MENARINI RICERCHE

DAVID

Dexketoprofen **A**nalgesic e**V**olution w**I**th trama**D**ol- **DAVID Study**

Statistical Analysis Plan

Version 1.0

Date: 14 February 2017

Study code	DEX-TRA-06
EudraCT-Number	2015-004152-22
Investigational Medicinal Product	Dexketoprofen trometamol 25mg /Tramadol hydrochloride 75mg
Development phase of study	III b
Sponsor	**MENARINI RICERCHE SpA** Via Sette Santi 1, 50131 Firenze (Italy)
Co-ordinating Investigator	[redacted]

SIGNATURE PAGE

Prepared by:

_____ _____

[redacted]
Senior Biostatistician
Menarini Ricerche S.p.A.

Approved by:

_____ _____

[redacted]
Date
Head of Data Management & Biostatistics
Menarini Ricerche S.p.A.

[redacted] [redacted]
Date
Study Physician
Menarini Ricerche S.p.A.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ADR	Adverse Drug Reaction
AE	Adverse Event
ALAT	Alanine Aminotransfetrase
ANCOVA	Analysis of Covariance
ASAT	Aspartate Aminotransferase
ß-HCG	β Human Chorionic Gonadotropine
BDRM	Blinded Data Review Meeting
BOCF	Baseline Observation Carried Forward
BP	Blood Pressure
CI	Confidence Interval
CM	Concomitant Medication
CPK	Creatine Phosphokinase
CRF	Case Report Form
CS	Clinically Significant
DKP.TRIS	Dexketoprofen Trometamol
ECG	Electrocardiogram
EOS	End of Study
ET	Early Termination/Withdrawal
GEE	General Estimating Equations
GGT	Gamma-Glutamyl Transpeptidase
Hb	Haemoglobin
HCT	haematocrit
HR	Heart Rate
ICF	Informed Consent Form
IMP	Investigational Medicinal Product
INR	International Normalized Ratio
ITT	Intention-to-Treat
LDH	Lactate Dehydrogenase
LOCF	Last Observation Carried Forward
MAO	Monoamine Oxidase
MCH	Mean Corpuscular Haemoglobin
MCHC	Mean Corpuscular Haemoglobin Concentration
MCV	Mean Corpuscular Volume
MMRM	Mixed Models Repeated Measures
NCS	Not Clinically Significant
NSAID	Non-Steroidal Anti-Inflammatory Drug
PAR	Pain Relief
PCA	Patient Controlled Analgesia
PGE	Patients Global Evaluation
PI	Pain Intensity
PID	Pain Intensity Differences

PP	Per Protocol
PT	Preferred Term
PTT	Partial Thromboplastin Time
RBC	Red Blood Count
RM	Rescue Medication
SAE	Serious Adverse Event
SADR	Serious Adverse Drug Reaction
SD	Standard Deviation
SE	Standard Error
SOC	System Organ Class
SPID	Sum of Pain Intensity Differences
TEAE	Treatment Emergent Signs and Symptoms
TOTPAR	Total Pain Relief
TRAM.HCI	Tramadol Hydrochloride
VAS	Visual Analogue Scale
VRS	Verbal Rating Scale
VS	Vital Signs
WBC	White Blood Count
WHO	World Health Organization
WOCF	Worst Observation Carried Forward

1 INTRODUCTION

The purpose of this document is to provide further details about the statistical analysis methods specified in study protocol for the Phase III-b study of dexketoprofen trometamol (DKP.TRIS) and tramadol hydrochloride (TRAM.HCI) fixed-combination in patients with moderate to severe acute pain after removal of impacted lower third molar. The study is sponsored by Menarini Ricerche S.p.A. A brief description of the study objectives and study design are given in sections 2 and 3 respectively. Subsequent sections include the definition of analysis populations, efficacy and safety endpoints including details about statistical methods and hypotheses.

Menarini Ricerche is developing a combination of DKP.TRIS and TRAM.HCI, two analgesics with different mechanisms of action, for the treatment of moderate to severe acute pain, based on the rationale that multimodal analgesia is necessary in most patients suffering from acute and chronic pain, particularly of moderate to severe intensity.

Pain is the most common symptom for which patients seek medical attention. Acute pain is typically associated with recent tissue injury and has a short duration. If not properly treated, prolonged activation of the pain pathways can lead to further neurophysiologic changes collectively called 'central sensitization' which may prolong recovery and in some cases convert acute pain into a chronic condition. Proper analgesic treatment can reduce this risk.

DKP.TRIS and TRAM.HCI are well-known and widely used analgesics which have been present in the market of most European countries for a very long time. The clinical development of the DKP.TRIS and TRAM.HCl fixed combination started with a phase I safety and pharmacokinetics (PK) study (DEX-TRA-PK), aimed at investigating the potential drug-drug interactions between the products as well as their tolerability when concomitantly administered as a single oral dose to healthy subjects. The study results indicated that there was no drug-drug interaction between DKP.TRIS and TRAM.HCl and that the concomitant administration was well tolerated. A phase II dose-finding study (DEX-TRA-02) was performed in the most frequently used model of acute nociceptive pain of moderate-severe intensity (impacted third mandibular molar tooth extraction) aimed at evaluating the analgesic efficacy and safety of DKP.TRIS (12.5 mg and 25 mg) and TRAM.HCl (37.5 mg and 75 mg) given as four different combinations and as single components. Results from DEX-TRA-02 allowed for the selection of DKP.TRIS 25 mg + TRAM.HCl 75 mg as the optimum combination of doses to be further evaluated in the subsequent phase III pivotal studies (DEX-TRA-04 and DEX-TRA-05).
DEX-TRA-04 and DEX-TRA-05 were two phase III registration studies aimed at evaluating the analgesic efficacy and safety of the DKP.TRIS/TRAM.HCl 25/75mg oral fixed-dose combination in comparison with each single component (tramadol given at a higher dose, 100mg), following single and repeated-dose administration, in two recognised models of visceral and somatic moderate to severe acute pain, namely abdominal hysterectomy and hip arthroplasty, respectively. The study results provided robust evidence of the superiority of DKP.TRIS/TRAM.HCl 25mg/75mg over the single components in the management of moderate to severe acute pain.

In this study, patients will receive one single oral dose of DKP.TRIS/TRAM.HCl 25mg/75mg, TRAM.HCl/paracetamol 75mg/650mg or placebo (in a 2:2:1 ratio).

RM consisting of ibuprofen 400mg (a maximum of 2 tablets, each one separated by a minimum interval of 4 hours) will be available on request during the 8-hour post-dose period. Ibuprofen 400mg is a safe and effective drug for the treatment of dental and post-operative pain (Ibuprofen SmPC).

This SAP is based upon the following study documents:

- Study Protocol, Version 2.0 (November 24, 2015)
- electronic Case Report Form (eCRF), Version 2.0 (May 20, 2016)

2 STUDY OBJECTIVES

2.1 Primary Objective

To assess the comparability of DKP.TRIS/TRAM.HCl and TRAM.HCl/paracetamol in terms of analgesic efficacy on moderate to severe pain following impacted lower third molar extraction.

2.2 Secondary Objective

To confirm the safety and tolerability profile of DKP.TRIS/TRAM.HCl following single dose administration.

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

DEX-TRA-06 is designed as a multicentre, randomised, double-blind, double-dummy, parallel-group, placebo and active-controlled, single dose, phase IIIb study, intended to be conducted in approximately 19 sites among Europe.

The study design includes a total of 3 treatment arms, with a 2:2:1 randomisation ratio:

- DKP.TRIS/TRAM.HCl
- TRAM.HCl/paracetamol
- Placebo

Male or female patients, aged more than 18 years, scheduled to undergo surgical extraction (under local anaesthesia) of at least one impacted lower third molar and who are otherwise healthy will be enrolled in the study.

Upon completion of the surgical extraction, those patients experiencing moderate to severe pain (Numerical Rating Scale, NRS ≥4) within 4 hours after the end of surgery will be considered eligible to progress with randomisation and to receive one single oral dose of the assigned study treatment.

Rescue medication (RM) consisting of ibuprofen 400mg (a maximum of 2 tablets, each one separated by a minimum interval of 4 hours) will be available on request during the 8-hour post-dose period.

The analgesic efficacy evaluation will be based on patients' e-Diary scores of pain intensity (PI) and pain relief (PAR) at pre-defined intervals during the 8-hour post-dose period; the patient global evaluation (PGE) of the study medication at the end of the assessment period; the time to onset of the analgesic effect within two hours post-dose and the first use of RM. The safety evaluation will include a physical examination, vital signs (VS) and recording of adverse events (AEs).

The individual study participation will have duration of approximately 3 weeks, encompassing:

- Screening period, for study eligibility assessment (within 2 weeks prior to randomisation), including the pre-surgery procedures to be completed at least one day prior to surgery and ending with the 4-hour qualification post-surgery.
- Randomisation and treatment administration (day 1, t_0) followed by a 8-hour pain and analgesic effect assessment period as follows:
  - from t_0 to t_{2h}, at the study centre.
  - from t_{2h} to t_{8h}, out of the study centre.

- End of study visit (6 ± 1 days after randomisation).

The design contemplates a sample size of 640 randomised patients. It is anticipated that approximately 800 patients will be needed to be screened in order to obtain 640 randomised patients, under the assumption that approximately 20% of the screened patients will not satisfy the eligibility criteria for randomisation. The enrolment will be competitive.

Figure 1: Schematic Study Outline

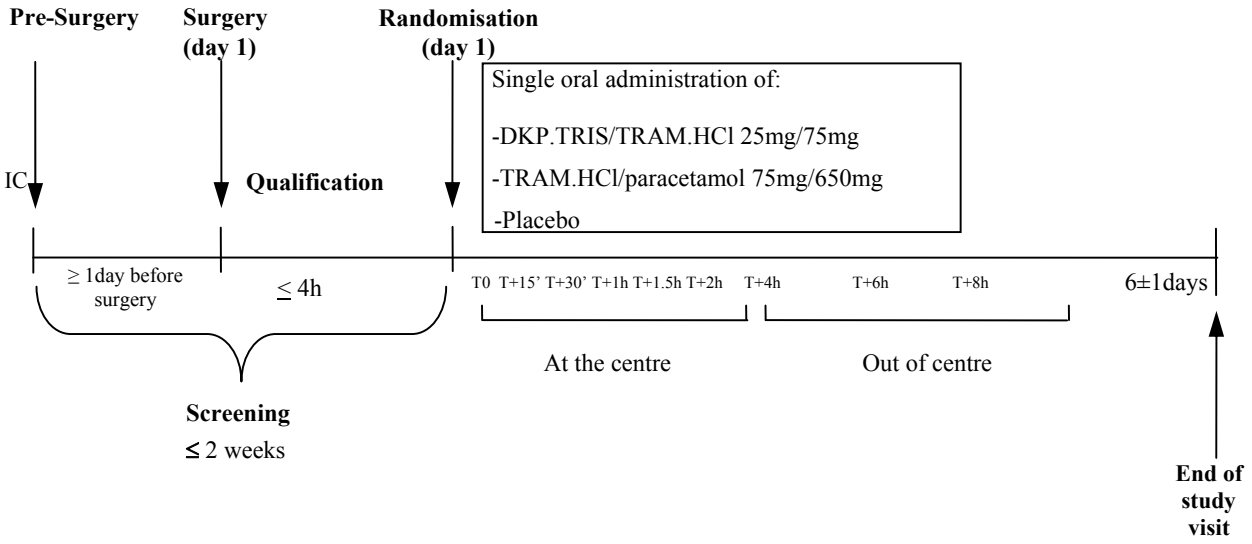


Table 1: Study Flow Chart

Study procedures	***Screening[a]***		***Randomisation, treatment administration and efficacy assessment period***									***End of Study***
	Prior to surgery	**Qualification (day 1)**	**At the centre (day 1)[h]**						**Out of centre (day 1)[i]**			**(6± 1 days after randomisation)**
			T-0 h	**T+15'**	**T+30'**	**T+1h**	**T+1.5h**	**T+2h**	**T+4h**	**T+6h**	**T+8h**	
Check of Informed consent[b]	X											
Demographics and Medical History	X											
Inclusion / Exclusion criteria	X	X										
Physical Examination	X											X
Vital sign (BP, HR)	X	X										X
Safety Laboratory Tests	X											
Pregnancy test (if applicable)	X	X[c]										X
e-Diary Instruction	X											
e-Diary dispensing & testing of its main functions		X										
Randomisation to treatment and Dosing			X									
PI assessment (NRS)[g]		X[d]	X[e]	X	X	X	X	X	X	X	X	
PAR assessment (VRS)[g]				X	X	X	X	X	X	X	X	
Patient's Global Evaluation											X[f]	
FPPAR and MPAR[g]			← **As soon as they are experienced** →									

Study procedures	***Screening[a]***		***Randomisation, treatment administration and efficacy assessment period***									***End of Study***
	Prior to surgery	**Qualification (day 1)**	**At the centre (day 1)[h]**						**Out of centre (day 1)[i]**			**(6± 1 days after randomisation)**
			T-0 h	**T+15'**	**T+30'**	**T+1h**	**T+1.5h**	**T+2h**	**T+4h**	**T+6h**	**T+8h**	
Dispensing of RM and RM diary			X									
RM and RM diary Return / Accountability												X
Return of e-Diary												X
AEs/CM recording	X	X	**While patients are at site**									X

[a] The entire screening period, including the pre-surgery evaluations, the surgical procedure and the qualification period, must be completed within 2 weeks prior to randomisation
[b] IC must be signed before any other study procedure is performed
[c] To be repeated on day 1 before randomisation if not performed the day before as screening procedure
[d] NRS-PI $\geq$ 4 is required for randomisation
[e] NRS-PI at T_0 is recorded as baseline PI
[f] Or immediately before RM intake
[g] To be completed upon RM intake, if any
[h] E-Diary recording is allowed within a ±3 minutes time window at T+15', T+30', T+1h, T+1.5h and T+2h
[i] E-Diary recording is allowed within a ±10 minutes time window at T+4h, T+6h and T+8h

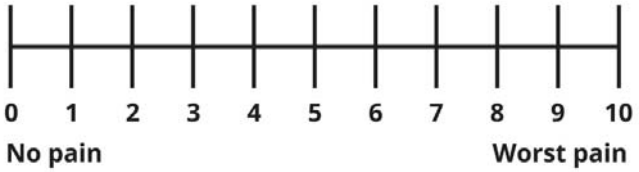
3.2 Efficacy and Safety Variables

3.2.1 Efficacy Variables

3.2.1.1 Pain Intensity - Numerical Rating Scale (NRS)

An 11-point numerical rating scale (ranging from 0 [no pain] to 10 [worst pain], as represented in figure 2) will be used for the subjective assessment of PI. Patients will be asked to select the number that best represents their PI at the time of the assessment.

Figure 2: Pain Intensity - Numerical Rating Scale (NRS)



NRS-PI will be measured:

- During the qualification period, in order to confirm eligibility, namely *moderate or higher PI (NRS ≥4);*
- immediately prior to the treatment administration (t_{0h}),
- at pre-specified time points during the 8-hour post-dose period, (i.e: at t15min, t30min, t1h, t1.5h, t2h, t4h, t6h and t8h).

3.2.1.2 Pain Relief – Verbal Rating Scale (VRS)

A 5-point VRS will be used for the 8-hour post-dose PAR subjective assessment.
Patients will be asked to answer the question 'How do you rate your pain relief?' and to select the adjective that best describes the level of pain relief on the e-Diary. The number associated with the adjective chosen by the patient will constitute the experienced pain relief as reported below:

Score	Pain Relief
0	No relief
1	A little (perceptible) relief
2	Some (meaningful) relief
3	Lot of relief
4	Complete relief

VRS-PAR will be measured at pre-specified time points during the 8-hour post-dose period, i.e. at t_{15min}, t_{30min}, t_{1h}, $t_{1.5h}$, and t_{2h}, t_{4h}, t_{6h} and t_{8h}.

3.2.1.3 Use of Rescue Medication

Time of first intake of rescue medication (Ibuprofen 400mg) will be recorded on the e-Diary. After this, patients will stop using the e-Diary.

3.2.1.4 Time to onset of analgesia

The [redacted] method will be used to assess the onset of analgesia, within the first 2 hours post dose (i.e.: while the patient is still at site).
Patients will be instructed to stop the first [redacted] by pressing the 'first perceptible pain relief' button when the relief from pain is first perceptible (i.e. at the moment they first feel any pain relief whatsoever) (FPPAR) and then to stop the second [redacted] by pressing the 'meaningful pain relief' button when the relief from pain becomes meaningful to them (MPAR).
In case of RM use during the first 2 hours, the patient will be no longer required to use the [redacted] and they will be considered as censored.
FPPAR and MPAR are assessed by [redacted] or the time when the patients report for the first time a pain relief score ≥1 and ≥2 respectively, whichever happens first.
Time to confirmed FPPAR (time to onset of analgesia) is defined as the time to FPPAR for those patients who achieve meaningful pain relief (MPAR).

3.2.1.5 Patient's Global Evaluation

Patients will be asked to give a subjective overall assessment (PGE) of the study medication by using a 5-point VRS.
Patients will be asked to answer the question: 'How would you rate the medication received for your pain?' and to select the adjective that best rate the study medication that they received for pain relief on the e-Diary. The number associated with the adjective chosen by the patient will constitute the global evaluation score, as reported below:

Score	PGE
1	Poor
2	Fair
3	Good
4	Very Good
5	Excellent

The PGE will be recorded at the end of the assessment period (at 8 hours post-dose or whenever the patient uses RM).

3.2.2 Safety Variables

Assessment of safety will be based on the evaluation of the following variables:

- Incidence, intensity (severity), seriousness and treatment-causality of TEAEs, i.e. AEs that occurred after study drug intake.

- Frequency of clinically significant changes in:
  - Physical examination (PE)
  - Vital Signs (VS)

  NOTE: for PE, the *screening period* values will be used as baseline; for VS, the values collected during the qualification period will be used as baseline.

3.2.2.1 Physical Examination, Height, Weight and Vital Signs

Physical examination (review of body systems) will be performed during the Screening period and at the End of study visit. Measurement of height (in whole centimeters) and weight (in kilograms to one decimal place) will be performed during the Screening period only.
Vital signs (i.e. BP [systolic and diastolic, mmHg] and HR [beats/min]) will be measured 1) during the screening period, 2) during the qualification period and, 3) at the End of Study visit.
Out of range measures will be repeated at the Investigator's discretion.
The collected data will be registered in the subject's source record and reported on the e-CRF.

3.2.2.2 Clinical Laboratory Evaluation

Blood laboratory safety tests will be performed locally, with blood samples (8ml approximately) collected in fasting conditions during the Screening period. The following parameters will be analysed:

- Haematology: Red blood cell (RBC) count, haematocrit (HCT), haemoglobin (Hb), mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), platelet count, white blood cells (WBC) count, and differential WBC count (neutrophil, lymphocyte, monocyte, eosinophil and basophil counts, absolute and %).

- Serum biochemistry: Sodium, potassium, chloride, total calcium, glucose, creatinine, urea, total protein, albumin, alkaline phosphatase, gamma-glutamyl transpeptidase (GGT), aspartate aminotransferase (ASAT), alanine aminotransferase (ALAT), total bilirubin, direct bilirubin, lactate dehydrogenase (LDH), and creatine phosphokinase (CPK).

- Coagulation test: prothrombin time, partial thromboplastin time (PTT), INR.

Urinalysis will be performed during the Screening period, using Roche Combur-10® dipsticks. The following parameters will be analysed: Specific gravity, pH, leucocytes, nitrite, protein, glucose, ketones, urobilinogen, bilirubin, blood **(**erythrocytes/haemoglobin**)**.

Pregnancy tests **(**when appropriate) will be performed during the Screening period, during the qualification period and at the End of Study visit, by determination of the urine β human chorionic gonadotropine (ß-HCG), using locally available commercial dipsticks.
Unscheduled re-tests can be performed as required according to the investigator's judgement.
The reference ranges of laboratory parameters described in this study will be submitted by each centre to the sponsor prior to beginning of this study and will be updated as appropriate.
The lab print-outs should be identified with the laboratory requisition number, patient number, patient's age, gender, and site number, as well as with the date and time of sample collection. All results should be reviewed by the Investigator with print-outs being dated and signed by the investigator and stored in the patient's record. Out of range values will be judged by the Investigator as "abnormal clinically significant" or as "abnormal not clinically significant" and stored in the patient's record. Note: abnormal evaluation can be repeated at the Investigator's discretion. Records will be transferred on the e-CRF.

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings to be included in the report will be independently checked for consistency, integrity and in accordance with standard Menarini Ricerche procedures.

4.2 General Presentation Considerations

'Baseline' is defined as the last available non-missing pre-treatment assessment (including unscheduled measurements if available and appropriate), unless otherwise specified.
'End of Study' is defined as having completed the EOS visit.
'Duration of Treatment' will be calculated relative to the date and time of randomization:

Duration of Treatment (h) = Last Treatment Date/Time - Randomization Date/Time.

The entire screening period comprehend the pre-surgery evaluation, the surgical procedure (Day 1 Screening) and the qualification period.

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum and number of observations, unless otherwise stated. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, median, lower quartile and upper quartiles (if required) will be reported to one more decimal place than the raw data recorded in the database. The SD will be reported to two more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.

Categorical data will be summarized in terms of the number of patients providing data at the relevant time point (n), frequency counts and percentages. Any planned collapsing of categories will be detailed in the SAP text and the data displays.

Percentages will be presented to one decimal place and will be calculated using n as the denominator.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. P-values less than 0.001 will be presented as "<0.001".

Confidence intervals will be presented to one more decimal place than the raw data.

All report outputs will be produced using SAS® version 9.3 or a later version in a secure and validated environment.

4.3 Analysis Populations

The **Safety Population** is defined as all patients randomized who have received the study treatment.

The **Intention-to-treat (ITT) Population** is defined as all patients randomized.

The **Per Protocol (PP) Population** will include all patients from the ITT population who did not experience relevant protocol violations related to the efficacy endpoints of primary interest.

The primary efficacy evaluation will be demonstrated on both PP and ITT populations in order to get major robustness in result of the non-inferiority study. Other efficacy analyses will be performed by using the ITT population, unless otherwise stated.

For the ITT population, in the event that a patient is allocated the incorrect study treatment as per the study randomization list, patients will be summarized and analyzed 'as randomized' i.e. by randomized treatment arm.

For the safety population, if a patient is allocated the incorrect study treatment as per the study randomization list, patients will be summarized and analyzed 'as treated' i.e. by allocated treatment arm.

The following summaries on the analysis populations will be provided:

- The number and percentage of patients in each analysis population by treatment arm and overall.
- A by-patient listing of analysis population details presented by treatment arm and will include: country, center, patient identifier, and inclusion/exclusion flags for each population. All patients screened should appear on this listing. If patient data has been partially excluded, visit should also appear on this listing.

4.4 Study Patients

4.4.1 Disposition of Patients

A clear accounting of the disposition of all patients who enter the study will be provided, from screening to study completion. In the event that a patient prematurely discontinues the study for any reason before the end of the treatment period, a final assessment of PGE should be collected on the e-Diary. An EOS visit will be completed for all patients who complete the study as well as those who prematurely terminate study participation after having received any dose of study treatment. This visit will also be performed if the patient withdraws their consent.

The number and percentage of patients screened, randomized, treated, terminating early (ET) and completing the study will be summarized by treatment arm and overall (Analysis population: All patients screened).

A by-patient listing of eligibility details, randomization details (including whether the blind was broken at discontinuation), visit dates and withdrawal/study completion details (including reason for discontinuation and duration of treatment prior to discontinuation) will be provided.

4.4.2 Protocol Deviations

Major or minor protocol deviations are defined as those deviations from the protocol likely to have an impact on the efficacy or safety of study treatments. The impact of major protocol deviations on the efficacy results will be investigated by assessing the robustness of the study results.
Before any analyses of the data, a blinded data review meeting (BDRM) will take place to identify protocol deviations. The category list of potential protocol violators is reported in Appendix 1.
Cases of major protocol violations will be discussed, and on a case-by-case basis it will be determined whether or not to exclude the patients from the PP population. The final decisions on which patients to include or exclude from the PP population will be finalized prior to unblinding.

Major protocol deviations and any action to be taken regarding the exclusion of patients or affected data from specific analyses will be described in a separate document.

The following protocol deviation summaries will be produced:

- The number and percentage of patients with a major protocol deviation by treatment arm and overall and by type of deviation (Analysis population: ITT)
- The number and percentage of patients with a major protocol deviation that brings to the exclusion from PP by treatment arm and overall and by type of deviation (Analysis population: ITT)
- The number and percentage of patients with a minor protocol deviation by treatment arm and overall and by type of deviation (Analysis population: ITT)
- A by-patient listing of all major protocol deviations.
- A by-patient listing of all minor protocol deviations.

4.5 Demographic and Other Baseline Characteristics

The demographic and baseline characteristic variables identified below will be summarized by treatment arm and overall for the ITT, PP and/or Safety population:

- Demographic data (ITT population):
    - Age
    - Sex
    - Race
    - Country Childbearing potential
    - Height
    - Weight

- Baseline characteristics (Safety population):

    - Physical examination
    - VS (HR, BP)
    - Laboratory safety tests (hematology, clinical chemistry, coagulation test, urinalysis)
    - Pregnancy test

By-patient listings of all demographic data and baseline characteristics as stated above will be provided.

Medical history will be classified using the Medical Dictionary for Regulatory Activities (MedDRA) Version 18.1, summarized by treatment arm and overall and listed.

Prior and concomitant medications as specified in the eCRF will be tabulated separately by treatment arm and overall. The proportion of patients receiving prior medications (medications ended within 4 weeks of the screening date) will be summarized by preferred medication name. A patient will be counted at most once for each prior preferred medication, even if the patient took the same medication on multiple occasions. Concomitant medications (medications started on/after the screening date) and Prior and Concomitant medications (medications started prior to the screening date but ended after this date or still ongoing) will be summarized similarly. Prior and concomitant medications will be classified according to the World Health Organization (WHO) Drug Dictionary version September 2015 and presented by ATC code level 3.

Medication start and stop dates/times will be compared to the date of screening date to allow medications to be classified as either prior only, both prior and concomitant, or concomitant only. Prior and concomitant and concomitant only medication will be summarized in the same table.

If medication start and/or stop dates are missing or partial, the dates will be compared as far as possible with the date of screening date. Medications will be assumed to be concomitant only, unless there is clear evidence (through comparison of partial dates) to suggest that the medication started within 4 weeks prior to the screening date. If there is clear evidence of that, the medication will be assumed to be both prior and concomitant, unless there is clear evidence to suggest that the medication stopped prior to the screening date. If there is clear evidence to suggest that the medication stopped prior to the screening date, the medication will be assumed to be prior only.

4.6 Efficacy Evaluation

4.6.1 Analysis and Data Conventions

4.6.1.1 Study Hypotheses

The primary efficacy variable will be analyzed on the PP and ITT populations to assess the non-inferiority hypothesis using an analysis of covariance (ANCOVA) for testing the differences in treatment efficacy, as quantified by Total Pain Relief over 6 hours after treatment intake (TOTPAR6), between DKP.TRIS + TRAM.HCl and TRAM.HCl + Paracetamol arms. The Total Pain Relief has been selected as primary variable since for acute pain studies utilizing TOTPAR to assess pain relief may be more sensitive to treatment effects than utilizing SPID to assess pain intensity.

ANCOVA is a statistical procedure that allows to include both categorical and continuous variables in a single model, and so is a general linear model which blends ANOVA and regression:

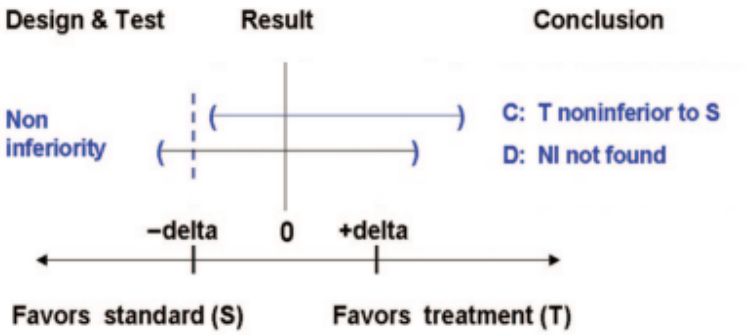
$$y_{ij} = \mu + \alpha_i + \beta(x_{ij} - \bar{x}) + \varepsilon_{ij}$$

where:

y_{ij} = jth replicate observation of response variable.
μ = mean value of response variable.
α_i = $\mu_i - \mu$.
β = combined regression coefficient.
x_{ij} = covariate value for the jth replicate observation from the ith level of factor.
$\bar{x}$ = mean value of covariate.
ε_{ij} = unexplained error associated with the jth replicate observation from the ith level of factor.

A non-inferiority test aims to demonstrate that results obtained from the test drug are not appreciably worse than the ones of standard drug. A non-inferiority margin (δ) has to be specified in order to characterize the largest difference that you consider to be dismissible (Figure 3).

Figure 3: Hypotheses of Non-inferiority tests.



The ANCOVA model will include terms of treatment and the baseline PI (NRS) with one-sided significance level of 2.5%. The following hypothesis will be tested:

$$H_0 : \mu_{TRAM.HCl\ /\ paracetamol} - \mu_{DKP.TRIS\ /\ TRAM.HCl} \geq \delta$$

against

$$H_1 : \mu_{TRAM.HCl\ /\ paracetamol} - \mu_{DKP.TRIS\ /\ TRAM.HCl} < \delta$$

with δ of 20% as non-inferiority limit and μ as the respective mean.

In case the non-inferiority is confirmed on PP and ITT populations, the superiority of DKP.TRIS + TRAM.HCl versus TRAM.HCl + Paracetamol will be tested on the ITT population by using the same ANCOVA model described above.

Hence, the following hypotheses are tested:

H(0): μ(DKP.TRIS + TRAM.HCl) = μ(TRAM.HCl + Paracetamol)
against H(1): μ(DKP.TRIS + TRAM.HCl) ≠ μ(TRAM.HCl + Paracetamol)

Comparison of DKP.TRIS + TRAM.HCl and TRAM.HCl + Paracetamol versus Placebo will be performed in order to confirm the model sensitivity on the ITT population.

The following hypotheses are tested:

H(0): μ(DKP.TRIS + TRAM.HCl) = μ(Placebo)
against H(1): μ(DKP.TRIS + TRAM.HCl) ≠ μ(Placebo)

H(0): μ(TRAM.HCl + Paracetamol) = μ(Placebo)
against H(1): μ(TRAM.HCl + Paracetamol) ≠ μ(Placebo)

4.6.1.2 Multi-center Studies

The term ‘Center’ will be used to define each investigator site. As center is not a stratification variable for the randomization, it is not included in the main model for the primary analysis. The center effect, as random, will be investigated and reported, if significant, in the final report as sensitivity and exploratory analysis..

4.6.1.3 Adjustments for Covariates

For the primary efficacy analysis the following variables are included as covariates in the ANCOVA model mentioned in §4.6.1.1 :

- Treatment (as the main effect).
- Baseline PI level.

Treatment will be classed as being one of the following three treatment arms with a 2:2:1 ratio:

Arm	***Single-dose phase***
1	DKP.TRIS + TRAM.HCl
2	TRAM.HCl + Paracetamol
3	Placebo

4.6.2 Handling of Dropouts or Missing Data

For the primary analysis, the method of last observation carried forward (LOCF) will be applied among patients who missed more than one consecutive value, otherwise the missing value will be replaced by the mean of the two non missing data collected respectively before and after the missing one. If PI is missed at t0h the value recorded during qualification procedure will be used as baseline.
After rescue medication intake, PI will return to its baseline (t0h) level and PAR to zero ("no relief") for all subsequent time points (i.e. baseline observation carried forward, BOCF).

In case non-inferiority is confirmed the primary efficacy endpoint will be also evaluated in the ITT population using other methods of imputation different from the ones described above as for sensitivity analysis:

1. After RM intake the value will be handle with LOCF imputation method instead of BOCF.
2. For missed values any method of imputation will be implemented.
3. For both missed values and values after RM intake any imputation will be used.

If a discrepancy between the Rescue Medication reported in the eCRF and the one registered in the eDiary is present, the Rescue Medication collected in the eCRF will be considered as correct and used for the analysis since is confirmed by the source data verification. Sensitivity analyses considering only the Rescue Medication recorded in eDiary will be also implemented at least for the primary endpoint and the main secondary endpoints.

4.6.3 Multiplicity

There is one primary endpoint comparison for non-inferiority (as detailed in Section 4.6.1.1) for this study. For the switch from non-inferiority to superiority no multiplicity adjustment is needed, since, as stated in CPMP/EWP/482/99, the superiority interpretation of results corresponds to a simple closed test procedure.
Other additional two comparisons between active treatments and placebo are planned in order to validate the study model. However as they are considered to be co-primary endpoints and they will therefore be interpreted together for purposes of the study, no adjustment for multiplicity is required.

4.6.4 Examination of Subgroups

The uniformity of the treatment effects for the primary efficacy variable will be examined for the following subgroups:

1. Gender (Male, Female)
2. Age (quartiles)
3. Country

The quartile values for age will be determined by examining the distribution of age at baseline.
The results of the primary statistical analyses within each subgroup category will be presented.

4.6.5 Primary Efficacy Variable – Mean TOTPAR ($TOTPAR_6$)

$TOTPAR_6$ is calculated as the weighted sum of the PAR-VRS scores, over 6 hours after the treatment intake.

The general formula for calculating $TOTPAR_x$ is as follows:

$$\mathbf{TOTPAR_x} = \sum [T(i) - T(i-1)] \times PAR(i) \quad \textbf{for i = 1, …, z}$$

where T(i) is the scheduled time,
T(0) = 0,
PAR(i) = PAR at time i.
z is the number of timepoints from the first one after T(0) to T(x).

This primary efficacy variable will be used for the assessment of the primary efficacy endpoint of non-inferiority of DKP.TRIS + TRAM.HC1 versus TRAM.HCl + Paracetamol.

The following SAS program will be implemented:

NB: Since the Placebo group is not considered in this efficacy analysis only the active treatment arms need to be prior selected.

The model uses the REML method (Restricted or Residual maximum likelihood) for the estimation of parameters. REML estimators are obtained by maximizing the part of the likelihood function that is invariant to the fixed effects part of the linear model. It requires the assumptions that the distribution of the dependent variable is normal, and the errors are normally distributed, uncorrelated and have an expected value of 0.

To check assumptions the following tests will be done:

To check the normal distribution of dependent variables within each treatment arm, QQplots and Histograms plots will be done as a graphical test. In addition the Shapiro-Wilk test will be performed. The null hypothesis of normal distribution, H_0, will be tested by using a significance level of 5%.
The following SAS procedure will be implemented:

To check the homogeneity of the variance of the dependent variables the Bartlett test will be done. The null hypothesis, H_0, is that all k population variances are equal. The alternative is that at least two are different. The null hypothesis is rejected at a significance level of 5% if the value of the Bartlett test statistic is greater than the corresponding critical value.

The following SAS procedure will be used:

To check whether the residuals of the utilized mixed model have a multivariate normal distribution and if heteroscedasticity is present, a graphical tests will be implemented by adding option "/ VCIRY" to the "MODEL" statement in the proc mixed reported above. For the multivariate normal distribution graph, the standardized residuals versus their ranks should form a straight line on the plot, which is visually assessed. For the heteroscedasticity graph, instead, if the points are equally distributed around the straight line and no trends are showed, the hypothesis of heteroscedasticity is rejected.

As alternative, if the assumption required by the REML method are not fulfilled or if the likelihood function cannot be evaluated, the MIVQUE0 (Minimum Variance Quadratic Unbiased Estimation) method is used, as this method doesn't require any particular assumptions.
The results of model assumptions will be reported.

Since, as explained in the §6.1.1.2, the ANCOVA model includes ANOVA and regression, it also assumes that the regression coefficients are homogeneous (the same) across the level of categorical independent variables. To check this assumption the interaction term of *Treatment*PI at Baseline* will be add to the covariates already included in the model as described in the §4.6.1.3.
If the interaction term is not significant (associated p-value >0.05) it means that the regression coefficients can be considered homogeneous and the dependent variables can be analyzed as described in §4.6.5. If, alternatively, the interaction term of *Treatment*PI at Baseline* is significant (associated p-value $\leq$ 0.05) it is not possible to interpret the relationship between the dependent variable and the treatment without referring also to the pain intensity at baseline. In this case a comparison of the treatment groups at different group levels of pain intensity (i.e. 5,7,10) needs to be done. The following SAS procedure will be implemented:

To assess the non-inferiority hypothesis, the 95% Interval Confidence of the $TOTPAR_6$ difference between the two active treatment arms (DKP.TRIS + TRAM.HC1 - TRAM.HCl + Paracetamol) will be produced by including the "*CL*" option in the '*estimate*' statement of Proc Mixed as showed above.
If the lower limit of CI is higher than the δ predefined, then the Non-Inferiority can be claimed. Otherwise DKP.TRIS + TRAM.HC1 has to be considered as inferior.
As δ , the 20% of the estimated mean of TRAM.HCl + Paracetamol will be used.

In case the non-inferiority is confirmed in both PP and ITT populations, the superiority of DKP.TRIS + TRAM.HC1 versus TRAM.HCl + Paracetamol will be tested on the ITT population by using the same model and SAS Program described above. $TOTPAR_6$ will be also analyzed by comparing the two active treatments versus the Placebo arm in order to validate the model sensitivity.

4.6.6 Secondary Efficacy Variables

All the secondary efficacy variables will be descriptively analysed. The superiority for DKP.TRIS + TRAM.HCl will be tested when applicable, through and ad hoc inferential analysis.
The following treatment comparison will be presented:

- DKP.TRIS + TRAM.HC1 vs TRAM.HCl + Paracetamol

For the $SPID_6$, as will be described in section 4.6.6.1.2, also the Non-Inferiority is tested as for the primary variable.

4.6.6.1 Pain Intensity

4.6.6.1.1 PI-NRS Scores at Rest

The PI-NRS score at rest will be analyzed by applying a longitudinal model with treatment, timepoint, the interaction between treatment and timepoint and baseline PI (NRS) as independent variables.

The following SAS program will be used:

[redacted]

The term *trt*timepoint* indicates the interaction between treatment and time. In clinical study where the hypothesis is that one method is non-inferior (or superior/equivalent) to the comparator method, an assessment of consistency over time for the treatment effect is often convenient. In practice there is an interaction whenever the effect of one factor depends on the level of another.
Initially a plot for the mean of the interest variable over time and displayed by treatment arm can be useful. If no *trt*timepoint* interaction is detected this can be deleted from the model and the overall treatment effect can be assessed marginally by collapsing over time.

Within-subject correlation is another important issue that must be considered, since we can not assume all data points to be independent as when there is only one observation per subject. A variance/covariance matrix has to be indicated in the *type* statement and, to choose the more appropriate structure, the AIC/BIC criteria (lower is better) will be used. The *rcorr* option displays the correlation matrix that can be also useful for the selection.

The underlying assumptions of normality of residuals will be examined by adding in the 'model' statement the *residual* option that produces an output containing the studentized residuals of the model. Then an univariate procedure will be applied to check the normality of them. A graph will be also produced by adding the *VCIRY* option in the 'model' statement.
If the normality of residuals is not confirmed as required for the estimation by the REML method , the MIVQUE0 method is applied, as this method doesn't require any assumptions.

In addition an analysis by timepoint will be implemented in order to check the efficacy of analgesia in the first 8 hours after the treatment intake.
The following SAS program will be used:

The underlying assumption for the model above will be check in the same way as described in section 4.6.5 .

NB: Since the Placebo group is not considered in the efficacy analysis reported above only the active treatment arms need to be prior selected.

4.6.6.1.2 SPIDs

The mean of the secondary efficacy variable $SPID_6$ (Sum of Pain Intensity Difference over 6 hours) will be analyzed for non-inferiority with the possibility to switch for superiority analogously to the primary efficacy variable. The same model and SAS program described in section 4.6.5 will be applied.
The $SPID_6$ at rest is calculated as the weighted sum of the PID values, over 6 hours after the treatment intake.

The general formula for calculating $SPID_j$ is as follows:

$$\mathbf{SPID_j} = \sum [T(i) - T(i-1)] \; x \; PID(i) \textbf{ for i = 1, …, z}$$

where T(i) is the scheduled time,
T(0) = 0,
PID(i) = PID at time i = PI (0) – PI (i).
z is the number of timepoints from the first one after T(0) to T(j).

The mean of SPID at rest over 2, 4 and 8 hours after the treatment intake ($SPID_2$ $SPID_4$ and $SPID_8$) will be analyzed only for the superiority of DKP.TRIS + TRAM.HC1 vs TRAM.HCl + Paracetamol as described in section 4.6.5. Patients in the Placebo treatment arm will not be considered.
Underlying assumptions will be examined in the same way as described in Section 4.6.5 .

4.6.6.1.3 %max SPID

[redacted]

4.6.6.1.4 Responders to treatment

A responder will be defined as having achieved at least 30% of PI-VAS at rest reduction versus baseline at each pre-specified time point.
The frequencies and relative percentages will be calculated using the PI-VAS values at 2, 4, 6 and 8 hours after the treatment intake.
The percentage of responders will be summarized by treatment and the response will be analyzed using a Chi-square test by comparing the active treatment arms.

The following SAS code will be applied for this analysis:

[redacted]

NB: Since the Placebo group is not considered in the efficacy analysis reported above only the active treatment arms need to be prior selected.

The underlying assumptions of adequate cell counts will be examined by reviewing data frequency and patterns using data the SAS procedure FREQ. If inadequate expected cell counts are present (i.e. less than five), then either a correction to the Chi-square test or an alternative test will be applied (e.g. Fisher's Exact Test).

In addition, in order to have a general post-dose overview, the percentage of responders over 8 hours after the treatment administration will be analysed using a general estimating equations (GEE) analysis. Response will be assessed for each of the following timepoints: 2, 4, 6 and 8 hours.
The following SAS code will be applied for this analysis:

[redacted]

$P(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$ [redacted] θ [redacted]

[redacted] $g(\theta) = \ln\left(\frac{\theta}{1-\theta}\right)$ [redacted]

[redacted]

NB: Since the Placebo group is not considered in the efficacy analysis reported above only the active treatment arms need to be prior selected.

4.6.6.2 Pain Relief

4.6.6.2.1 VRS-PAR score

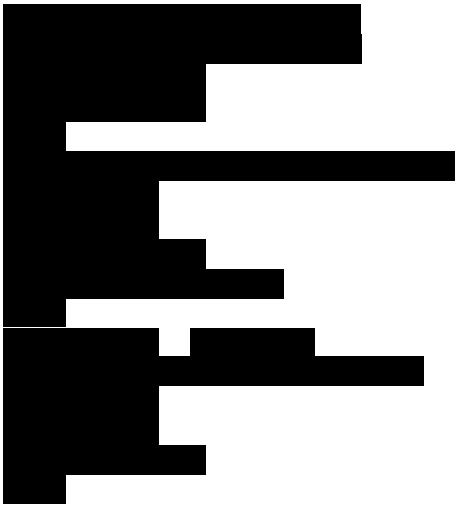
The VRS-PAR score will be analyzed by timepoint by applying a Wilcoxon rank-sum test (or Mann–Whitney–Wilcoxon test), that is a nonparametric test with the null hypothesis that the probability of an observation from the treatment X exceeding an observation from the second treatment Y is equal to the probability of an observation from Y exceeding an observation from X, against the alternative hypothesis that these probability are different.
In an easily way the set of hypothesis can be summarized saying that under H_0 the distributions of both treatment populations are equal against the alternative hypothesis that they are not.

Since the responses are assumed to be numeric discrete and the alternative is restricted to a shift in location, the Hodges-Lehmann estimate can be computed by comparing the median for each treatment.
A Wilcoxon rank-sum test, so, can be interpreted as showing a difference in medians. Under this location shift assumption, we can also interpret the Wilcoxon rank-sum test as assessing whether the Hodges–Lehmann estimate of the difference in central tendency between the two populations differs from zero.

The following SAS code will be applied for this analysis:

The underlying assumptions for the above analysis (independence of data and similar distributions in the groups being compared) will be examined by the Fisher exact test for the independence and by the Kolmogorov-Smirnov, with the relative Empirical Distribution Function (EDF) plot, to test the equality of distribution functions. In addition also the histogram plots for the two group distributions, interpolated by the estimated Kernel function, will be reported. If these examinations indicate that an assumption has been violated then an appropriate data transformation may be applied.

The following SAS procedure will be implemented:



4.6.6.2.2 TOTPARs

Underlying assumptions will be examined in the same way as described in Section 4.6.5.

4.6.6.2.3 %max TOTPAR

4.6.6.2.4 Percentage of responders

A responder will be defined as having achieved at least 50% max TOTPAR.
The mean will be calculated using the % max TOTPAR values at [redacted] 8 hours after the treatment intake.
Responders will be analyzed by using the same statistical model and SAS syntax described in section 4.6.6.1.4 .
Since the Placebo group is not considered in the efficacy analysis reported above only the active treatment arms need to be prior selected.

4.6.6.3 Patient Global Evaluation

The Patient Global Evaluation (PGE) scores will be collected 8 hours after the treatment intake or immediately prior the Rescue Medication intake.

PGE scores will be analyzed using a Wilcoxon rank-sum test (or Mann–Whitney–Wilcoxon test) by applying the Hodges-Lehmann method as already described in section 4.6.6.2.1 .

The following SAS code will be applied for this analysis:

[redacted]

The underlying assumptions for the above analysis will be examined with the same techniques reported in section 4.6.6.2.1 (without 'by' statement) .

NB: Since the Placebo group is not considered in the efficacy analysis reported above only the active treatment arms need to be prior selected.

4.6.6.4 First Perceptible Pain Relief and Meaningful Pain Relief

4.6.6.4.1 Percentage of First Perceptible Pain Relief and Meaningful Pain Relief

The First Perceptible Pain Relief (FPPAR) and the Meaningful Pain Relief (MPAR) are assessed by pressing the specific stop-watch buttons or alternatively assessed when the Pain Relief score is, respectively, ≥ 1 and ≥ 2 within the first 2 hours after the treatment intake.
In case a patients declare a MPAR as first, then the FPPAR will be considered assessed as well and it will correspond in term of time with the MPAR.
If the MPAR is assessed then the FPPAR is defined as confirmed.

Percentage of patients who achieved FPPAR, confirmed FPPAR and MPAR within 30 minutes, 1 hour and 2 hours will be reported and analyzed by using the same statistical methods and SAS syntax described in section 4.6.6.1.4 . Since the Placebo group is not considered in the efficacy analysis reported above only the active treatment arms need to be prior selected.

4.6.6.4.2 Time to First Perceptible Pain Relief and Meaningful Pain Relief

The time to FPPAR and MPAR is defined as the time elapsed between treatment administration and the time of FPPAR and MPAR confirmation.
In case a patient declares a Meaningful Pain Relief (MPAR) as first, then the FPPAR will be considered assessed as well and will correspond in term of time with the MPAR.
If the MPAR is assessed then the FPPAR is defined as confirmed.
The formula for the calculation is:

$$\text{Time to FPPAR/ MPAR} = t_{FPPAR/MPAR} - t_{TI}$$

where $t_{FPPAR/MPAR}$ is the time (hours and minutes) of FFPAR/MPAR assessment,
t_{TI} is the time (hours and minutes) of the study drug intake.

In the case of no FPPAR and/or MPAR assessment, time to FPPAR/ MPAR will be calculated as the time that occurred between t_{TI} and the end of the time of interest (2 hours after the t_{TI}) and the relative patient will be considered as censored. In case a Rescue Medication intake occurs or patient stops to use the e-diary within two hours after the treatment intake, the corresponded patient will be considered as well as censored with a time associated equal to the time of Rescue Medication intake or the time of last assessment available, respectively.

Time to FPPAR/ MPAR will be compared using a log-rank test. The 25th, 50th and 75th percentiles of the Kaplan-Meier estimate with associated 95% CI's will be produced and presented with the log-rank p-value. In addition, a Kaplan-Meier plot of will be presented. Time to confirmed FPPAR will be also analyzed in the same way.
The following SAS code will be applied for this analysis:

The underlying assumptions for the above analysis will be examined by reviewing the data and checking the assumption of proportional hazards by analysing the interaction term '*treatment*log(duration)*', where the logarithmic transformation is implemented to maintain the linearity. If the interaction term is not significant the hazards can be considered as proportional.
In addition the estimated hazard functions plot by using the Epanechnikov kernel-smoothed estimates will be produced as a graphical check.
Should these examinations indicate that an assumption has been violated then an alternative method of analysis (e.g. Wilcoxon Test) may be applied.

The following SAS Syntax will be implemented:

4.6.6.5 Rescue Medication

4.6.6.5.1 Percentage of use of Rescue Medication

Percentage of patients who required Rescue Medication within the first 2, 4, 6 and 8 hours after the treatment intake.
Percentages will be analyzed by using the same method and SAS syntax described in section 4.6.6.1.4 .
Underlying assumptions will be examined in the same way as described in Section 4.6.6.1.4 .

- In order to handle some discrepancies between the Rescue Medications collected in the eCRF and the one recorded in the eDiary the following approaches will be adopted: The first RM between eCRF and eDiary
- The RM recorded in eCRF
- The RM recorded in eDiary

4.6.6.5.2 Time to Rescue Medication

The time to RM is defined as the time elapsed between treatment administration and RM intake. The formula for calculating time to RM is:

$$\text{Time to RM} = t_{RM} - t_{TI}$$

where t_{RM} is the time (hours and minutes) of RM intake,

t_{TI} is the time (hours and minutes) of study drug intake.

In the case of no rescue medication intake, time to RM will be calculated as the time that occurred between t_{TI} and the end of the time period of interest (8 hours after t_{TI}) and the relative patient will be considered as censored. Since the first RM intake have to be reported both in e-diary and in the paper diary if a patient stops to use the e-diary prematurely, he will not be considered as censored.
In order to take into account any inconsistency between the electronic and the paper diaries, three type of time of first RM intake will be calculated by considering the RM reported in the e-diary, the RM reported in the paper diary and the first reported RM between electronic and paper diaries.

Time to RM will be analyzed by using the same method and SAS syntax reported in section 4.6.6.4.2 . Underlying assumptions will be examined in the same way as described in Section 4.6.6.4.2 .

In order to handle some discrepancies between the Rescue Medications collected in the eCRF and the one recorded in the eDiary the following approaches will be adopted:

- The first RM between eCRF and eDiary
- The RM recorded in eCRF
- The RM recorded in eDiary

4.7 Safety Evaluation

All safety summaries and analyses will be based upon the safety population as defined in Section 4.3. No formal statistical testing will be performed and only descriptive statistics provided unless otherwise noted. The number of non-missing values and the mean with the respective SD (for continuous variables) or the percentage (for ordinal/categorical variables) will be reported. Missing safety data will not be imputed.

4.7.1 Adverse Events

AEs will be coded using MedDRA Version 18.1.

AEs will be summarized by primary System Organ Class (SOC), Preferred Term (PT) and treatment. An overview of AEs will be provided, with the number and percentage of patients reporting an event, presented for the following categories:

- Any AEs
- Any serious AEs
- Any AEs leading to study discontinuation.
- Any AEs leading to death

Additionally, the following summaries of AEs will be presented by SOC and PT:

- The number of AEs and the number and percentage of patients with AEs by treatment and overall.

TEAE will be defined as those adverse events that occur for the first time or worsen in terms of intensity (severity)/ seriousness after the drug intake.

Where dates are missing or partially missing, adverse events will be assumed to be treatment-emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the adverse event started prior to the study treatment administration.

TEAE will be summarized by primary System Organ Class (SOC), Preferred Term (PT) and treatment. An overview of TEAEs will be provided, with the number and percentage of patients reporting an event, presented for the following categories:

- Any TEAE
- Any serious TEAE
- Any mild/moderate/severe TEAE
- Any causally related TEAE (Adverse drug reaction [ADR])
- Any TEAE leading to study discontinuation.
- Any TEAE leading to death

Additionally, the following summaries of TEAE will be presented by SOC and PT:

- The number and percentage of patients with TEAE by treatment and overall.
- The number and percentage of patients with TEAE by intensity (mild, moderate and severe).
- The number and percentage of patients with TEAE by causality (related/not related).

Any TEAE is defined as related (ADR) if the causality category falls into one of the following: certainly related, probably related, possibly related and unassessable/unclassifiable. Otherwise TEAE will be considered as not related.

ADRs will be summarized by primary System Organ Class (SOC), Preferred Term (PT) and treatment. An overview of ADRs will be provided, with the number and percentage of patients reporting an event, presented for the following categories:

- Any ADR
- Any serious ADR
- Any mild/moderate/severe ADR
- Any ADR leading to study discontinuation.
- Any ADR leading to death

Additionally, the following summaries of ADR will be presented by SOC and PT:

- The number and percentage of patients with ADR by treatment and overall.
- The number and percentage of patients with ADR by intensity (mild, moderate and severe).

For each patient and each event, the worst intensity recorded will be attributed and used in the by-intensity summaries. Similarly, the worst causality (most related to treatment) will be attributed and used in the by-causality summaries. If intensity or causality is missing, the worst case will be assumed.

A by-patient listing of all AEs (including non-TEAE) will be provided together with a listing of TEAE.

4.7.2 Deaths, Serious Adverse Events, and Other Significant Adverse Events

The number and percentage of deaths (if any) during the study by treatment arm (if numbers allow) will be reported.
Serious AEs (SAEs) will be summarized by primary System Organ Class (SOC), Preferred Term (PT) and treatment. An overview of SAEs will be provided, with the number and percentage of patients reporting an event, presented for the following categories:

- Any SAE
- Any mild/moderate/severe SAE
- Any SAE leading to study discontinuation.
- Any SAE leading to death

4.7.3 Clinical Laboratory Evaluation

Each patient's continuous hematology, biochemistry and coagulation laboratory values will be flagged as "low", "normal", or "high" relative to the normal ranges provided by the laboratory, or "unknown". Laboratory values will also be flagged for abnormality (normal, abnormal – not clinically significant [NCS] or abnormal-clinically significant [CS]).

Urinalysis parameters are reported according to Roche Combur-10 dipstick leaflet. An overall investigator judgment for abnormality (normal, abnormal – not clinically significant [NCS] or abnormal-clinically significant [CS]) is also reported.

The observed values (at screening) of haematology, serum biochemistry and coagulation test results will be summarized by treatment arm and overall, using descriptive statistics, including mean, SD, min and max values. Also the investigator judgment for each parameter will be summarized based on the following categories: normal, abnormal-NCS and abnormal-CS.

The observed values (at screening) of urinalysis and pregnancy test (at screening and end of study) results will be presented in a frequency table by parameter, treatment arm and overall. For urinalysis , also the overall investigator judgment will be summarized based on the following categories: normal, abnormal-NCS and abnormal-CS.

A by-patient listing of all laboratory data and pregnancy test data will be presented.

Unscheduled laboratory values will only be presented in listings.

4.7.4 Vital Signs, Physical Findings and Other Observations Related to Safety

Vital signs (BP and HR) will be summarized by treatment arm and overall, for each time point, using descriptive statistics, including mean, SD, minimum and maximum values and also the abnormality categories: normal, abnormal-NCS and abnormal-CS. The change from baseline in each vital sign parameter will also be summarized.

Physical examination (at screening and at end of study visit), will be summarized by treatment arm and overall, using descriptive statistics, including the abnormality categories: normal, abnormal-NCS and abnormal-CS.

By-patient listings of all vital sign parameters, physical examination results and ECG results will be presented.

Unscheduled values will only be presented in listings.

4.8 Determination of Sample Size

A sample size of 230 patients per active arm is considered adequate to demonstrate the non-inferiority of DKP.TRIS/TRAM.HCl oral fixed combination compared to TRAM.HCl/paracetamol assuming a non-inferiority margin of 20%, a power of 80% and an overall significance level of 2.5 % (1-side). The mean $TOTPAR_6$ for TRAM.HCl/ paracetamol is assumed [redacted]

Assuming about 10% of major protocol violators, and 20% of screening failure rate a total of 640 patients need to be randomised and approximately 800 patients are expected to be screened.

128 patients in placebo arm are sufficient to demostrate the superiority of both TRAM.HCl/paracetamol and DKP.TRIS/TRAM.HCl vs placebo for model sensitivity

The sample size it is expected to be adequate to demostrate the superiority if a clinical significant benefit will be achieved by DKP.TRIS/TRAM.HCl vs. TRAM.HCl/paracetamol.

The following SAS procedure has been implemented for the calculation of sample size:

[redacted]

4.9 Changes in the Conduct of the Study or Planned Analysis

N/A

5 GENERAL DEFINITIONS

5.1 Data validation

[redacted] 2016 4.1 will be used, as Electronic Data Capture system for data entry, by site personnel and for data cleaning and data locking by the Menarini Data Management team.

5.2 Computer systems and software to be used in the analysis

SAS v.9.3 by SAS Institute Inc., Cary, NC, USA.
Database and SDTM and ADAM datasets will be created by using SAS Clinical Data Integration Studio version 4.9.
All tables and listings will be produced using PROC REPORT or procedure specific output displays using output delivery system (ODS) within SAS version 9.3 and using SAS Office analytics..

5.3 Coding systems used

5.3.1 Clinical Terms

Concomitant diseases, medical procedures, and AEs will be coded with MedDRA 18.1

5.3.2 Drugs

Drugs will be coded with WHODrug September 2015

5.4 Report type, language and format

The statistical output will be in Microsoft Word 2007 format and in English language.

- Dates will be presented as DDMMMYY.
- Numeric values will be decimal-point aligned.
- Counts and percentages: <group 1> XXX (XX.X%)
- Descriptive statistics:

N	XXX
Mean	XXX.X
SD	XXX.XX
Median	XXX.X
Minimum	XXX
Maximum	XXX

- Character values will be left aligned.

6 TABLES FIGURES AND LISTINGS

6.1 Statistical analysis report (TFL)

The TFL (Tables, Listings and Figures) will follow the list of tables, plots, and listings of section 6.3 and 6.4, which are intended to provide the overall idea of the general output and ordering of the TFL and will not necessarily be reproduced in the final TFL document.

6.2 Tables, headings, and footnotes

All tables stratified by treatment and gender will also contain the treatment overall column. The tables repeated for analysis population will be produced only once if the analysis populations are identical. All tables are provided for ITT population unless specified otherwise.

Line 1: Study code: DEX-TRA-06
Line 3: *Table/Listing/Figure n: Table name (Study population)*
Line 4: *Table/Listing/Figure n.n: Table name*
Line 5: *Table/Listing/Figure n.n.n: Table name* [if applicable]
Line 6: *Table/Listing/Figure n.n.n.n: Table name* [if applicable]
Line 7: *Table/Listing/Figure n.n.n.n.n: Table name* [if applicable]
Footnote: *Relevant notes (if any)*

6.3 List of Tables and Figures

1 Patient disposition

Note: all tables by treatment/overall for ITT population (where applicable)

- 1.1 Overall patient disposition
- 1.2 Presence of patients at study visits
- 1.3 Protocol violations by category

2 Baseline and general characteristics

Note: all tables by treatment/overall for ITT population (otherwise specified)

- 2.1 Demographics and pregnancy test
- 2.2 Medical history by MedDRA SOC and PT
- 2.3 Pain Intensity at rest scores at Qualification and T0h
- 2.4 Physical examination results –investigator judgement Prior to Surgery
- 2.5 Vital Signs Prior to Surgery and at Qualification
- 2.6 Safety Laboratory Test Prior to Surgery
- 2.7 Prior and concomitant medication by ATC-Code
  - 2.7.1 Prior medication by ATC-Code
  - 2.7.2 Concomitant medication by ATC-Code

3 Safety analysis

Note: all analyses will be done on safety population by treatment/overall, unless otherwise stated

- 3.1 Consumption of rescue medication
- 3.2 AEs
  - 3.2.1 Overview of AEs
  - 3.2.2 AEs by MedDRA SOC, PT
  - 3.2.3 AEs by MedDRA SOC, PT and intensity
- 3.3 TEAEs
  - 3.3.1 Overview of TEAEs
  - 3.3.2 TEAEs by MedDRA SOC, PT
  - 3.3.3 TEAEs by MedDRA SOC, PT and intensity
- 3.4 ADRs
  - 3.4.1 Overview of ADRs
  - 3.4.2 ADRs by MedDRA SOC, PT
  - 3.4.3 ADRs by MedDRA SOC, PT and intensity
- 3.5 Serious TEAEs
  - 3.5.1 Overview of Serious TEAEs by Investigator
  - 3.5.2 Serious TEAE (by Investigator) by MedDRA SOC, PT
  - 3.5.3 Serious TEAE (by Investigator) by MedDRA SOC, PT and intensity
  - 3.5.4 Overview of Serious TEAEs by Sponsor
  - 3.5.5 Serious TEAE (by Sponsor) by MedDRA SOC, PT
  - 3.5.6 Serious TEAE (by Sponsor) by MedDRA SOC, PT and intensity
- 3.6 Vital signs :
  - 3.6.1 Descriptive Statistics and Change Vs baseline at End of Study
  - 3.6.2 Investigator Judgement at End of Study
- 3.7 Physical examination:
  - 3.7.1 Descriptive Statistics and Change Vs baseline at End of Study
  - 3.7.2 Investigator Judgement at End of Study
- 3.8 Pregnancy test at End of Study

4 Efficacy analysis

Note: all tables by treatment/overall for ITT and PP population. For the estimation of non-inferiority the upper limit of 95% Confidence Interval of $TOTPAR_6$ difference will be compared with $\delta = 20\%$ of estimated mean of TRAM.HCl+Paracetamol.

- 4.1 Primary analysis
  - 4.1.1 Descriptive statistics
    - 4.1.1.1 Descriptive statistics for $TOTPAR_6$
    - 4.1.1.2 Descriptive statistics for $TOTPAR_6$ (PP Population)
  - 4.1.2 ANCOVA model for $TOTPAR_6$
    - 4.1.2.1 Test of Fixed Effects
      - 4.1.2.1.1 Test of Fixed Effects for $TOTPAR_6$
      - 4.1.2.1.2 Test of Fixed Effects for $TOTPAR_6$ (PP Population)
    - 4.1.2.2 Estimation and difference of Least Squares Means

4.1.2.2.1 Estimation and difference of LSM for $TOTPAR_6$
4.1.2.2.2 Estimation and difference of LSM for $TOTPAR_6$ (PP Population)

4.2 Secondary analysis
4.2.1 Continuous variables

4.2.1.1 Pain Intensity score
4.2.1.1.1 Descriptive statistics and change vs. Baseline by timepoint
4.2.1.1.2 ANCOVA model (longitudinal analysis)
4.2.1.1.2.1 Test of fixed effect
4.2.1.1.2.2 Estimation and difference of LSM
4.2.1.1.3 ANCOVA model (analysis by timepoint)
4.2.1.1.3.1 Test of fixed effect
4.2.1.1.3.2 Estimation and difference of LSM

4.2.1.2 SPIDs at 2, 4 ,6 and 8 hours
Note: For SPID6 analysis will be performed also on the PP population to check non-inferiority as for TOTPAR6
4.2.1.2.1 Descriptive statistics for SPIDs
4.2.1.2.2 ANCOVA model
4.2.1.2.2.1 Test of fixed effect for SPID6
4.2.1.2.2.2 Estimation and difference of LSM for SPID6
4.2.1.2.2.3 Test of fixed effect for SPID2,SPID4 and SPID8
4.2.1.2.2.4 Estimation and difference of LSM for SPID2,SPID4 and SPID8
4.2.1.2.3 Descriptive statistics of SPID6 (PP Population)
4.2.1.2.4 ANCOVA model (PP Population)
4.2.1.2.4.1 Test of fixed effect for SPID6
4.2.1.2.4.2 Estimation and difference of LSM for SPID6

4.2.1.3 %max SPIDs at 2, 4, 6 and 8
4.2.1.3.1 Descriptive statistics for %max SPIDs
4.2.1.3.2 ANCOVA model
4.2.1.3.2.1 Test of fixed effect
4.2.1.3.2.2 Estimation and difference of LSM

4.2.1.4 TOTPARs at 2, 4 and 8 hours
4.2.1.4.1 Descriptive statistics for TOTPARs
4.2.1.4.2 ANCOVA model
4.2.1.4.2.1 Test of fixed effect
4.2.1.4.2.2 Estimation and difference of LSM

4.2.1.5 %max TOTPARs at 2, 4, 6 and 8
4.2.1.5.1 Descriptive statistics for %max TOTPARs
4.2.1.5.2 ANCOVA model
4.2.1.5.2.1 Test of fixed effect

4.2.1.5.2.2 Estimation and difference of LSM

4.2.2 Categorical variables

4.2.2.1 Pain Relief score
4.2.2.1.1 Descriptive statistics and change vs. Baseline by timepoint
4.2.2.1.2 Wilcoxon rank-sum test

4.2.2.2 Patient Global Evaluation
4.2.2.2.1 Descriptive statistics
4.2.2.2.2 Wilcoxon rank-sum test

4.2.3 Binary variables

4.2.3.1 Pain Intensity Responders at 2, 4, 6 and 8 hours
4.2.3.1.1 Frequency table and percentage
4.2.3.1.2 Chi-Square Test by timepoint
4.2.3.1.3 GEE model
4.2.3.1.3.1 Test of fixed effect
4.2.3.1.3.2 Estimation and difference of LSM

4.2.3.2 TOTPAR responders at 2, 4, 6 and 8 hours
4.2.3.2.1 Frequency table and percentage
4.2.3.2.2 Chi-Square Test by timepoint
4.2.3.2.3 GEE model
4.2.3.2.3.1 Test of fixed effect
4.2.3.2.3.2 Estimation and difference of LSM

4.2.3.3 FPPAR within 30 minutes, 1 hour and 2 hours
4.2.3.3.1 Frequency table and percentage
4.2.3.3.2 Chi-Square Test by timepoint

4.2.3.4 Confirmed FPPAR within 30 minutes, 1 hour and 2 hours
4.2.3.4.1 Frequency table and percentage
4.2.3.4.2 Chi-Square Test by timepoint

4.2.3.5 MPAR within 30 minutes, 1 and 2 hours
4.2.3.5.1 Frequency table and percentage
4.2.3.5.2 Chi-Square Test by timepoint

4.2.3.6 Use of Rescue Medication within 2,4, 6 and 8 hours
4.2.3.6.1 Frequency table and percentage
4.2.3.6.2 Chi-Square Test by timepoint

Figures

6.4 Patient data listing

Key variables (KV) are the variables present in each patient data listing: Subject number, age, gender, planned treatment, study population (Yes/No for Safety, ITT and PP). Screening Failures will be excluded from the listing.
The following list is intended to provide the overall idea of the general output.

L.1 Patient disposition

KV, date/time of informed consent, date of screening, date/time of first study treatment, date of end of participation, reported term for the disposition event, standardized disposition term.

L.2 Demography

KV, race, race specification, ethnicity, country, childbering potential (yes/no), actual treatment

L.3 Baseline characteristics

KV, pain intensity at baseline, pain intensity at qualification.

L.4 Medical History

KV, start date of medical history event, end date of medical history event, ongoing, body system or organ class/reported term for the medical history, HLGT, HLT, LLT.

L.5 Previous and Concomitant Medication

KV, medication class, medication name, indication, dose form, dose, dosing frequency per interval, route of administration, start date of medication, end date of medication, ongoing, ATC1 (first level group) /ATC2 (second level group) /ATC3 (third level group), category (prior, concomitant, prior and concomitant)

L.6 Drug Exposure

KV, actual treatment, box number planned, box number received, start date/time of treatment

L.7 All Adverse Events

KV, date/time of first study treatment, actual treatment, treatment emergent analysis (yes/no), reported term for the adverse event, body system or organ class, causality, pattern of adverse event, concomitant or additional treatment given (yes/no), start date/time of adverse event, end date/time of adverse event, ongoing (yes/no), action taken with study treatment/outcome of adverse event, serious event (yes/no)/severity, lowest level term/high level term/high level group term

L.8 Laboratory results

KV, parameter, baseline (yes/no), category (coagulation test, haematology, serum chemistry, urinalysis), completion status, visit name (pre-surgery or unscheduled), investigator's judgement (normal, abnormal NCS or abnormal CS), date/time of specimen collection, result in standard format, range in standard units.

L.9 Pregnancy Test Data

KV, result test prior to surgery, date of specimen collection at baseline, result test at V1, date of collection at V1, result test at V4, date of collection at V4

L.10 Vital Signs Data

KV, parameter, baseline (yes/no), visit name (pre-surgery, surgery, end of study), investigator's judgement (normal, abnormal NCS or abnormal CS), actual treatment, date/time of measurements, result in standard units (standard units), change surgery-end of study.

L.11 Physical Examination Data

KV, visit name (pre-surgery, unscheduled, end of study), date of examination, actual treatment, investigator's judgement (normal, abnormal NCS, abnormal CS or NOT DONE)

L.12 PI-NRS (Numerical Rating Scale)

KV, PI value (original data),, timepoint, change from baseline, RMediary, RMeCRF, RMediaryLOCF, RMeCRFLOCF, missing.

L.13 All SPID Data

KV, SPID2, derivation type SPID2, SPID4, derivation type SPID4, SPID6, derivation type SPID6, SPID8, derivation type SPID8.

L.14 All Mean % max of SPID Data

KV, Mean % max of SPID2, derivation type Mean % max of SPID2, Mean % max of SPID4, derivation type Mean % max of SPID4, Mean % max of SPID6, derivation type Mean % max of SPID6, Mean % max of SPID8, derivation type Mean % max of SPID8.

L.15 All TOTPAR Data

KV, Total Pain Relief over 2 hours, derivation type TOTPAR2, Total Pain Relief over 4 hours, derivation type TOTPAR4, Total Pain Relief over 6 hours, derivation type TOTPAR6, Total Pain Relief over 8 hours, derivation type TOTPAR8.

L.16 All Mean % max of TOTPAR Data

KV, Mean % max of TOTPAR2, derivation type Mean % max of TOTPAR2, Mean % max of TOTPAR4, derivation type Mean % max of TOTPAR4, Mean % max of TOTPAR6, derivation type Mean % max of TOTPAR6, Mean % max of TOTPAR8, derivation type Mean % max of TOTPAR8.

L.17 PAR-VRS (Verbal Rating Scale)

KV, parameter (PAR), PAR value (original value), timepoint, RMediary, RMeCRF, RMediaryLOCF, RMeCRFLOCF, missing.

L.18 Patient Global Evaluation

KV, patient global evaluation, date/time of finding, actual treatment.

L.19 Rescue Medication

KV, date/time of first exposure to treatment, rescue medication required within 4 hours (yes/no), rescue medication required within 6 hours (yes/no), rescue medication required within 8 hours (yes/no), standardized medication name, ATC1, ATC3, dose form, start date/time of medication, end date/time of medication, comment.

L.20 FPPAR MPAR and CFPPAR

KV, time to FPPAR, censor FPPAR(yes/no), censoring description FPPAR, time to MPAR, censor MPAR(yes/no), censoring description MPAR, censor CFPPAR(yes/no), event or censoring description CFPPAR.

L.21 FSURG

KV, first tooth ID, time anaesthetic, time last suture, duration of surgery, number of teeth, sum of surgical trauma rating scale, second tooth ID, third tooth ID, fourth tooth ID.

7 REFERENCES

1. DEX-TRA 06 Study protocol, Final Version 2.0, November 24, 2015 , Menarini Ricerche, S.p.A

2. International Conference on Harmonisation ICH Topic E9: Statistical principles for clinical trials. 1998.
3. International Conference on Harmonisation ICH Topic E10: Choice of control group and related issues in clinical trials. 2001