23 in each treatment group will also be

summarised at each assessment point. Statistical comparison of the clinically relevant ≥30% improvement threshold of RMDQ-23 will be by logistic regression with explanatory variables including treatment group and baseline RMDQ-23 score. Missing data could skew this result if some subjects fail to attend one or more visits. For the purpose of the logistic regression analysis only, where patients have simply missed a visit (or visits) LOCF will be used to impute a result. The number needed to treat will also be presented alongside the logistic regression analyses. If there are convergence issues with the logistic regression, Firth regression methods will be employed; if this fails to resolve convergence issues, then no estimates will be presented. The approach will be footnoted if Firth regression methods had to be used.

6.4 Exploratory Efficacy Analyses – Parts A and B

6.4.1 Leg Pain NRS

Leg pain is recorded on an 11-point NRS scale where 0 = "no pain" and 10="the worst pain you can imagine". As with the LBP NRS there are three scores at each assessment, leg pain (sciatica) intensity now, worst leg pain (sciatica) intensity in the last 14 days and the average leg pain (sciatica) intensity in the last 14 days. These three scores will be averaged at each assessment to give an overall Leg Pain NRS Score.

The collection points for Leg Pain NRS are those for the LBP NRS.

The observed values and changes from Baseline in Leg Pain Scores will be listed and summarised at every assessment point and compared between treatment groups by a repeated measures analysis in the same way as the LBP NRS scores, using just a 2-sided hypothesis test at 95% level. These estimates will all come from the same MMRM, which includes timepoints up to the relevant cut-off, for example up to Month 6 if looking at the early readout for Month 6 and Month 12 for the final analysis.

6.4.2 Hours with Low Back Pain

Recorded at screening, Day 1 pre-dose and on days 30 (1 month), 60 (2 months, for Part A only),90 (3 months), 180 (6 months), 270 (9 months) and 365 (12 months) or Early Termination. Subjects record the average number of hours per day (over the last 28 days) that they experienced low back pain and also the number of days over the last 28 days they had experienced low back pain. Hours with low back pain over the last 28 days will be derived by multiplying the number of days with low back pain (over the last 28 days) with the average number of hours per day with low back pain (over the last 28 days). These assessments will be listed and observed values and changes from Baseline at each timepoint will be summarised as continuous data with no formal statistical comparison.

6.4.3 Global Perceived Effect

The Global Perceived Effect of treatment will be recorded at days 30 (1 month), 60 (2 months, for Part A only), 90 (3 months), 180 (6 months), 270 (9 months) and 365 (12 months) or Early Termination. This will be scored on a 7-point scale with the subject asked at each timepoint to rate the overall drug effect in relation to severity at baseline. To analyse the data the scores will be ordered as follows:

1. Much worse

2. Somewhat worse
3. A bit worse
4. Unchanged
5. A bit better
6. Somewhat better
7. A lot better.

At each assessment the scores will be listed and summarised as frequencies and percentages and also listed.

6.4.4 Subject Diary

Recorded at screening, Day 1 pre-dose and on days 30 (1 month), 60 (2 months, for Part A only), 90 (3 months), 180 (6 months), 270 (9 months) and 365 (12 months) or Early Termination.

The subject Diary will include information on analgesics taken during the study which will be listed for the SS. This will include the drug name and number of days the treatment was taken during the seven days before each study visit; This is recorded by the subject each day for 7 days prior to actual visit date for paper diaries and Day 1/365 for e dairies; for other e diary collections it is 7 days prior to visit date calculated from baseline. The drug names will be coded as per the WHO-DRUG version detailed in Section 9.6 and presented alongside frequencies and percentage of subjects who recorded using them since the previous visit and also listed. The frequency and percentage of subjects who recorded using Opioids, NSAIDs and Other medications of interest since the last visit will also be presented and listed; the medications will be identified by review of the patient reported medications in the dairy prior to any early readout and prior to final database lock.

The number of days on any analgesic since the previous visit will be presented for both groups as continuous data.

6.4.5 Derived RMDQ-24

The Derived RMDQ-24 is an exploratory efficacy variable for Part A and Part B. Further details on the analysis of this endpoint are in Section 6.3.2.

6.4.6 MRI Data

In Part A of the study an MRI and DCE-MRI scan to assess the volume and severity of vertebral body endplate oedema and inflammation will be performed at Screening, 6 Months and 12 Months. In Part B of the study an MRI and DCE-MRI were also performed at 6 Months (this was confirmed based on the Part A data) and 12 Months. The results of these scans will be added electronically to the database, not entered into the eCRF. The findings of the scans will be listed and summarised.

Nordic Modic Analysis

The Nordic Modic Protocol is a measure of the volume of bone oedema coded as follows: 0 = normal, 1 = located to the endplate only, 2 = less than 25% of vertebral body volume, 3 = 25–50% of vertebral body volume and 4 = More than 50% of vertebral body volume. At each assessment point there will be two Nordic Modic Protocol scores, one for the vertebra which

is above (upper vertebra) the affected disc and one for the vertebra which is below (lower vertebra) the affected disc. These scores will be reported separately and the most severe of the upper and lower will be reported as appropriate. The listings will make clear which is the affected disc and which vertebra is being scored. The summary tables detailed below will all be presented for both the upper and the lower vertebra and the most severe of the upper and lower as appropriate, either as a single table where space permits or as separate tables.

The Nordic Modic Protocol values will be listed and summarised as frequencies and percentages at each assessment and for both the upper and lower vertebra in relation to the affected disc; the most severe of the upper and lower vertebra will also be reported. Additional summary tables will present frequencies and percentages of subjects with grades 0-1 and 2-4 by visit and for both the upper and lower vertebra in relation to the affected disc; the most severe of the upper and lower vertebra will also be reported. Shift tables will compare the grades of 0-1 versus 2-4 for baseline versus each post-baseline visit with a McNemar test p-value for change in the grades from Baseline to the visit and will also compare individual Nordic Modic Protocol Scores for baseline versus each post-baseline visit and for both the upper and lower vertebra, as well as the most severe of the upper and lower.

McNemar p-values will be calculated with baseline grade as the reference (in the 2x2 table set up, these would be considered as the columns).

The volume of PP353 administered will be summarised by the Nordic Modic Protocol value, the most severe score being used if there is a difference between the upper and lower vertebra. Since there are multiple injections administered in Part B the summary table will be per injection. Subjects randomised to placebo will only receive sham injections.

The number of patients with osteophytes, localized endplate defects and irregular endplates at baseline will be summarised by vertebra and overall. Also, the number and frequency of patients with only Modic 1 changes, Modic 1 and 2 changes, as well as those that additionally have Modic 3 changes at baseline will be summarised by vertebra and overall. This information will also be listed.

A table of RMDQ23 response/non-response at Month 12 by most severe Nordic Modic Protocol Score category at Baseline will be produced.

DEMRIQ and DIXON Analysis

The volume of each lesion will be determined from Dixon sequences (volume of the entire lesion) and the quantitative volume (DEMRIQ volume or volume of pixels in the lesion that are inflamed) determined from the DCE-MRI sequences. Both volume measures will be analysed separately as follows: The total lesion volume per subject will be calculated by summing the volume of all lesions within a vertebra. This volume and the changes from baseline will be listed and summarised as continuous data at each assessment point as a total volume across the upper and lower vertebra in relation to the affected disc. The change from baseline will be compared between treatment groups by analysis of covariance with the baseline value as the covariate. The change from baseline analysis will be repeated using the largest lesion at baseline per vertebrae and the largest lesion at baseline across both vertebrae.

The change from baseline analysis will also be repeated for DEMRIQ-ME, for the total volume of DEMRIQ-ME across both vertebrae, the largest lesions DEMRIQ-ME volume at

baseline per vertebrae and the largest lesions DEMRIQ-ME at baseline across both vertebrae; this analysis will be repeated for the DEMRIQ-IRE also. The DEMRIQ-ME and DEMRIQ-IRE volumes and changes from baseline will also be listed.

The number of lesions by visit, as well as the number of new lesions at months six and 12 in comparison to the Baseline visit will be summarised across vertebra above and below the affected disc.

Whether patients had DCE-MRI or not at baseline will be summarised.

Due to the exploratory nature of this analysis the first real subject data from the MRI scans will arrive when the Group A subjects reach month 6. There may be up to six subjects at the Month 6 visit in Part A. At this point IAG will prepare a report on 'Utility of DCE-MRI' and based on this, further exploratory analysis may be defined and performed. Any further exploratory analysis of MRI data will be defined in a separate data analysis plan, and may include, but will not be limited to, analysis of a primary lesion, as well as analysis of the DEMRIQ-ME mean, the DEMRIQ-IRE mean and the time to peak. A decision will be also be taken whether DCE-MRI should continue for Part B subjects at 6 months.

All exploratory efficacy outputs will be produced for the FAS.

6.4.7 Baseline Cytokines

If sufficient data are available, correlation plots of sponsor selected cytokines at baseline and Month 12 results of efficacy parameters of interest may be produced for the FAS.

Cut-off points for elevation may be identified and sponsor selected descriptive tables subgrouped by non-presence of elevation or presence of elevation; no inferential statistics will be produced.

6.5 Exploratory Efficacy Analyses – Part B only

6.5.1 LBP NRS

The proposed MMRM analysis may be repeated for a sensitivity analysis using e-diaries only. The proposed MMRM analysis may also be repeated for a sensitivity analysis using Pain Catastrophising Score (PCS) at Baseline (as defined in section 4.8) as a covariate in the model.

As a further exploratory analysis, the primary efficacy endpoint will also be assessed in a binary form, whether the subject has had ≥30% improvement in LBP NRS at Months 1,3, 6, 9 and 12. This will be assessed using logistic regression, with treatment group and baseline LBP NRS as covariates; the number and percentage of subjects that achieve ≥30% improvement in LBP NRS at Months 1, 3, 6, 9 and 12 will also be summarised. For the purpose of the logistic regression analysis only, where patients have simply missed a visit (or visits) LOCF will be used to impute a result. The number needed to treat will also be presented alongside the logistic regression analyses. If there are convergence issues with the

logistic regression, Firth regression methods will be employed; if this fails to resolve convergence issues, then no estimates will be presented. The approach will be footnoted if Firth regression methods had to be used.
For this 30% improvement exploratory endpoint, for the logistic regression only, a sensitivity analysis may be provided where the method for missing data imputation is changed from LOCF to Non-Responder Imputation, whereby if a subject does not have a valid assessment for analysis at a timepoint, they will be treated as not having improved for the relevant threshold at the timepoint (i.e. considered as non-response).

The above analysis will be repeated but using LBP NRS ≤ 2 as the response. In addition, the primary efficacy endpoint will also be assessed in a binary form, whether the subject has had ≥50% improvement in LBP NRS at Months 1,3, 6, 9 and 12.

The analyses specified for the 30% improvement above will also be repeated, using the mean percent change from baseline value of LBP NRS, taken for all subjects at Months 1, 3, 6, 9 and 12 who have a global perceived effect score of "Somewhat better" or "A lot better". The same analyses will also be performed, using the mean percent change from baseline value of LBP NRS, taken for all subjects at Month 1, 3, 6, 9 and 12 who have a global perceived effect score of "A lot better".

A correlation plot will also be produced for LBP NRS Change from Baseline and Percent Change from baseline and global perceived effect at Months 6, 9 and 12.
A table stratified by global perceived effect responder status will be produced for the percent change from baseline and change from baseline of LBP NRS at all scheduled assessment points, for each categorisation of global perceived effect response; "Somewhat better" or "A lot better", or "A lot better" only.

6.5.2 RMDQ-23

For logistic regression only, a sensitivity analysis may be provided where the method for missing data imputation is changed from LOCF to Non-Responder Imputation, whereby if a subject does not have a valid assessment for analysis at a timepoint, they will be treated as not having improved for the relevant threshold at the timepoint (i.e. considered as non-response).

6.5.3 Global Perceived Effect

For part B only, as a further exploratory analysis, global perceived effect will also be assessed in a binary form, using the following two response definitions

1. A responder is defined as having a global perceived effect score of 6 or more ("Somewhat better" or "A lot better")
2. A responder is defined as having a global perceived effect score of 7 ("A lot better").

The number and percentage of subjects that achieve response will be summarised for each of these response definitions and will be analysed using logistic regression, with treatment group as an explanatory variable. For the purpose of the logistic regression analysis only, where patients have simply missed a visit (or visits) LOCF will be used to impute a result.

The number needed to treat will also be presented alongside the logistic regression analyses. If there are convergence issues with the logistic regression, Firth regression methods will be employed; if this fails to resolve convergence issues, then no estimates will be presented. The approach will be footnoted if Firth regression methods had to be used.

6.5.4 7-day average LBP NRS daily score

Recorded at Day 1 pre-dose and on days 30 (1 month), 90 (3 months), 180 (6 months), 270 (9 months) and 365 (12 months) or Early Termination. Each daily score is recorded by the subject each day for 7 days prior to actual visit date for paper diaries and Day 1/365 for e dairies; for other e diary collections it is 7 days prior to visit date calculated from baseline. Each daily score is recorded on an 11-point NRS scale where 0 = “no pain” and 10=“the worst pain you can imagine”. If at least 4 daily scores are collected then the daily scores will be averaged to give the 7-day average LBP NRS daily score.

The summaries and listings provided for the 14-day recall LBP NRS will be repeated for the 7-day average LBP NRS daily score, using just a 2-sided hypothesis test at 95% level for repeated measures analysis; all outputs will be produced for the FAS. Any estimates coming from MMRM will all come from the same MMRM, which includes timepoints up to the relevant cut-off, for example up to Month 6 if looking at the early readout for Month 6 and Month 12 for the final analysis.

A sensitivity analysis may be performed to only include assessments where all 7 days were filled in.

As a further exploratory analysis, the 7-day average LBP NRS daily score will also be assessed in a binary form with response as 7-day average LBP NRS daily score ≤ 2 at Months 3, 6, 9 and 12. This will be assessed using logistic regression, with treatment group and baseline 7-day average LBP NRS daily score as covariates; the number and percentage of subjects that achieve 7-day average LBP NRS daily score ≤ 2 at Months 3, 6, 9 and 12 will also be summarised. For the purpose of the logistic regression analysis only, where patients have simply missed a visit (or visits) LOCF will be used to impute a result. The number needed to treat will also be presented alongside the logistic regression analyses. If there are convergence issues with the logistic regression, Firth regression methods will be employed; if this fails to resolve convergence issues, then no estimates will be presented. The approach will be footnoted if Firth regression methods had to be used.

For this 7-day average LBP NRS daily score ≤ 2 exploratory endpoint, for the logistic regression only, a sensitivity analysis may be provided where the method for missing data imputation is changed from LOCF to Non-Responder Imputation, whereby if a subject does not have a valid assessment for analysis at a timepoint, they will be treated as not having improved for the relevant threshold at the timepoint (i.e. considered as non-response).

6.5.5 MRI Data

Shift tables for the Nordic Modic Analysis may be repeated for a sensitivity analysis using e-diaries only.For the actual Nordic Protocol Scores (as opposed to change from baseline) a sensitivity analysis may be performed based on the grade groupings used in Albert et al

(2013b) in the event that a subject may be allocated to a different grade based on those used in Albert et al (2013b).

These summary tables will all be presented for both the upper and lower vertebra in relation to the affected disc; the most severe of the upper and lower vertebra will also be reported.

7. SAFETY VARIABLES

The investigator will evaluate safety by adverse events monitoring (including injection site reactions), vital signs, physical examination, ECG and clinical laboratory tests.

All listings and summary tables of safety data will be presented for the Safety Set.
Note that Screening and Day 1 pre-dose will not be presented in summaries – for pre-dose assessments, only the baseline will be presented as defined in Section 4.8 and included as appropriate.

7.1 Adverse Events

A treatment emergent adverse event (TEAE) is defined as any event not present prior to the initiation of the treatments or any event already present that worsens in either intensity or frequency following exposure to the treatments. Any adverse events which are reported after the signing of the informed consent to the time of administration of PP353 are pre-treatment adverse events and will be listed with TEAEs but will not be summarised. In the listings of adverse events each event will be flagged either Yes (TEAE) or No (pre-treatment). Separate listings will be produced for SAEs, AEs leading to withdrawal and AEs leading to death.

Subjects will be assessed for the occurrence of adverse events throughout the study. Adverse events will be assessed at each study visit up to 12 months or at the final visit in the case of subjects who are withdrawn earlier. All adverse events encountered during the study will be reported in the eCRF.

AEs will be coded using the most recent MedDRA Dictionary. Coding will include the system organ class (SOC) and preferred term (PT). All verbatim descriptions and coded terms will be listed for the AEs. One listing will be produced for all AEs (serious and not serious) and a separate listing will also be produced for SAEs only. Adverse events leading to withdrawal and AEs leading to death will also be listed separately.

An overall summary of the number and percentage of subjects with TEAEs will be presented, including the number and percentage of subjects with any TEAE, with related TEAEs, with serious TEAEs, with serious related TEAEs, with TEAEs leading to withdrawal or death and related TEAEs leading to death.

All TEAEs will be summarized by SOC and PT. For each subject, multiple occurrences of the same event will be counted only once within a SOC and PT (i.e. subject events). The denominator for percentages will be the total population of the respective treatment groups. Summaries will be presented by decreasing frequency of subjects with each event in the PP353 treatment group.

The same summaries of subject events by SOC and PT will be repeated for related TEAEs, serious TEAEs, TEAEs by relationship, serious TEAEs by relationship, fatal TEAEs by relationship, TEAEs leading to withdrawal, and non-serious TEAEs occurring in more than 2 subjects. A summary of TEAEs by maximum severity (mild, moderate or severe), SOC and PT will also be presented, and repeated for serious TEAEs. For each subject, multiple occurrences of the same event will be counted only once at their maximum severity within a SOC and PT.

7.2 Vital Signs

Systolic and diastolic blood pressure, heart rate, body temperature and respiratory rate will be assessed at screening, day 1 pre-dose, every 30 minutes post injection for the first two hours and at subsequent visit days – collections may not be taken if not clinically indicated as per Schedule of Activities (Appendix A in the protocols) . The values at each assessment point and the change from Baseline will be listed and summarised as continuous data for the safety set, with no statistical comparison between the treatment groups.

7.3 Physical Examination

A complete physical examination will be conducted at screening and day 1 pre-dose (including weight at both timepoints and height only at screening). Height and weight will also be collected at Month 12. (or early Termination) An abbreviated physical examination will be performed at some post-Baseline timepoints as per Appendix A in the protocols.

Only abnormal findings will be listed, and summarised for the SS as frequencies and percentages by treatment group. Height and Weight will be listed only.

7.4 ECG

A 12-lead ECG is to be obtained at screening and at specified timepoints post treatment. The following will be recorded in the eCRF: heart rate, RR interval, QRS duration, PR interval, uncorrected QT interval, and QT interval corrected based on the machine readings of the individual ECG tracings. Values and change from Baseline for these parameters will be listed and summarised for the SS as continuous data with no statistical comparison between treatment groups. One listing will show all data including whether values were outside the normal range and whether clinically significant. A second listing will show the same information but only for values which were outside the normal range. If a subject has any values that are outside of the normal range for a given variable, then all values will be shown for that subject for that variable.

The number and percentage of subjects with values that are outside the normal range will be summarised as frequencies and percentages, at each assessment point as per Appendix A in the protocols. If a variable has any values that are outside of the normal range, then all assessment points for that subject and that variable will be displayed in the table. The normal ranges will be heart rate: 40 - 120 bpm, RR interval: 600 – 1200 ms, QRS duration: 80 - 120ms, PR interval: 120 – 200 ms, and QTcB interval: ≤480ms regardless of gender.

7.5 Laboratory Tests

Blood and urine (dipstick) samples will be collected for laboratory testing of clinical chemistry, haematology and urinalysis. All data observed results and changes from Baseline will be listed and summarised as continuous data for assessments outlined as per Appendix A in the protocols. A second listing will show the same information but only for values which were outside the normal range. If a subject has any values that are outside of the normal range for a given variable, then all values will be shown for that subject for that variable.

Since multiple laboratories are used the values will be standardised as SI units.

Tables will show the shift from baseline to each assessment point as low, normal, high missing and total. The normal ranges will be those set by the individual laboratories.

Any positive urine pregnancy test results will be listed for the SS.

8. PHARMACOKINETIC VARIABLES AND ANALYSIS

The following section is an extract from the full PK analysis plan: "Final PK Plan (Persica) v1.0 17 April 2019". If the PK analysis plan is updated, the following sections will only be updated if there is an impact on the information below.

In Part B, under Protocol Amendment 5.0 and 6.0, only the first 12 randomized subjects will be required to undergo full PK testing. Under Protocol Amendment 4.0, all subjects will be required to undergo full PK testing.

Plasma concentration data will be provided by the Sponsor in Excel format. Actual blood sampling times relative to the administered dose (consolidated with the plasma concentration data, demographic data and randomisation schedules) will be provided by the Sponsor.

8.1 Pharmacokinetic parameter estimation

Definition and estimation of PK parameters are described in KinetAssist's SOPs. PK parameters will be estimated for each subject using a fully validated version of WinNonlin Phoenix (Version 8.1or higher). The following parameters (but not limited to) will be derived, where appropriate, from the individual plasma concentration versus time profiles of PP353.

Parameter	Definition
C_{max}	The maximum observed concentration.
t_{max}	The time at which C_{max} was apparent.

AUC_{0-t}	The area under the concentration versus time curve from time zero to the sampling time at the last quantifiable concentration (C_t) at t_{last} (the time of the last quantifiable concentration) calculated by the mixed linear/log trapezoidal rule.
$AUC_{0-\tau}$	The area under the concentration versus time curve within a dosing interval calculated by the mixed linear/log trapezoidal rule (Part B only).
λ_z	The apparent terminal rate constant, estimated using the negative slope of the least square regression analysis of the log concentration versus time data for the terminal linear portion of the curve (Part A).
$t_{½}$	The apparent terminal half-life, calculated from $Log_e 2 / \lambda_z$ (Part A).
$AUC_{0-\infty}$	The area under the concentration-time curve estimated from time zero to infinity as the sum of the two areas: AUC_{0-t} and AUC_{extrap}, where AUC_{extrap} is calculated as C_t / λ_z (Part A only).
CL/F	The systemic clearance calculated from plasma calculated as: Dose / $AUC_{0-\infty}$ (Part A) or Dose / $AUC_{0-\tau}$ (Part B), where $AUC_{0-\tau}$ is estimated after the last dose.
Tdur	Duration above a prescribed threshold (5 ng/mL).

Consideration will be given to the estimation of λ_z and corresponding $t_{½}$ values. Three or more points are required within the terminal phase for λ_z and $t_{½}$ to be estimated. The following additional variables will be tabulated to aid identification of potentially unreliable estimates of $t_{½}$ and $AUC_{0-\infty}$:

Number data points	The number of data points used in the calculation of λ_z.
λ_z lower	The lower limit on time for values included in the calculation of λ_z.
λ_z higher	The upper limit on time for values included in the calculation of λ_z.
λ_z period	Estimated as (λ_z higher - λ_z lower) / $t_{½}$. Values < 2 will indicate that λ_z and corresponding $t_{½}$ estimates are potentially unreliable.

Rsq	the square of the correlation coefficient for the apparent terminal phase regression line, adjusted for the number of data points.
AUC_{extrap}	The extrapolated area from t_{last} to infinity estimated as C_t/ λ_z.

Actual sampling times will be used for the PK analysis, excluding interim reports. Where the actual sampling time is not recorded, the nominal sampling time will be used.

Plasma concentrations below the limit of quantification of the assay (BLQ), will be taken as zero. All calculations will be made using raw (unrounded) data.

8.2 Pharmacokinetic analysis

Plasma concentration data will be summarised by sampling time, sampling day and dose level, as appropriate; PK parameters will be summarised by, sampling day and dose level.

All individual plasma and PK parameter estimates will be listed and summarised. Mean and individual plasma concentration versus time profiles will be illustrated using both linear-linear and logarithmic-linear scales.

8.3 Statistical Analysis

Summary statistics will include number of subjects (N), number of observations (n), arithmetic mean, SD and coefficient of variation. Summaries for plasma concentrations will also include n(LLOQ), the number of samples <LLOQ.

Summaries for the PK parameters will also display the median, minimum and maximum. In addition, with the exception of t_{max}, the geometric mean and geometric coefficient of variation (CV)

where $CV = \sqrt{\exp(SD_{\ln}{}^2) - 1} * 100$ ($SD_{\ln}$ is the standard deviation of the natural logarithmically transformed data) will be reported for all PK parameters.

Between-subject variability will be based on geometric mean CVs.

Any further exploratory analysis of the relationship between exposure and efficacy will be defined following a review of the PK data.

9. SUMMARY OF STUDY POPULATION DATA

Data listings will include all subjects. Only the reason for discontinuation will be listed for subjects classed as screening failures (i.e. a subject who had signed informed consent and had failed any inclusion or exclusion criteria or withdrew before any treatment had started). All

listings will include the subject identifier and the study populations the subject was included in.

9.1 Subject Disposition

All subject disposition information described below will be listed for the ES. The number and percentage of subjects in each analysis population (defined under Study Populations) will be summarized for the ES. The number and percentage of subjects completing the study and the number and percentage of subjects withdrawing from the study, including the reasons for withdrawal, will be summarised for the ES for Part A and the RS for Part B.

9.2 Demographics and Baseline Characteristics

Descriptive summaries of subjects' demographics will be presented for the RS and FAS and will be listed for the ES. Demographics will include: age (years), gender, height (cm), weight (kg) and BMI (kg/m^2). Age will be calculated from the year of birth relative to the date of informed consent if year of birth is recorded. Baseline values for duration of LBP (years), LBP NRS score, 7-day average LBP NRS daily score where available, RMDQ-23, Derived RMDQ-24, ODI, Leg Pain NRS and PCS score will all be summarised at Baseline for the SS. The PCS scale score at Baseline (as defined in section 4.8) will also be listed. Age, height, weight, BMI and the baseline values of the pain severity scores will be summarised as continuous data, EudraCT age categories and gender as frequencies and percentages of the respective study populations. The duration of CLBP will be calculated from the medical history where possible.

No statistical comparisons will be carried out on baseline data.

9.3 Baseline Cytokines

If the data are available, this will be summarised as continuous data for the SS, including box plot presentations also on the SS; All data observed results will also be listed.

9.4 Medical History

Medical History terms will be coded using the most recent MedDRA Dictionary. All medical history will be summarised for the SS as the number and percentage of subjects reporting each medical category. All verbatim descriptions will be listed for the ES in Part A and the RS in Part B.

9.5 Dosing and Extent of Exposure

All study drug administration data will be listed for the SS. The quantity of PP353 delivered and the percentage of the planned dose of PP353 will be summarised as continuous variables for the FAS and PPS populations in the treated arm of the study.

For Part B under Protocol Amendment 4.0,5.0 and 6.0, Placebo subjects receive a sham injection for both administrations.

9.6 Non-study Medications and Therapies

Non-study medications will be coded using the most recent World Health Organization Drug Dictionary WHO-DRUG Global and the Anatomical Therapeutic Chemical (ATC) classification system. Coding will include the drug class and drug name.

All prior and concomitant medication will be reported and are defined as below:

Prior medication: Started and stopped before first study drug administration.
Concomitant medication: Has at least one day in common with the intervention or follow-up period.

All non-study medications will be listed for the RS, including verbatim descriptions, coded terms, whether the medication was prior or concomitant and whether it was ongoing on the date of first injection. Summaries will be presented by drug class and drug name and by decreasing frequency of subjects taking medication in the active treatment group. For each subject, multiple records of the same concomitant medication will be counted once within a drug class and drug name. The denominator for percentages of the drug classes and the drug names will be the total of the respective study populations.

Separate tables will be produced of prior and concomitant medication for the SS. A subject could have a record of the same drug in both prior and concomitant medication if the drug was stopped before the intradiscal injection took place but was restarted later in the study.

Additional tables will be produced of prior and concomitant medication which was used to treat CLBP for the SS.

9.7 Other Variables

Following each intradiscal injection subjects are monitored for injection site reactions on an Injection Site Reaction Grading Scale, defined precisely in section 5.2.7 of the Protocol. Pain, tenderness, erythema and swelling will be recorded on a 4-point scale as, None, Mild, Moderate or Severe. If they record that there were no injection site reactions, all four symptoms will be scored as “None”. The severity scores will also be recorded as adverse events. If “No” is reported in response to “Were there any injection site reactions?” there will be no AE record.

The scores of pain, tenderness, erythema and swelling will be summarised for each injection as frequencies and percentages for the safety set.

10. CHANGES IN THE STATISTICAL METHODS FROM THOSE STATED IN THE PROTOCOL

As an additional explanatory endpoint, a binary categorisation of the primary end point of LBP NRS to look at 30% improvement, LBP-NRS ≤ 2 and a percentage improvement to be determined from global perceived effect responders has been included as it may inform endpoints to use in subsequent studies. 7-day average LBP NRS daily score ≤ 2 has also been added for the same reason.

Specified that the Kenward-Roger method for denominator degrees of freedom will be used for the primary analysis (though alternative methods may be explored), as the protocol did not specify which method to use.

The Month 1 timepoint has also been included for the following endpoints for summaries: LBP NRS and RMDQ-23.

11. REFERENCES

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. Eur Spine J. 2013b; 22:697–707.

Bråten LCH, Rolfsen MP, Espeland A, Wigemyr M, Aßmus J, Froholdt A, Haugen AJ, Marchand GH, Kristoffersen PM, Lutro O, Randen S, Wilhelmsen M, Winsvold BS, Kadar TI, Holmgard TE, Vigeland MD, Vetti N, Nygaard ØP, Lie BA, Hellum C, Anke A, Grotle M, Schistad EI, Skouen JS, Grøvle L, Brox JI, Zwart JA, Storheim K; AIM study group. 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 Oct 16;367:l5654. doi: 10.1136/bmj.l5654. Erratum in: BMJ. 2020 Feb 11;368:m546. PMID: 31619437; PMCID: PMC6812614.

Siddiqui O, Hung HM, O'Neill R. MMRM vs. LOCF: A comprehensive comparison based on simulation study and 25 NDA datasets. J Biopharm Stat. 2009; 19:227-246.

12. TABLES, FIGURES, LISTINGS & OUTPUT

All tables and figures and listings will be described in a separate "Tables Manual" which will form an attachment to this SAP. This will include shells of the planned table, example charts and details of the planned subject listings.