Clinical Trial Protocol

Clinical Trial Protocol Number	AMZ001-007
Title	A Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Trial to Evaluate the Efficacy and Safety of Once Daily Diclofenac Gel AMZ001 [redacted] in the Treatment of Pain and Symptoms of Knee Osteoarthritis
Phase	3
IND Number	116375
EU CT Number	2024-517404-11-00
Sponsor	Amzell BV Name: [redacted] Address: Siriusdreef 31, Hoofddorp 2132WT, The Netherlands Telephone Number: [redacted]
Sponsor Legal Representative in the European Union	Amzell BV Siriusdreef 31 Hoofddorp 2132WT The Netherlands Telephone Number: +31 23 55 60 460
Clinical Trial Protocol Version	09 May 2025 Version 2.0
Replaces Version	26 March 2025 Version EU1.1

– Confidential –

Protocol Table of Contents

Table of In-Text Tables

Table of In-Text Figures

List of Abbreviations

AE	Adverse Event
ACR	American College of Rheumatology
ANCOVA	Analysis of Co-variance
AUC	Area Under the Curve
BMI	Body Mass Index
CI	Confidence Interval
CIPT	Cumulative Irritant Patch Test
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CRF	Case Report Form
CRO	Clinical Research Organization
COX	Cyclooxygenase
COXIB	Inhibitors of cyclooxygenase
CPSI	Chronic Pain Sleep Inventory
DCF-Na	Diclofenac Sodium
ECG	Electrocardiogram
eGFR	Estimated Glomerular Filtration Rate
EP	European Pharmacopoeia
EOT	End of Treatment
EQ-5D-5L	5-Level EuroQol-5 Domain Questionnaire
EQ-VAS	EQ Visual Analogue Scale
ET	Early Termination
FDA	Food and Drug Administration
GCP	Good Clinical Practice

GMP	Good Manufacturing Practice
IA	Intra-Articular
IB	Investigator's Brochure
ICE	Intercurrent Events
ICF	Informed Consent Form
ICH	International Conference on Harmonization
ICOAP	Intermittent and Constant Osteoarthritis Pain
IEC	Independent Ethics Committee
IMP	Investigational Medicinal Product
IRB	Institutional Review Board
ITT	Intention To Treat
IWRS	Interactive Web Response System
LOE	Lack of Efficacy
LPLV	Last Participant's Last Visit
LSR	Local Skin Response
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not At Random
mITT	Modified Intention to Treat
NRS	Numerical Rating Scale
NSAID	Non-steroidal Anti-inflammatory Drug
OA	Osteoarthritis

OMERACT-OARSI	Osteoarthritis Research Society International Standing Committee for Clinical Trials Response Criteria Initiative and the Outcome Measures in Rheumatology Committee
OTC	Over the Counter
PGA	Patient Global Assessment
PK	Pharmacokinetic
PRO	Patient Reported Outcome
QoL	Quality of Life
RIPT	Repeat Insult Patch Test
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SLS	Sodium lauryl sulphate
SNRI	Serotonin- and Norepinephrine Reuptake Inhibitor
SUSAR	Suspected Unexpected Serious Adverse Reaction
USP	United States Pharmacopoeia
WOMAC	Western Ontario and McMasters University Osteoarthritis Index

1 Synopsis

Clinical Trial Protocol Number	AMZ001-007
Title	A Randomized, Double-Blind, Multi-Center, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Once Daily Diclofenac Gel AMZ001 [redacted] in the Treatment of Pain and Symptoms of Knee Osteoarthritis
Trial Phase	3
IND Number	116375
FDA covered trial	Yes
EudraCT Number	2024-517404-11-00
Sponsor	Amzell BV [redacted] Address: Siriusdreef 31, 2132WT Hoofddorp, The Netherlands
Sponsor Legal Representative in the European Union	Amzell BV Siriusdreef 31 2132WT Hoofddorp The Netherlands Telephone Number: +31 23 55 60 460
Trial centers	Approximately 25 sites
Planned trial period (first participant in-last participant out)	First participant in: Jan 2025 Last participant out: Sep 2025

Objectives:

Primary Objective

- To compare the efficacy of diclofenac gel AMZ001 versus placebo for the treatment of pain, in terms of change in pain intensity evaluated by the WOMAC pain sub-score of the target knee, at week 2

Secondary Objectives

- To evaluate changes in pain, in terms of the WOMAC pain sub-score of the target knee, at weeks 1, 3, 4, and 6
- To evaluate changes in daily pain, in terms of the 11-point Pain Numeric Rating Scale (NRS) of the target knee until week 6.
- To evaluate the onset of action of AMZ001, based on self-rated average pain scores in terms of the 11-point NRS 1 hour, 4 hours, 12 hours, 24 hours after the first drug application
- To evaluate changes as measured by Physician Global Assessment (PGA) at week 3 and 6
- To evaluate changes in sleep, as measured by the Chronic Pain Sleep Inventory (CPSI), at weeks 3 and 6
- To evaluate changes in physical functioning, in terms of the WOMAC function sub-score, at weeks 1, 2, 3, 4 and 6
- To evaluate symptoms of osteoarthritis in terms of the WOMAC total score of the target knee, at weeks 3 and 6
- To evaluate changes in quality of life, as measured by the EQ-5D-5L, at weeks 3 and 6
- To evaluate changes as measured by the Osteoarthritis Research Society International Standing Committee for Clinical Trials Response Criteria Initiative and the Outcome Measures in Rheumatology Committee (OMERACT-OARSI) responder criteria, at weeks 3 and 6
- To evaluate the use of rescue medication
- To evaluate the safety and tolerability of AMZ001

Methodology:
This is a multi-center, randomized, double-blind, placebo-controlled, parallel group, 6-week trial [redacted] diclofenac gel AMZ001 once daily versus [redacted] placebo once daily. Approximately 540 participants will be randomized in a 1:1 to one of the two double-blind treatment groups:

1. AMZ001 Diclofenac Gel [redacted] as 2 pump actuations [redacted] once in the morning, for each knee affected [redacted]

2. Placebo Gel, [redacted] as 2 pump actuations [redacted] once daily in the morning per knee affected.

The double-blinded treatment phase consists of 42 days of treatment with one baseline visit (Day 1), 5 visits (weeks 1, 2, 3, 4, and 6) and safety follow-up phone call visit (week 8).

During the treatment period, the participants will apply IMP gel on the treated knee joint(s) once daily at approximately the same time each day (every 24 ±1 hours), in the morning.

For the purpose of facilitating standardization in study procedures and relevant sampling in the study, participants are required to apply the IMP gel in the morning.

The statistical analysis will be performed at the end of the double-blinded treatment phase. Results from the treatment phase will be analyzed and reported. Study staff and participants will remain blinded during the study.

Brief description of trial procedures

After written informed consent is obtained, participants will be screened for eligibility. The screening period consists of two visits 1a and 1b. During the Screening Period, participants will first undergo a medical history interview to verify that they meet demographic and clinical criteria for inclusion into the study. Participants will self-evaluate pain, after adequate wash-out of any pain medication, on a Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain sub-score (five questions) questionnaire for both knees and hips. If both knees meet radiographic and pain criteria, the reported pain score of the contralateral knee should not exceed that of the target knee. If hip pain of either hip exceeds that of the target knee, the participant will not be enrolled into the study.

A physical examination including vital signs, measurement of sitting diastolic and systolic blood pressure, heart rate, and measurement of height and weight will be performed. Participants not meeting the defined criteria will not be enrolled into the study.

X-ray images of the knees will be obtained using a Fixed-Flexion frame (Synaflexor© or similar) and a central radiologist reader will grade osteoarthritis on the Kellgren-Lawrence scale.

Blood will be collected for hematology and biochemical safety analysis. Pregnancy test will be performed for women of child-bearing potential at screening and baseline visits.

Upon meeting all eligibility criteria and completing all screening assessments, a target knee will be assigned, and participants will be randomized into one of two groups. Presence of bilateral OA of the knees is allowed. If both knees fulfil the diagnostic, radiographic, and symptomatic inclusion criteria, both knees should be treated with IMP throughout the treatment period, but the more painful knee is selected as the target knee. If significant pain of the contralateral knee arises during the course of the trial, IMP treatment should be initiated on the contralateral knee in addition to the target knee and continued on both knees through the remainder of the trial. Initiation of treatment of the contralateral knee should only be initiated at study visits, and not between study visits.

Participants will be provided with rescue medication (paracetamol tablets) for use in case of intolerable pain during the entire trial period. Non-rescue analgesic medication is prohibited during the trial period.

In the trial, WOMAC, and other questionnaires will be completed on Day 1 (Baseline), and at weeks 1, 2, 3, 4, and 6, or as indicated in the schedule of assessments. Prior to patient reported outcomes at study visits, participants will be trained on the placebo response reduction mitigation script.

In order to "wash out" the effect of rescue medication, the participants are not allowed to take rescue medication within 24 hours prior to any study visit.

Planned number of participants:
Approximately 540 participants will be randomized. 270 will receive AMZ001 once daily, and 270 will receive placebo once daily. It is estimated that approximately 20% of randomized participants will discontinue the trial during the double-blind placebo-controlled trial period.

Primary endpoint:
The primary endpoint of this trial is the change from baseline in the WOMAC pain sub-score (questions 1 to 5) in the target knee as evaluated at week 2.

Key secondary endpoints:

- Change from baseline in the WOMAC pain sub-score at week 1, in the target knee
- Change from baseline in daily pain at Day 8, Day 7, Day 6, Day 5, Day 4, Day 3, Day 2, and Day 1 using an 11-point Numeric Rating Scale (NRS)

Other Secondary Endpoints:

- Partial AUCs of the observed change from baseline based on pain recorded at 24 hours, 12 hours, 4 hours and 1 hour, after the first drug application using an 11-point Numeric Rating Scale (NRS)
- Change from baseline in PGA at week 3 and 6
- Change from baseline in the Chronic Pain Sleep Inventory at week 3 and 6
- Change from baseline in the weekly average of daily pain recorded using an 11-point pain NRS until week 6 in the target knee
- Change from baseline in the WOMAC pain sub-score during the trial in the target knee (week, 3, 4 and 6)

- Change from baseline in the WOMAC function sub-score at during the trial, (weeks 1, 2, 3, 4 and 6)
- Change from baseline in the WOMAC total score at weeks 3 and 6
- Changes from baseline in quality of life as assessed using the EQ-5D-5L at weeks 3 and 6
- Proportion of participants meeting the OMERACT-OARSI responder criteria at weeks 3 and 6
- Proportion of participants meeting at least a 30 % reduction from baseline in WOMAC pain sub-score at Week 3 and 6.
- Proportion of participants meeting at least a 50 % reduction from baseline in WOMAC pain sub-score at Week 3 and 6.
- Time (days) to first use of rescue medication since baseline
- Average daily amount of rescue medication taken between visits
- Weekly proportion of days with rescue medication use at Weeks 1-6

Safety endpoints:

- Nature, incidence and severity of adverse events (AEs)
- Changes in laboratory safety parameters, vital signs, 12-lead electrocardiogram (ECG) parameters, and body weight.
- Nature, incidence, and severity of skin reactions at the application site

Inclusion and exclusion criteria:

Inclusion Criteria:

1. Participant is able to read and understand the language and content of the study material, understand the requirements for study visits, and is willing to provide information at the scheduled evaluations and appropriate written informed consent has been obtained.
2. Femorotibial osteoarthritis (OA) of the knee, according to the American College of Rheumatology (ACR) clinical and radiographic criteria (Altman et al. 1986)
3. Radiological OA grade 2, or 3 of the target knee, using the Kellgren-Lawrence method (Kellgren & Lawrence 1957) as graded by central, independent reading of X-ray

obtained during screening, or on a recent (within 6 months) X-ray image which fulfils the specifications for central reading.

4. Age between 40 years and 85 years at the time of screening, both included of either sex.

5. Pain score rated on an 11-point numerical rating scale of the target knee of ≥ 20 out of 50 in response to the WOMAC pain sub-score (5 questions), at the time of screening and baseline.

6. The WOMAC pain sub-score on the contralateral knee must not exceed the one on the target knee, regardless of the eligibility of the contralateral knee , at the time of screening and baseline.

7. At screening Visit 1a, participants report that their typical OA knee pain in one or both knees when not using medication is ≥ 4 out of 10.

8. Daily OA knee pain diary average numerical rating scale (NRS) score of ≥ 4 and ≤ 9 in the target knee, for the 7 days immediately preceding baseline (Day 1). The average calculation is based on the recorded scores during this entire period with a requirement of at least 4 days of data recorded.

9. Women of child-bearing potential must use a highly effective method of contraception (hormonal contraceptives, intrauterine device, vasectomy, bilateral tubal occlusion, sexual abstinence) from enrolment up to at least 3 months after the study end. Postmenopausal status is defined as being amenorrheic for at least 1 year prior to screening. Sexually active men with a female partner of childbearing potential must agree to use condom from enrolment up to at least 3 months after the study end.

10. Knee pain in the target knee for at least 14 days of the preceding month (periarticular knee pain due to OA and not due to non-OA conditions such as bursitis, tendonitis, etc.) based on participant report.

11. Except for OA, the participant is in reasonably good health as determined by the Investigator.

Exclusion criteria:

1. Known or suspected hypersensitivity to or previous hypersensitivity reactions to diclofenac or of the excipients in either of the investigational products.

2. Patients in whom asthma, angioedema, urticaria, or acute rhinitis are precipitated by acetylsalicylic acid or other non-steroidal anti-inflammatory drugs (NSAIDs).

3. Any physical impediment to gel application on the target knee.

4. Intra-articular delivery of corticosteroids or hyaluronic acid in the target knee within 6 months of screening or into any other joint within 30 days of screening.

5. High dose (equivalent to > 10 mg of prednisone/day) systemic corticosteroid treatment of more than 14 days during the past 6 months prior to screening.

6. Major surgery/ arthroscopy of the target knee within the previous year prior to screening or aspiration of effusion of the target knee within 12 weeks prior to enrollment.

7. Planned surgery of the target knee within the next 12 months.

8. Use of a currently unapproved investigational drug, device or biologic within 3 months prior to screening.

9. Presence of concomitant non-osteoarthritic disease affecting either knee, such as rheumatoid arthritis, psoriasis, gout or clinically relevant pseudogout, if there is reason to believe that the disease(s) may significantly interfere with the interpretation of the clinical response to the study drug.

10. Perioperative period in the setting of coronary artery bypass graft surgery.

11. Current malignancy or treatment for malignancy within the past five years, with the exception of non-melanoma skin cancer, unless affecting the target knee area, or carcinoma in situ events.

12. Any other abnormal laboratory results or significant medical conditions that the Investigator believes should preclude the participant's participation in the trial.

13. Secondary OA of the target knee, previous procedures or trauma affecting joint homeostasis including total meniscectomy or septic arthritis, or any other serious condition leading to secondary OA of the target knee.

14. Reported incidence of any of the following diseases: known OA of the hip(s) if pain in either or both hip(s) exceeds that of the target knee using the WOMAC hip pain sub-score, or presence of significant radicular back pain.

15. Presence of a defined cut-off of pain variability during the screening period of 1.5 out of 10:

    -Pain Numeric Rating Scale (NRS) scores must be recorded for the target knee on at least 4 out of the 7 days immediately preceding baseline (Day 1)

-Observed standard deviation of the Daily osteoarthritis (OA) knee pain diary average NRS intensity score for 7 days immediately preceding baseline (Day 1) must not exceed 1.5

16. Estimated glomerular filtration (eGFR) rate < 30 mL/min using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI 2021) formula.
17. Generalized skin irritation, previous skin reactions upon use of topical NSAIDs, current skin irritation or redness at the planned site of gel application, or significant skin disease including psoriasis, as judged by the investigator.
18. Use of any topical medication on the planned application site within 15 days of the time of randomization.
19. Any use of opioid medication within 6 weeks before the screening visit.
20. Use of duloxetine, pregabalin, or gabapentin within 4 weeks before the screening visit.
21. History of alcohol or drug abuse within the past year prior to randomization.
22. Pregnant and breastfeeding women are excluded for participation.

Investigational Medicinal Products:

Double-blind therapy:
Diclofenac sodium [redacted] (AMZ001) gel or matching placebo gel in an 87 g metered dose dispenser, [redacted] gel to be applied once daily at approximately the same time each day (every 24 ±1 hours) in the morning.

Rescue medication:
Acetaminophen/Paracetamol tablets 500 mg. The dosage of acetaminophen/paracetamol that the participants will be allowed to take per day will be defined according to the standard of care in the countries where the trial will be carried out; however, the maximum dose should not exceed 1 gram per dose and 4 grams per day.

Planned trial and treatment duration per participant:
The treatment phase consists of 6 weeks (42 days) of treatment with one Baseline Visit (Day 1), 5 visits at weeks 1, 2, 3, 4, and 6, and safety follow-up phone call visit (week 8). The screening period is up to 28 days from the time of providing written consent. In total, the expected duration of the trial participation of a randomized participant will be up to 84 days.

Statistical methods:
Two primary estimands addressing different aspects of the trial objectives will be applied to the analyses of the primary and secondary efficacy endpoints, focusing on estimating "treatment attributable" effect of trial product, one for continuous and one for binary endpoints.

Additionally, two secondary estimands with the aim of estimating "efficacy" of trial product will also be applied to the primary endpoint and potentially to the secondary efficacy endpoints.

The primary estimand for the primary endpoint of change from baseline in WOMAC pain sub-score at Week 2 will be the "treatment attributable" estimand. The analysis model will be ANCOVA of change from baseline at the corresponding visit with baseline as covariate and treatment group & pooled center as covariates.

Data after treatment discontinuation (for treatment related reasons) will be discarded (hypothetical strategy), and imputed under Missing Not At Random (MNAR) (return-to-baseline). Same will be done for data impacted by use of prohibited medication/non-permitted therapies with high expected impact on efficacy evaluation. Missing data arising from study withdrawals will be handled based on relation to treatment, either MNAR return-to-baseline if possibly related otherwise MAR. Prohibited medication use with low expected impact or use of rescue medication will be used as is (treatment policy). Excessive use of rescue medication will be treated as a case of prohibitive medication use and classified based on anticipated effect on efficacy evaluation. Missing data will be imputed using multiple imputation, based on sequential use of an ANCOVA model for Week 1 and 2 (and later weeks for Secondary Endpoints). The imputations within each treatment arm will be based on a model including baseline value, pooled center (if feasible), and values from previous weeks' models.

For each of the imputed datasets, the endpoint will be analyzed using an ANCOVA model adjusted for baseline value, treatment group and pooled center. The estimates and standard errors from all the imputed datasets analyses will be combined using Rubin's rule to form a unique point estimate and standard error. The estimated treatment difference to placebo will be presented with a 95% confidence interval (CI).

The power calculation is based on data from a previous 4-week double-blind study of AMZ001 QD or BID vs. placebo in a study population highly similar to that proposed in the current trial.

The assumed difference between AMZ001 QD and placebo is based on a sub-group analysis of AMZ001-006, consisting of participants with WOMAC pain scores of ≥ 20 and ≤ 45 out of 50 at both screening AND baseline visits. In AMZ001-006, this subgroup comprised 89.4% of the total (mITT) population.

Assuming a common standard deviation of 19.1 (out of 100) in the WOMAC pain sub-score, analysis of 216 evaluable subjects per treatment group will result in minimally detectable difference between placebo and active treatment of 5.9 (out of 100) with a two-sided 5% level of significance, with 90% power. As all participants randomized to receive IMP are included in the efficacy analysis (ITT population), this estimate is considered a conservative estimate.

Table 1: Schedule of Trial Procedures

Activity/Assessment	Screening 1a[1]	Screening 1b[1]	Baseline V2	Double-blind Treatment Phase					V8 Follow-up (phone visit)
				V3	V4	V5	V6	V7 (EOT) / Early termination (ET)	
Study Week	-	-	**0**	**1**	**2**	**3**	**4**	**6**	**8**
Study Day	**-28 to – 7[2]**	**-13 to -7[2]**	**1**	**8**	**15**	**22**	**29**	**43**	**57**
Visit Window (days)	-	-	**0**	**+/- 2**	**+/- 2**	**+/- 2**	**+/- 2**	**+/- 3**	**+/- 2**
Informed Consent	X								
Medical history	X								
Inclusion/exclusion criteria	X	X	X						
Collection of "typical OA pain when not taking analgesics"	X								
Randomization			X						
Demographic data	X								
Physical examination	X								
Collection of substance use	X								
Knee X-ray (both knees)	X								
Placebo Response Mitigation [4]	X[3]	(X)	X	X	X	X	X	X	
WOMAC pain sub-scale questionnaire both knees and hips [5]	X[3]	(X)							
WOMAC Questionnaire			X[6]	X	X	X	X	X	
CPSI			X			X		X	
PGA			X			X		X	
First day pain-NRS (5 times) [7]			X						
Daily Pain NRS		→	→	→	→	→	→	→	
Rescue medication diary			→	→	→	→	→	→	
EQ-5D-5L			X			X		X	
Selection of target knee			X						
Height and weight	X							X[8]	
Vital signs	X		X	X	X	X	X	X	
12-lead ECG	X							X	
Adverse events	X[9]	X	X	X	X	X	X	X	X

Activity/Assessment	Screening 1a[1]	Screening 1b[1]	Baseline V2	Double-blind Treatment Phase					V8 Follow-up (phone visit)
				V3	V4	V5	V6	V7 (EOT) / Early termination (ET)	
Study Week	-	-	**0**	**1**	**2**	**3**	**4**	**6**	**8**
Study Day	**-28 to – 7[2]**	**-13 to -7[2]**	**1**	**8**	**15**	**22**	**29**	**43**	**57**
Prior and concomitant medication	X	X	X	X	X	X	X	X	X
Study drug dispensation			X	X	X	X	X		
Dispensation of IMP "instructions for use" pamphlet			X						
Instruction of study drug application and frequency of use			X						
Skin tolerability assessment[10]			X	X	X	X	X	X	
Application of study drug on site under staff supervision			X	X	X	X	X		
Weighing of kit(s) for compliance assessment (IMP accountability)[11]			X	X	X	X	X	X	
Rescue medication dispense			X	(X)[12]	(X)[12]	(X)[12]	(X)[12]		
Rescue medication accountability/collection empty bottles[13]				(X)	(X)	(X)	(X)	X	
Rescue medication Wash-out	(X)	(X)	X	X [14]	X [14]	X [14]	X [14]	X [14]	
Hematology	X[15]		X					X	
Biochemical safety analysis	X		X					X	
Pregnancy test[16]	X		X					X	

EQ-5D-5L: 5-level EuroQol-5 Domain. NRS: Numeric Rating Scale. PGA: Patient Global Assessment. CPSI: Chronic Pain Sleep Inventory. VAS: Visual analogue scale. WOMAC: The Western Ontario and McMaster Universities Osteoarthritis Index.

1. If participant is washed out of analgesics at Screening Visit 1a, the examinations scheduled for screening visits 1a and 1b can be performed on the same day but at least 7 days before baseline.
2. If visit 1a and 1b are performed on the same day, the visit window that should be followed is that of visit 1a.
3. Assessed without analgesic medication for five half-lives of the analgesic. If no analgesic washout at visit, this assessment to be done at visit 1b.
4. Before any pain questionnaire is completed, site staff will read a standardized script, designed with the purpose of improving the accuracy of pain reporting, and managing the participant's expectation of treatment benefit.
5. WOMAC pain subscale assessment of hips should only be performed at the screening visit.
6. WOMAC for baseline visit will be done on both knees to assess eligibility.
7. Pain NRS at baseline, 1 hour, 4 hours, 12 hours, 24 hours after drug application on the day of the first IMP administration.
8. Only weight at Visit 7.
9. Conditions previously unknown to the participant at the time of the screening visit will be recorded as medical history. All subsequent events will be recorded as adverse events.
10. Skin assessment to be performed prior to application of study drug at every visit. Skin tolerability assessment includes the local skin response grading scale and the skin tolerability symptoms.
11. Any suspension of the study treatment must be recorded in the eCRF.
12. To be dispensed if previously dispensed supply of rescue medication was depleted and empty packaging returned to site.
13. Collection of remaining rescue medication at Visit 7. Empty packaging of previously dispensed kit should be returned to the site if new rescue medication is requested.
14. Rescue medication washout of 24 hours prior to every visit.
15. At screening INR will be measured.
16. Serum test at screening visit, urine pregnancy test at all other visits.

2 Sponsor, Investigators and Trial Administrative Structure

Sponsor

The clinical trial will be sponsored by: Amzell BV, Siriusdreef 31, Hoofddorp 2132WT, The Netherlands.

Sponsor Responsibilities

The sponsor will be responsible for overseeing all aspects of the study.

Investigators and Trial Sites

The trial will be conducted at approximately 25 trial sites.

Signature pages for the Protocol Lead and the Coordinating Investigator as well as a list of Sponsor and CRO responsible persons are in Appendix E-G.

The trial will appear in the following clinical trial registries: clinicaltrials.gov, euclinicaltrials.eu.

Details of structures and associated procedures will be defined in separate, study specific documents.

3 Background Information

3.1 Osteoarthritis Pain

Osteoarthritis (OA) is the most common form of arthritis and is one of the leading causes of disability in the world, with more than 10% of the population aged 50 and above having symptoms of the disease (Abramson *et al.* 2006, Bijlsma *et al.* 2011). The hallmarks of the disease are joint pain, reduced physical function, and progressive degeneration of articular cartilage. Drugs to reduce the symptomatic burden of OA include oral and topical non-steroidal anti-inflammatory drugs (NSAIDs), paracetamol, opioids, and intra-articular (IA) corticosteroid injections. Existing oral NSAID therapies are associated with risks of serious complications of chronic use (McAlindon *et al.* 2014), while increased, long-term use of opioids has led to an increase in patients with opioid addiction, a situation described as an "opioid crisis" in the United States of America (Frenk *et al.* 2015). The use of IA corticosteroid injection is considered efficacious but with limited duration of action (McAlindon *et al.* 2014), lack of long-term symptomatic benefit, and may lead to deleterious acceleration of cartilage loss (McAlindon *et al.* 2017). Novel biological agents targeting nerve growth factor are being developed but have been found to be associated with increased risks of rapidly progressive osteoarthritis (Dakin *et al.* 2019, Schnitzer *et al.* 2019). The topical application of NSAIDs, including diclofenac sodium, is considered a viable and effective alternative treatment option, with a lower likelihood of users experiencing systemic adverse events owing to the low levels of systemic absorption following topical application (Brunner *et al.* 2005). However, topical formulations require several daily applications of considerable amounts of gel to achieve a modest efficacy (Bookman *et al.* 2004, Barthel *et al.* 2009, Baraf *et al.* 2010). Thus, development of new and improved locally applied analgesics with a faster onset and longer duration of action, while requiring fewer applications of a lower volume of gel to be applied, is needed.

3.2 AMZ001 (Diclofenac Sodium [redacted])

NSAIDs, including diclofenac, inhibit cyclooxygenase (COX), which reduces pain and inflammation through the inhibition of prostaglandins. However, the COX enzyme is also present in gastric mucosa where it stimulates gastroprotective prostaglandins. Systemic long-term unselective blockade of COX-1 and COX-2 leads to risks of developing gastric ulcer, but increased risks of cardiovascular and renal complications have also been described (Whelton 2000). The clinical utility of NSAIDs in treatment of osteoarthritis pain is well recognized and the first diclofenac analgesic was introduced in 1973. Several formulations for different routes of administration of diclofenac sodium have been developed, including topical formulations. Diclofenac as a molecule possesses excellent physicochemical properties for skin and deep tissue penetration (Cordero *et al.* 1997, Cordero *et al.* 2001). Although systemic absorption of topically applied diclofenac is low, the concentration of diclofenac has been reported to be higher in the local muscle tissue after topical administration compared with systemically administered diclofenac (Brunner *et al.* 2005). Clinical trials of currently approved diclofenac sodium (DCF-Na) 1% gel formulations found that four daily gel applications were required to demonstrate a statistically significant efficacy compared to placebo (Barthel *et al.* 2009, Baraf *et al.* 2010).

A novel diclofenac sodium gel formulation, AMZ001, has been developed with the aim of improving the onset and duration of action, as well as user preference and satisfaction, by

reducing the required daily frequency of gel application. The gel is a homogeneous, transparent, non-staining hydroalcoholic gel containing [redacted] diclofenac sodium. The excipients of the gel are well known and have been used in previously approved topical products.

3.3 Summary Of Relevant Nonclinical Data

Non-clinical data comparing commercial Voltaren® (diclofenac sodium 1% gel) with AMZ001 assessing the permeation rate of human skin samples indicates that the permeation rate of AMZ001 was up to two times higher than that of Voltaren® gel (Figure 1).

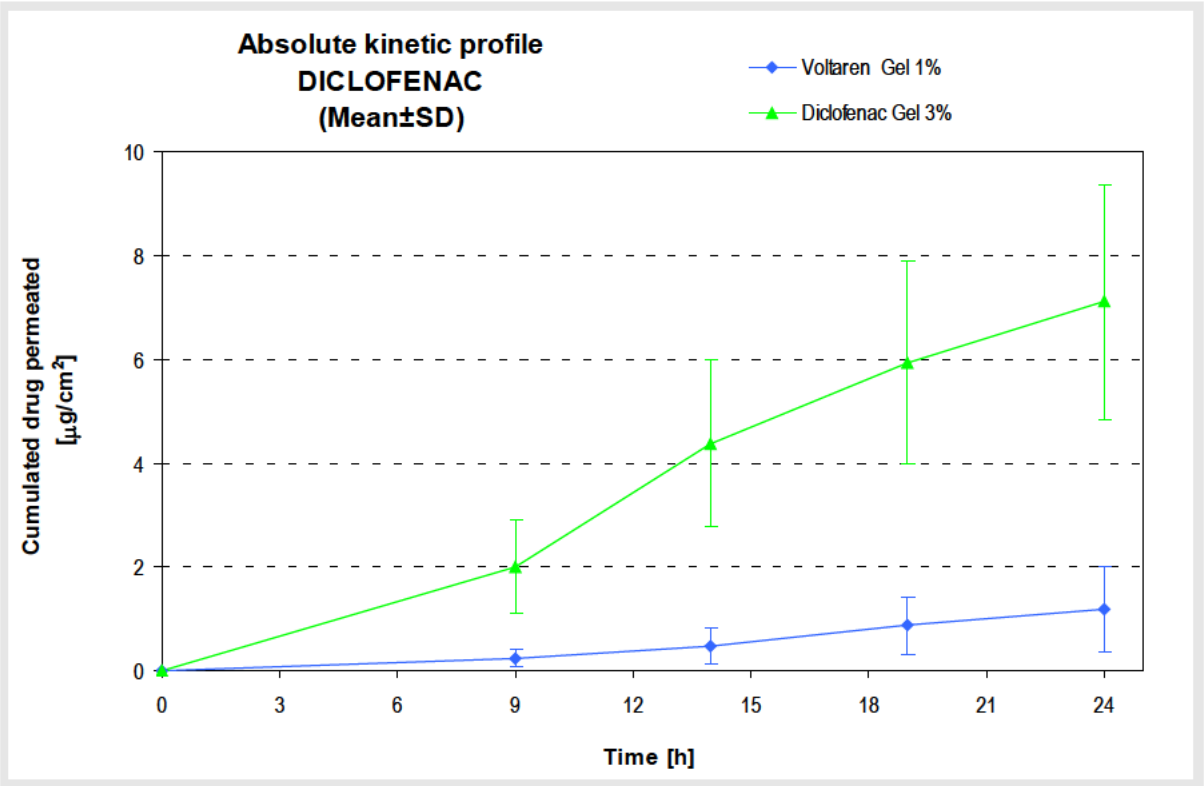


Figure 1: *In vitro* permeation study comparing AMZ001 and Voltaren® (diclofenac sodium 1% gel) using frozen human skin

As the pharmacology of diclofenac sodium is considered well-known, no nonclinical pharmacology or pharmacotoxicological experiments have been conducted with AMZ001.

The toxicology assessments of AMZ001 have instead been focused on assessing the dermal toxicology of AMZ001. The local tolerability of AMZ001 was evaluated in rabbits and minipigs, and skin sensitization tests were carried out in guinea pigs. Results showed that the formulation [redacted] was irritating in the carpal and tarsal joints of the rabbit and minipig, and that skin sensitization occurred in the guinea pig, while application of gel to the dorsal region of the minipig was well tolerated. Refer to the Investigator's Brochure for a detailed description.

3.4 Summary Of Relevant Clinical Data

One phase 2 trial and four phase 1 trials with AMZ001 have been completed: A 21-day, randomized, controlled study to evaluate the skin irritation potential of AMZ001, using a Cumulative Irritant Patch Test (CIPT) design, in 35 healthy participants (AMZ001-001); a randomized, controlled study to evaluate the sensitizing potential of AMZ001, using a Repeat Insult Patch Test (RIPT) design in 200 healthy participants (AMZ001-002), two PK studies (000084 and AMZ001-005) in healthy participants and one phase 2 proof-of-concept study (AMZ001-006) .

AMZ001-001 was a randomized, single-center, controlled, evaluator-blinded, within-participant comparative study of the investigational product AMZ001 and its drug-free vehicle, Pennsaid® 2% (comparator product), and positive (commercially available 0.5% sodium lauryl sulfate [SLS]) and negative (0.9% saline) controls under semi-occlusive conditions. The primary objective of this study was to determine the potential of AMZ001 to cause skin irritation after repeated topical applications under controlled conditions.

No participants discontinued the patch applications due to irritation and no AEs that could be considered study product related were reported. Under the exaggerated conditions of this 21-day cumulative irritation study, AMZ001, AMZ001 vehicle, and Pennsaid® 2% were considered safe but poorly tolerated.

In AMZ001-002, the potential of AMZ001, AMZ001 vehicle, and Pennsaid® 2% diclofenac gel to induce sensitization by repeated topical application onto healthy skin was determined in 200 healthy volunteers. The investigational products with occlusive patches were applied 9 times to adjacent sites on the infrascapular region on the back during an induction phase of three weeks, followed by a resting phase with no application of IMP. After the resting period, the participants were challenged by one 48-hour patch application of the same product to a naïve skin area. Dermal reactions at the application sites were assessed clinically using a visual scale that rates the degree of erythema, edema, and other signs of cutaneous irritation.

No participants in the study experienced a reaction indicative of sensitization, as classified by the PI, to AMZ001, AMZ001 vehicle, Pennsaid® 2%, 0.2% sodium lauryl sulfate (SLS), or 0.9% saline.

In conclusion, under the conditions of the study, the tested products (AMZ001 and AMZ001 vehicle) were safe, and no participants exhibited reactions consistent with sensitization to the tested products.

Study 000084 was designed to compare the PK of three different doses of AMZ001 [redacted] to Voltaren® (32.0 g of Voltaren containing 320 mg of DCF-Na) topical gels in 12 healthy volunteers after a 5-day bilateral treatment regimen. Participants were treated according to a fixed-sequence schedule (A-B-C-D) with a 16-day washout period between treatments.

The study concluded that AMZ001 proved to be an acceptable pharmaceutical alternative to Voltaren® (diclofenac sodium 1% gel). The 4.6 g AMZ001 application achieved a similar concentration of diclofenac systemically as Voltaren® gel with the benefit of fewer daily applications (twice daily vs. four times daily). AMZ001 9.2 g/day had fewer daily applications

(twice daily vs. four times daily), a 2-fold higher concentration of diclofenac systemically compared with Voltaren® gel, and a good safety profile with no skin reactions.

Study AMZ001-005 was designed to determine the exposure of diclofenac in healthy participants following four different dose-frequency combinations of AMZ001 (diclofenac gel [redacted]) in comparison with marketed product Voltaren® (diclofenac sodium 1% gel) after repeated topical application. Seventy-five (75) participants were included in the study and were randomized to parallel groups of fifteen (15) participants to one of five treatments.

The study concluded that the exposure to diclofenac was not significantly different with AMZ001 compared to Voltaren®, despite using AMZ001 in doses lower than those of Voltaren®. In addition, the mean for the $AUC_{(0\text{-}24h)}$ at treatment E - [redacted] AMZ001 [redacted] applied as 2 pump actuations per knee twice daily (BID) on both knees - was significantly different from that at treatment C - [redacted] AMZ001 [redacted] applied as 2 pump actuations per knee once daily (QD) in the morning on both knees.

Study AMZ001-006 was a phase 2 proof-of-concept, placebo-controlled, double-blind, dose-ranging trial to evaluate the efficacy and safety of AMZ001 [redacted] once or twice daily compared to placebo during 28 days. The trial included a single-blind Voltaren (diclofenac sodium 1% gel) four times daily comparator.

In this trial, both once- and twice-daily application of AMZ001 gel was found to perform nominally statistically significantly better than placebo on several measures of pain and quality of life, whereas no significant differences between the two application frequencies of AMZ001 were found. Thus, further confirmatory research for establishment of the efficacy and safety of AMZ001 once and twice daily is warranted (Bihlet *et al.* 2020).

One single center phase 1, three period cross over, multidose trial comparing the bioavailability and safety of two dose levels of AMZ001 and one dose level of Voltaren Arthritis Pain (diclofenac sodium 1% gel) in healthy participants has been completed and the study report is ongoing (AMZ001-008).

Refer to the Investigator's Brochure (IB) for further information about the nonclinical and clinical programs and Guidance for the Investigator.

This clinical trial will be conducted in compliance with the clinical trial protocol, International Conference on Harmonization Good Clinical Practice (ICH GCP), EU Clinical Trials Regulations (EU-CTR) and any additional applicable regulatory requirements.

Based on the available nonclinical and clinical data to date, the conduct of the trial specified in this protocol is considered justifiable.

3.5 Benefits and Risks of Participation.

This clinical trial will be conducted in compliance with the clinical trial protocol, ICH GCP, and all additional applicable regional regulatory requirements.

Osteoarthritis is the most common joint disorder and cause of chronic disability in the US. The number of individuals directly impacted by symptomatic OA is projected to increase substantially due to the aging population and rising obesity (Zhang Y, 2010). Currently no disease-modifying treatments are available for OA, and available symptom-directed treatments include oral and topical non-steroidal anti-inflammatory drugs (NSAIDs), paracetamol, opioids, and IA corticosteroid injection. While these are either considered modestly efficacious (paracetamol), associated with serious health risks by long term use (oral NSAIDs), or risks of addiction (opioids), or efficacious but with limited duration of action (IA corticosteroids). The topical application of NSAIDs, including diclofenac sodium, is considered a viable and effective alternative treatment option.

The potential benefits of study participation are that participants with knee OA may experience a reduction in symptoms. No other benefits of participation are anticipated.

The study population is required to have a certain degree of pain due to their underlying condition. As the trial is placebo-controlled, approximately half of the population will receive a pharmacologically inert treatment and may not derive any symptomatic relief of the investigational treatment. In order to address this situation, rescue medication in the form of paracetamol is allowed to be used throughout the trial.

The potential risks of study participation comprise those associated with AMZ001. Based on the observations of the pre-clinical and clinical studies, and studies of topical diclofenac products, the most common adverse reactions were mild to moderate, and mostly concerning local skin reactions.

Due to the alcohol content of the product (45% ethanol and 3% MA), frequent applications to the skin may cause irritation and dry skin. Adverse drug reactions reported in 7 –12% of patients treated with similar types of diclofenac-containing gels were mainly application site reactions including, but not limited to, dermatitis, pruritus, erythema, paresthesia, dryness, vesicles, irritation, and papules.

Participants participating in the study will be also exposed to the discomfort associated with the study procedures, such as collection of blood samples and knee X-ray for screening. For blood samples the risk of pain or local bruising and infection at the site where blood is drawn for laboratory tests. There is also a small risk of a fainting episode, which can occur as a reaction to giving blood. As for the x-ray, this study involves radiation exposure from 1 radiograph of the knees at screening. As part of everyday living, everyone is exposed to a small amount of background radiation that comes from soil, rocks, outer space, and within the body itself. Each radiograph in this study exposes subjects to approximately 0.005 millisievert (mSv), equivalent to about one half day of background radiation in the US. (Thurston, 2009) This radiation exposure may not be necessary for the subjects' medical care, but it is necessary to obtain the research information desired.

Risks to subjects will be minimized by clinical safety oversight performed by centralized review and on-site monitoring.

Taking the above information into account, an assessment of risks and benefits supports the current study designed to investigate AMZ001 as a potential therapy for patients with knee OA.

4 Trial Objectives

4.1 Primary Objective

The primary objective of the study is:

- To compare the efficacy of diclofenac gel AMZ001 versus placebo for the treatment of pain, in terms of change in pain intensity evaluated by the WOMAC pain sub-score of the target knee, at week 2.

4.2 Secondary Objectives

The secondary objectives of the study are:

- To evaluate changes in pain, in terms of the WOMAC pain sub-score of the target knee, at weeks 1, 3, 4, and 6
- To evaluate changes in daily pain, in terms of the 11-point Pain Numeric Rating Scale (NRS) of the target knee until week 6.
- To evaluate the onset of action of AMZ001, based on self-rated average pain scores 1 hour, 4 hours, 12 hours, 24 hours after the first drug application.
- To evaluate changes as measured by Physician Global Assessment (PGA) at week 3 and 6
- To evaluate changes in sleep, as measured by the Chronic Pain Sleep Inventory (CPSI), at weeks 3 and 6
- To evaluate changes in physical functioning, in terms of the WOMAC function sub-score of the target knee, at weeks 1, 2, 3, 4 and 6
- To evaluate the symptoms of osteoarthritis in terms of the WOMAC total score of the target knee, at weeks 3 and 6
- To evaluate changes in quality of life, as measured by the EQ-5D-5L, at weeks 3 and 6
- To evaluate changes as measured by the OMERACT-OARSI responder criteria, at weeks 3 and 6
- To evaluate the use of rescue medication
- To evaluate the safety and tolerability of AMZ001

4.3 Estimands

The estimand framework presented below provides the systematic approach to the definition of the treatment effect under investigation in the trial. Estimands consists of 5 attributes: Analysis population, endpoint attribute, treatment condition, intercurrent events, and population-level summary.

The following intercurrent events (ICEs), which can occur after the initiation of IMP and affect either the existence or the interpretation of the measurements associated with the clinical question of interest, is addressed by the defined estimands:

- Treatment discontinuation (for treatment related reasons, e.g. related adverse event (AE), lack of efficacy (LOE)).
- Use of short-term analgesic, protocol defined rescue medication in the form of paracetamol/acetaminophen, including during the 24-hour window prior to a study visit.
- Excessive use of rescue medication.
- Use of prohibited medication/non-permitted therapies with low expected impact.
- Use of prohibited medication/non-permitted therapies with high expected impact.

Two estimands will be applied to the primary and secondary efficacy endpoints, one for continuous endpoints, the other for binary. They each aim to answer the following clinical question of interest:

Clinical question of interest:

What is the estimated treatment effect of Diclofenac Gel versus Placebo

- if participants would continue to be treated for the entire study duration, unless
- they discontinue IMP due to treatment related reasons,
- without the benefit of additional treatment, except for low impact medication.

The aim is to estimate the treatment effect of Diclofenac Gel vs Placebo attributable to the randomized treatment (“treatment attributable” estimand) alone or in combination with trial rescue medication or other analgesic therapies, assuming these don’t have strong impact on efficacy.

Additionally, two secondary estimands will be applied to the primary and secondary efficacy endpoints, one for continuous endpoints, the other for binary. They each aim to answer the following clinical question of interest:

What is the estimated treatment effect of Diclofenac Gel versus Placebo

- if participants would continue to be treated for the entire study duration,

- had they not initiated any additional treatments, except for low impact medication.

The aim is to estimate the treatment effect of Diclofenac Gel vs Placebo if randomized treatment was taken as instructed ("efficacy" estimand) alone or in combination with trial rescue medication or other analgesic therapies, assuming these don't have strong impact on efficacy.

4.3.1 Primary estimand for continuous endpoints

The estimand for continuous endpoints is described by the following attributes:

- Analysis population: All randomized participants (Intention-to-Treat (ITT) set).
- Variable: Change from baseline or AUC of change from baseline in the corresponding efficacy score.
- Treatment condition: Randomized treatment, i.e., AMZ001 or Placebo, with or without use of dispensed rescue medication (paracetamol/acetaminophen) or other analgesic therapies not deemed to have a large impact on trial outcome.
- Handling of intercurrent events and missing data post ICE:
    - Treatment discontinuation: Discard data after ICE (hypothetical) and impute under MNAR (return-to-baseline) as this ICE is viewed as treatment failure.
    - Use of short-term analgesic, protocol defined rescue medication in the form of paracetamol/acetaminophen, including during the 24-hour window prior to a study visit: Regardless of use (treatment policy).
    - Excessive use of rescue medication: Will be considered as an instance of prohibited medication use and classified accordingly as either low or high expected impact.
    - Use of prohibited medication/non-permitted therapies with low expected impact: Regardless of use (treatment policy).
    - Use of prohibited medication/non-permitted therapies with high expected impact: Discard data affected by ICE (hypothetical) and impute under MNAR (return-to-baseline) as this ICE is viewed as treatment failure.
- Population-level summary: Mean difference between treatment groups.

Rationale for the estimand:

Treatment discontinuation for treatment related reasons indicates treatment failure, hence imputations are done based on MNAR, imputing to baseline, which essentially assumes that after such ICE, subject's condition returns to the levels observed at baseline.

For rescue medication use and use of prohibited medications/non-permitted therapies with low expected impact the treatment policy approach is chosen as these are anticipated to have low impact on efficacy evaluation. On the other hand, use of prohibited medication/non-permitted therapies with high expected impact is viewed as treatment failure, hence the same strategy as for treatment discontinuation possibly related to treatment is used.

Sensitivity analyses that examine the impact of alternative handling of ICEs may be employed and these will be described in SAP, e.g., treatment policy strategy for prohibited meds with high anticipated impact, treatment policy strategy for treatment discontinuation related to treatment, etc.

4.3.2 Primary estimand for binary endpoints

- The estimand for binary endpoints is described by the following attributes Analysis population: All randomized participants (ITT-set).

- Variable: Binary yes/no variable, describing whether a participant achieved response or not, where discontinuation of treatment or use of prohibited medication/non-permitted therapies with high expected impact is viewed as another mode of failure.

- Treatment condition: Randomized treatment, i.e., Diclofenac or Placebo, with or without use of dispensed rescue medication (paracetamol/acetaminophen) or other analgesic therapies not deemed to have a large impact on trial outcome.

- Handling of intercurrent events:

    - Use of short-term analgesic, protocol defined rescue medication in the form of paracetamol/acetaminophen, including during the 24-hour window prior to a study visit: Regardless of use (treatment policy).

    - Excessive use of rescue medication: Will be considered as an instance of prohibited medication use and classified accordingly as either low or high expected impact.

    - Use of prohibited medication/non-permitted therapies with low expected impact: Regardless of use (treatment policy).

- Population-level summary: Odds ratio between treatment groups.

Rationale for the estimand:

Similar to the estimand for continuous endpoints, discontinuation of IMP for treatment related reasons or use of prohibited medication/non-permitted therapies with high impact indicate failure of randomized treatment, hence are considered non-response for the binary outcome. The rest of rationale follows the same logic as for the estimand for continuous endpoints.

4.3.3 Secondary estimands

The secondary estimands for continuous and binary endpoints will treat all data post treatment discontinuation/use of prohibited medication with high impact as missing (hypothetical) and

data post such intercurrent events will be imputed assuming MAR. This also means that for the binary estimand, treatment discontinuations or use of prohibited medication/non-permitted therapies with high expected impact are interpreted as ICEs, instead of being incorporated in the variable definition through the composite strategy where such events are understood as non-response.

5 Investigational Plan

5.1 Overall Trial Design and Plan

The trial is a multi-center, double-blind, randomized, placebo-controlled trial of AMZ001 for the treatment of pain and symptoms of knee OA. The purpose of the trial is to evaluate the efficacy and safety and tolerability of once daily application of [redacted] AMZ001 during a 42-day period in participants with radiographic and symptomatic knee OA in either one or both knees. A total of approximately 540 participants are planned to be randomized. The participants will be randomized in a ratio of 1:1 to one of the two treatment groups, with approximately 270 participants in each treatment group.

The trial consists of two phases: screening and treatment, as illustrated in Figure 2

The planned total duration of the trial for each randomized participant will be up to 84 days:

- Screening period, up to 28 days prior to randomization
- Treatment period of 6 weeks (42 days) following randomization
- Follow-up period of 2 weeks (14 days) following the last dose of the IMP

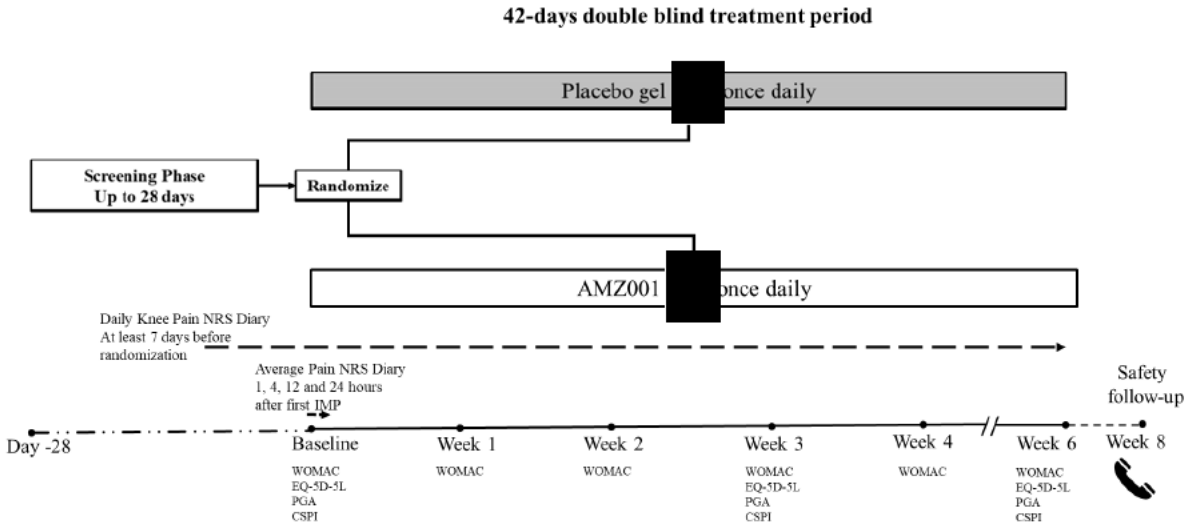


Figure 2: Trial Outline

Screening Period

The screening period will comprise a period of up to 28 days during which a participant's eligibility for the trial will be determined, beginning with the signature of the informed consent form (ICF) (Visit 1a). Participants who have taken analgesic medication at Visit 1a will be scheduled to return for complete screening procedures (Visit 1b) after a washout of all analgesic medications. Participants not taking analgesic medication at the time of ICF signature may complete Visit 1a and 1b screening procedures at one visit. Screening will include demographics, collection a typical knee pain when not taking analgesics and completion of the WOMAC pain sub-score (5 questions) for both knees and hips and record of the daily pain on the knees on an 11-point pain NRS during a period of at least 7 days. Participants who have a WOMAC pain sub-score in at least one knee of 20 to 50 after a predefined analgesic washout period will continue with other screening procedures. During the screening, medical history and prior and concomitant medication use data will be collected, the participants will undergo a physical examination, including 12-lead ECG, vital signs, measurement of sitting diastolic and systolic blood pressure, heart rate, and measurement of height and weight. Blood will be collected for safety laboratory analyses (hematology and biochemistry). Pregnancy test at both screening visit (serum) and baseline prior randomization (urine) pregnancy test will be performed for women of child-bearing potential. X-ray images of the knees will be obtained using a Fixed-Flexion frame (Synaflexor© or similar) and a central radiologist reader will grade OA on the Kellgren-Lawrence scale. If participants fulfill all inclusion and no exclusion criteria (see Sections 5.3.1 and 5.3.2), they will be randomized into the trial.

Double-blind Treatment Period

During the treatment period, all participants will apply the IMP onto the anterior, medial and lateral areas (not the posterior) of the painful knee(s) as described in Section 6.2.2, for 42 days starting on the day of randomization. The treatment phase begins at the baseline visit (Day 1, Visit 2), when eligibility is confirmed, the participant is randomized, and the first IMP will be applied. The first application will be supervised at the clinical center, and subsequent applications will be self-administered, except for the visit days where the application will also be supervised at the clinical center. Participants will apply IMP at approximately the same time in the morning once daily (every 24 ±1 hours). In addition to the baseline visit, the participants will attend five study visits during the double-blind treatment period, at Weeks 1, 2, 3, 4, and 6, at which efficacy assessments will be made.

At baseline the target knee will be selected according to the radiographic and symptomatic criteria, more on the selection of the target knee in section 5.4.2. If both knees fulfill the criteria, both knees should be treated, but the knee with the highest WOMAC pain score will be selected as the target knee. The target knee must be unchanged throughout the trial regardless of changes in symptoms of the target knee during the trial. Daily NRS-pain and WOMAC will only be recorded for the target knee. If, at any time during the treatment period of the trial, the participant refers increase of pain in the contralateral knee and this is assessed by the investigator as significant, IMP treatment should be initiated for that knee also. If IMP treatment of the contralateral knee is initiated at any time during the trial, the treatment must continue throughout the treatment period, regardless of whether analgesia is achieved, pain for the contralateral knee will not be recorded. Treatment of the contralateral knee must only be initiated at study visits, and not during the course of the trial on the participant's own initiative.

AEs and concomitant medications and procedures will be monitored throughout the trial, and vital signs will be measured at each trial visit. Safety assessments will also include 12-lead

electrocardiograms (ECGs), and safety laboratory analyses (hematology and biochemistry) collected at different timepoints throughout the trial. Height will be measured at screening visit only. Data regarding any surgical procedures of the treated knees will also be collected.

A detailed schedule of study procedures/assessments is provided in Table 1.

Follow-up Period

Two weeks after the last site visit and the last dose of investigational product has been administered, site staff will perform a phone call to collect information concerning new or changed (Serious) Adverse Events (AE), and changes in concomitant medication.

Trial Endpoints

Trial endpoints and their analyses are presented in Section 8.3.

Interim Analysis

No interim analysis is planned.

5.2 Discussion of Trial Design

Overall Trial Design

The double-blind, randomized, parallel group, placebo-controlled design of the trial is considered the reference standard study design for evaluating the efficacy of medical interventions in most disease areas.

The trial consists of a screening period of up to 28 days, a 6-week treatment period and a 14 day follow-up period. The screening period of up to 28 days is considered appropriate for assessment of eligibility, including acquisition and central, independent evaluation of knee X-rays for radiographic eligibility criteria. The treatment phase of 6 weeks (42 days) is considered sufficient to evaluate the potentially clinically meaningful efficacy, safety and tolerability of AMZ001. The follow-up period of 14 days is considered sufficient to capture any delayed adverse events that might occur after the IMP has been stopped and considered washed out, and to evaluate the development after treatment cessation of AEs which were ongoing at the time of treatment completion.

Study Population

The proposed study population is highly similar to that of other trials evaluating topical NSAIDs as analgesic treatment of symptomatic knee osteoarthritis (Bookman, Williams and Shainhouse, 2004; Baraf *et al.*, 2010; Wadsworth, Kent and Holt, 2016), and to that of the previous proof-of-concept trial with AMZ001-006. Participants with theoretical additional health risk factors including active skin disease at the planned site of application, known gastric ulcer, malignancies, poor renal function among other risk populations, are excluded from participation. Furthermore, interference from other conditions likely to confound the efficacy outputs of the trial is sought minimized by excluding participants with concomitant non-OA inflammatory diseases affecting either knee, secondary OA or previous trauma or surgical intervention of the target knee, as well as OA of the hips with significant pain, significant

radicular back pain or active migraine, among other or if pain of the contralateral knee exceeds the pain of the target knee.

Diclofenac may induce a teratogenic/fetotoxic potential risk from the second trimester, and is contraindicated during the third trimester of pregnancy (FDA 2009). Since AMZ001 is a novel formulation of diclofenac with known excipients, it would be considered unlikely to have human teratogenicity/fetotoxicity in early pregnancy. As a precautionary measure, women of childbearing potential will be allowed to participate in the trial, provided they adhere to the use a highly effective method of contraception, as described in Appendix B, throughout the trial and for 3 months after the last study treatment application. Although diclofenac is not considered to have mutagenic potential, as a precautionary measure, sexually active men with a female partner of childbearing potential must agree to use condom from enrolment up to at least 3 months after the study end.

Dose and Dose Regimen

The trial aims to evaluate whether topical AMZ001 applied in small quantities (once daily) can demonstrate a significant relief of pain and symptoms of OA as compared with placebo.

Based on the nonclinical data and clinical data from studies with healthy volunteers and OA patients, once daily application of AMZ001 is considered likely to reach therapeutic periarticular tissue levels of diclofenac similar or superior to those observed with four times daily application of commercial diclofenac sodium gel formulations with a reduced total administered dose. The same dose will be applied to all participants in the same anatomical region. However, due to variations between participants, the total surface area for application will differ among these. Despite the difference, these variations are not expected to impact the efficacy of the product or its local effect in the knee, as the total dose of diclofenac permeating the skin and reaching the target tissue is considered likely to be independent of the area on which the gel is applied.

The duration of IMP treatment of 42 days is considered appropriate for evaluation of efficacy of AMZ001. The safety of topical diclofenac administration is considered well established, and no new safety signals are expected with AMZ001. Regardless, the safety and tolerability of a local exposure to diclofenac using the novel formulation of AMZ001 is less well described and will be assessed in the current trial. Specifically, local tolerance will be assessed by a procedure evaluating the skin of the participants, as described in Section7.6.3.

Placebo Gel as Comparator

The placebo gel is identical to that of AMZ001 but without the active constituent diclofenac sodium. The placebo gel will be indistinguishable from AMZ001 used in the other double-blinded treatment group by both participants and the study team, which allows direct comparisons between the efficacy data of AMZ001 and placebo upon study completion.

As subjects participating in the trial have clinically relevant pain in their target joint, particularly those randomized to receive placebo may need additional analgesic treatment during the trial. The use of rescue medication is a well-known method of allowing access to a standardized analgesic which also allows detailed quantification of used medication. The limited duration of the trial together, and the low risks associated with using placebo gel as a

comparator, supports that the risk does not outweigh the benefits of allowing a firm reference for comparison for the active study groups, and the use of the placebo gel is thus considered appropriate.

Use of rescue medication

As participants participating in the trial have clinically relevant pain in their target joint, participants may need additional analgesic treatment during the trial. The current trial allows for use of rescue medication in such cases. The use of rescue medication is a well-known method of allowing access to a standardized analgesic.

Trial Efficacy Assessments

Evaluation of WOMAC pain sub-score as the primary endpoint

It is well recognized that OA pain is multifactorial, consisting of elements of psychosocial and biologic factors such as local inflammation leading to lowered nociceptor firing thresholds (Dieppe and Lohmander 2005). The standard method of measuring pain in clinical trials is using validated patient-reported outcomes (PRO). In OA research, the WOMAC psychometric tool (Bellamy *et al.* 1988), invented in 1981, is considered one of the most well-described, validated and recognized PROs to assess the impact of OA on pain, joint stiffness and physical function (FDA 1999). The pain sub-score of the WOMAC index consists of five questions evaluating both pain on movement and pain while idle and has been used as the primary endpoint of clinical trials evaluating the efficacy of topical diclofenac sodium products (Barthel *et al.* 2009, Baraf *et al.* 2010). The total WOMAC score, including measures of physical function, pain, and joint stiffness, will be assessed as a secondary endpoint. The WOMAC Index is further described in Section 7.4.1.

The primary objective of the study is to evaluate the efficacy of AMZ001 compared to placebo for the treatment of pain evaluated by WOMAC pain sub-score at week 2. AMZ001 has been developed for OTC use, similar to other diclofenac topical products it aims to alleviate knee pain as a short treatment period, with a proposed treatment duration of 7-21 days. For this reason, the primary endpoint will be set at week 2.

Other evaluations of pain, physical function and quality of life

In addition to the WOMAC, other questionnaires will be used to explore the effects of AMZ001 treatment on the symptoms associated with OA.

- The OMERACT-OARSI Responder evaluation (Pham *et al.* 2004) (Section 7.4.2) will be assessed as a secondary endpoint in the current trial.

- The daily pain will be collected using an 11-point NRS with a 24-hour recall period. The time of onset of pain relief will be assessed based on recorded pain using an 11-point NRS (Section 7.4.3)

- Sleep quality will be assessed using the CPSI (Ayearst *et al.* 2012). (Section 7.5.2).

- The number of days using, and total use of rescue medication will be evaluated as an endpoint, as the use of rescue medication is expected to be associated with occurrence of pain and thus lack of efficacy of the IMP during the trial.

- The quality of life of the participants will be assessed using the EQ-5D-5L (Drummond *et al.* 1998) (Section 7.5.1).

5.2.1 Inclusion of Special Populations

The trial population will consist of adults aged between 40 and 85 years. The inclusion of elderly participants is considered justified because this is the group most frequently affected by OA, in addition to the IMP being topically applied, by which systemic exposure is expected to be limited.

5.3 Selection of Trial Population

The trial will randomize approximately 540 participants, in an appropriate number of centers located in the multiple countries.

Only persons meeting all inclusion criteria, and no exclusion criteria may be enrolled in the trial as participants. Prior to performing any trial or trial-related assessments not part of the participant's routine medical care, the Investigator will ensure that the participant has provided written informed consent.

5.3.1 Inclusion Criteria

For inclusion in the trial, all the following criteria must be fulfilled:

1. Participant is able to read and understand the language and content of the study material, understand the requirements for study visits, and is willing to provide information at the scheduled evaluations and appropriate written informed consent has been obtained.

2. Femorotibial OA of the knee, according to the American College of Rheumatology (ACR) clinical and radiographic criteria (Altman *et al.* 1986).

3. Radiological OA grade 2 or 3 of the target knee, using the Kellgren-Lawrence method (Kellgren and Lawrence 1957) as graded by central, independent reading of X-ray obtained during screening, or on a recent (within 6 months) X-ray image which fulfills the specifications for central reading.

4. Age between 40 years and 85 years at the time of screening, both included of either sex.

5. Pain score rated on an 11-point numerical rating scale of the target knee of ≥ 20 out of 50 in response to the WOMAC pain sub-score (5 questions), at the time of screening and baseline.
6. The WOMAC pain sub-score on the contralateral knee must not exceed the one on the target knee, regardless of the eligibility of the contralateral knee, at the time of screening and baseline.

7. At screening Visit 1a, participants report that their typical OA knee pain in one or both knees when not using medication is ≥ 4 out of 10.
8. Daily OA knee pain diary average numerical rating scale (NRS) score of ≥ 4 and ≤ 9 in the target knee, for the 7 days immediately preceding baseline (Day 1). The average calculation is based on the recorded scores during this entire period with a requirement of at least 4 days of data recorded.
9. Women of child-bearing potential must use a highly effective method of contraception (hormonal contraceptives, intrauterine device, vasectomy, bilateral tubal occlusion, sexual abstinence) from enrolment up to at least 3 months after the study end. Postmenopausal status is defined as being amenorrheic for at least 1 year prior to screening. Sexually active men with a female partner of childbearing potential must agree to use condom from enrolment up to at least 3 months after the study end.
10. Knee pain in the target knee for 14 days of the preceding month (periarticular knee pain due to OA and not due to non-OA conditions such as bursitis, tendonitis, etc.) based on participant report.
11. Except for OA, the participant is in reasonably good health as determined by the Investigator.

5.3.2 Exclusion Criteria

Participants are not eligible for the trial if they fulfill any of the following exclusion criteria:

1. Known or suspected hypersensitivity to or previous hypersensitivity reactions to diclofenac or of the excipients in either of the investigational products
2. Patients in whom asthma, angioedema, urticaria, or acute rhinitis are precipitated by acetylsalicylic acid or other non-steroidal anti-inflammatory drugs (NSAIDs).
3. Any physical impediment to gel application on the target knee.
4. Intra-articular delivery of corticosteroids or hyaluronic acid in the target knee within 6 months of screening or into any other joint within 30 days of screening.
5. High dose (equivalent to > 10 mg of prednisone/day) systemic corticosteroid treatment of more than 14 days during the past 6 months prior to screening.
6. Major surgery /arthroscopy of the target knee within the previous year prior to screening or aspiration of effusion of the target knee within 12 weeks prior to enrollment.
7. Planned surgery of the target knee within the next 12 months.
8. Use of a currently unapproved investigational drug, device or biologic within 3 months prior to screening.
9. Presence of concomitant non-osteoarthritic disease affecting either knee, such as rheumatoid arthritis, psoriasis, gout or clinically relevant pseudogout, if there is reason

to believe that the disease(s) may significantly interfere with the interpretation of the clinical response to the study drug.

10. Perioperative period in the setting of coronary artery bypass graft surgery.

11. Current malignancy or treatment for malignancy within the past five years, with the exception of non-melanoma skin cancer, unless affecting the target knee area, or carcinoma in situ events.

12. Any other abnormal laboratory results or significant medical conditions that the Investigator believes should preclude the participant's participation in the trial.

13. Secondary OA of the target knee, previous procedures or trauma affecting joint homeostasis including total meniscectomy or septic arthritis, or any other serious condition leading to secondary OA of the target knee.

14. Reported incidence of any of the following diseases: known osteoarthritis of the hip(s) if pain in either or both hip(s) exceeds that of the target knee using the WOMAC hip pain sub-score, or presence of significant radicular back pain.

15. Presence of a defined cut-off of pain variability of 1.5 out of 10:

    · Pain NRS scores must be recorded for the target knee on at least 4 out of the 7 days immediately preceding baseline (Day 1)

    · Observed standard deviation of the Daily OA knee pain diary average NRS intensity score for 7 days immediately preceding baseline (Day 1) must not exceed 1.5

16. Estimated glomerular filtration rate (eGFR) < 30 mL/min using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI 2021) formula.

17. Generalized skin irritation, previous skin reactions upon use of topical NSAIDs, current skin irritation or redness at the planned site of gel application, or significant skin disease including psoriasis, as judged by the investigator.

18. Use of any topical medication on the planned application site within 15 days of the time of randomization.

19. Any use of opioid medication within 6 weeks before the screening visit.

20. Use of duloxetine, pregabalin, or gabapentin within 4 weeks before the screening visit.

21. History of alcohol or drug abuse within the past year prior to randomization.

22. Pregnant and breastfeeding women are excluded for participation.

5.4 Criteria for Initiation of Trial Treatment

5.4.1 Pain assessment

The participant's pain level will be assessed during the screening period, and assessment of eligibility will be based on the baseline visit pain report.

As an essential part of the eligibility determination process, participants will be asked to complete the pain sub-score of the WOMAC index for both knees and hips (5 questions). The WOMAC questions which determine eligibility must be obtained after the patient has provided informed consent and, if taking pain medications, has withheld them, as described below. Participants who are taking pain medications and who are potentially eligible for the trial will be asked to withhold all pain medications including paracetamol (acetaminophen), NSAIDs including inhibitors of cyclooxygenase (COXIBs) for at least 5 half-lives of the given medication before the screening (Visit 1b) assessments. Aspirin (acetylsalicylic acid) used for cardiovascular protection in doses up to 100 mg/day does not need to be withheld. Participants must have a total score of $\geq$20 points for the target knee WOMAC pain index at the time of screening (after washout of analgesic medications) and baseline. If the WOMAC hip score for either hip at the screening visit exceeds that of the proposed target knee at the same visit, the participant is not eligible for randomization. If the hip pain score is identical to or lower than that of the target knee, the participant may be randomized if otherwise eligible

5.4.2 Selection of target knee

The target knee must meet all specified inclusion criteria relating to the knee at the time of screening: Radiographic (ACR OA criteria, Kellgren-Lawrence grade), symptomatic (WOMAC pain sub-score at screening and baseline, significant knee pain history), and must not meet any of the exclusion criteria, i.e., prohibited medical treatment, surgery, presence of other significant disease affecting the knee or presence of secondary OA.

If both knees meet the specified criteria at the baseline visit, the knee with the highest WOMAC pain score will be selected as the target knee.

If both knees meet the specified pain-criteria, but only one of the two meets the radiographic criterion, the fully eligible knee is selected as the target knee.

If both knees qualify, but have the same WOMAC pain score, the knee with the highest Kellgren-Lawrence grade should be selected as the target knee.

If both knees qualify, have the same WOMAC pain score and the same Kellgren-Lawrence grade, the Investigator may select either knee as the target knee in consultation with the Subject.

The target knee will remain the same throughout the trial, regardless of changes to parameters influencing the selection of target knee during the trial.

Both knees will be assessed separately on the WOMAC pain sub-score for the screening and baseline visits.

5.5 Criteria for Participant Withdrawal

5.5.1 Withdrawal from Trial Therapy

Participants are free to discontinue the trial at any time without giving their reasons. The Investigator may withdraw a participant at any time if this is considered to be in the participant's best interest.

A participant must be withdrawn in the event of any of the following:

- Withdrawal of the participant's consent
- Participant's participation in another trial

If a participant is discontinued at any time after entering the study, the investigator should make a reasonable effort to ascertain the reason(s) while fully respecting the subject's rights, as a subject is not obliged to give (a) reason(s) for withdrawing. The investigator or site personnel should make every effort to contact the subject to assess his or her status and complete the termination page in the eCRF.

If a participant fails to attend scheduled trial assessments, the Investigator must determine the reasons and the circumstances as completely and accurately as possible. In case of premature withdrawal from the trial, the Early termination visit should be performed, with the procedures scheduled for Week 6/Visit 7 at the time of withdrawal whenever possible, with focus on the most relevant assessments. Participants who withdraw from the trial will not be replaced.

Treatment with the IMP must be discontinued in the event of any of the following:

- Occurrence of an exclusion criterion which is clinically relevant and affects the participant's safety, if discontinuation is considered necessary by the Investigator and/or Sponsor.
- Occurrence of AEs, if discontinuation of trial drug is desired or considered necessary by the Investigator and/or the participant.
- Occurrence of pregnancy (see Section 7.6.2 for reporting of pregnancies and follow-up of participants who become pregnant during the trial).
- Substantial non-compliance or protocol violation.

In case of early IMP discontinuation, the reason for discontinuation (if it is known) will be registered in the eCRF and source documentation. Apart from procedures directly related to the IMP treatment, all procedures as planned at visits will remain unchanged for those participants who discontinue study treatment and choose to remain in the study, including daily use of rescue medication, concomitant medication and pain NRS. Safety and efficacy assessments will take place as described in Schedule of Trial Procedures (see Section 7.1 and Table 1). For participants continuing in the trial without IMP treatment, the investigator or site personnel should systematically and consistently attempt to contact participants who fail to actively maintain contact with the site.

5.6 Premature Termination of the Trial

The clinical trial may be terminated prematurely or suspended at the request of Health Authorities, or if new safety or efficacy information leads to an unfavorable risk benefit judgment for any IMP. The Sponsor may discontinue the trial if it becomes unjustifiable for medical or ethical reasons, for poor enrollment, or because of discontinuation of clinical development of an IMP or withdrawal of an IMP or comparator from the market for safety reasons.

Health Authorities and Independent Ethics Committees (IECs)/Institutional Review Boards (IRBs) will be informed about the discontinuation of the trial in accordance with applicable regulations.

5.7 Definition of End of Trial

The end of the trial is defined as the last participant's last visit (LPLV). The last visit of the trial is Visit 8.

A clinical trial protocol may not be considered closed as long as:

- Any participant is still receiving any IMP,
- Visits specified by the protocol are still taking place,
- Procedures or interventions according to the protocol are still being undertaken in any participant.

6 Investigational Medicinal Product and Other Drugs Used in the Trial

The term "Investigational Medicinal Product" (IMP) refers to AMZ001 and placebo gel being tested in the clinical trial. Rescue medication refers to paracetamol (acetaminophen) 500 mg tablets as described in Section 6.4.

6.1 Description of the Investigational Medicinal Product

Clinical trial materials are manufactured in accordance with Good Manufacturing Practice (GMP). All clinical trial materials not currently approved and marketed will be supplied with Certificates of Analysis and/or Release Statements. Trial drugs will be identified by trial code, kit number, and Sponsor name.

IMP packaging and labeling are described in Section 6.6; IMP preparation, handling, and storage are described in Section 6.7.

6.1.1 AMZ001

Diclofenac sodium European Pharmacopoeia (EP)/US Pharmacopoeia (USP) is a white to slightly yellowish crystalline powder that is slightly hygroscopic. The diclofenac sodium drug substance is manufactured by Amoli, Mumbai, India.

AMZ001 is a completely homogeneous, transparent and non-staining hydroalcoholic gel, containing [redacted] diclofenac sodium. [redacted] The excipients are well known and have been used in previously approved topical products. AMZ001 Diclofenac gel is manufactured by Orion (Turku, Finland).

Nomenclature

Common name: Diclofenac sodium

Chemical name: 2-[2-(2,6-dichloroanilino) phenyl] acetic acid sodium; or benzeneacetic acid, 2-[(2,6-dichlorophenyl)amino]-, monosodium salt

CAS No.: 15307-79-6

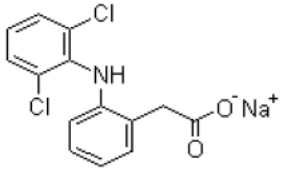
Amzell code: AMZ001

Trade name: Not applicable

Molecular formula: $C_{14}H_{10}Cl_2NNaO_2$

Molecular weight: 318.13

Structural formula:



Stability has been documented for AMZ001 in Lablabo metering dispensers (Lablabo.com) supporting a shelf life of 36 months when stored not above 25°C. The product is supplied in an 87 g metered dose dispenser.

6.1.2 Placebo

Placebo will be supplied as a gel visually identical to the active product and will be supplied in packaging identical to that for AMZ001. The excipients will be identical to those in AMZ001.

6.2 Dosage and Administration

During the trial, all participants will receive 6 weeks (42 days) of treatment with IMP for application [redacted] once daily, per knee.

6.2.1 Application Schedule

AMZ001 or placebo [redacted] once daily – double-blind treatment phase

Each randomized participant will be randomized to once daily double-blind treatment regimens, as shown in Figure 2. The participants will be dispensed one metered dose-dispenser for each knee to be treated at each dispensation visit. Participants should be instructed to apply 2 pump actuations of the IMP [redacted], per knee, at approximately the same time each morning (every 24 ±1 hours), i.e., once daily. Participants will be instructed to not apply the IMP at home on visit days and wait for the visit where the application will be supervised by the site staff after the skin assessments have been done.

For the purpose of facilitating standardization in study procedures and relevant sampling in the study, participants are required to apply the IMP gel in the morning.

In case a planned application is missed, the participant should be instructed to wait until the next planned time of application before applying the next dose as planned. The participant should be instructed to contact the study staff in case of doubts regarding missed doses.

6.2.2 Method of Application

AMZ001 and Placebo

The gel should be applied onto the knee and be allowed to dry. The skin should be clean and dry, free of soap, perfume or disinfectant. The gel should be spread and rubbed in a gentle manner onto the anterior, medial and lateral areas, (not the posterior) of the painful knee(s) and the application site should not be covered with clothes for at least ten minutes after application of the gel. Application of gel to the posterior area of the knee must not occur in the central skin fold.

Showering or bathing should be avoided for at least 1 hour before applying and at least 4 hours after applying the gel. External heat and/or occlusive dressings should not be applied to the treated knee(s). Exposure of the treated knee(s) to sunlight should be avoided.

Participants should be instructed to wipe their hands with a paper towel following application of the gel and wash them with soap and water.

Further details regarding the application procedure are provided in the Instructions for Use pamphlet.

6.2.3 Temporary Suspension of Study Treatment

In cases of intolerable adverse events, the study treatment may be temporarily suspended for a period of up to 14 consecutive days, for evaluation of resolution or clinical improvement of the event(s) in question. If the events resolve or improve to an acceptable level, the planned treatment should be re-started and continued for the remainder of the trial period. If the reason for suspension recurs with a comparable or aggravated severity after restarting study treatment, the Investigator may decide to discontinue the study treatment permanently. The timing and duration of any suspension of study treatment should be reported to the study staff at each study visit and recorded in the eCRF.

6.3 Assignment to Treatment Groups

Eligible participants will be randomized to one of two treatment groups with an allocation ratio of 1:1 between the double-blinded treatment groups (270 participants in each group). Participants will be required to apply [redacted] the IMP once daily, per treated knee.

The screening/randomization procedure will be centrally managed through an electronic Interactive Web Response System (IWRS) integrated in the eCRF. To enroll and randomize a new participant, the Investigator/authorized personnel will have to access the eCRF by entering their username and password and filling in the requested data.

During the screening visit, the participant will be given a unique Subject ID number. The participant will be allocated to a treatment group through the IWRS after the Investigator or delegated person confirms eligibility.

6.4 Auxiliary medicinal products to be Used

The trial allows paracetamol/acetaminophen tablets 500 mg to be used as analgesic rescue medication in case of intolerable pain during the double-blind treatment phase of the trial. The dosage of paracetamol that the participants will be allowed to take per day will be defined according to the standard of care in the countries where the trial will be carried out; however, the maximum dose should not exceed 1 g per dose and 4 g per day.

Paracetamol/acetaminophen is an authorized product and will be used and stored in the clinical trial in accordance with the terms of its marketing authorization and approved label as per the countries the trial will take place. The supply of paracetamol will be coordinated centrally, using locally sourced auxiliary medicinal products for USA and HK, for European countries where it will be sourced centrally. Each supplied rescue medication kit will contain a maximum of 100 tablets containing each a maximum of 500 mg of paracetamol/acetaminophen.

The total use of rescue medication will be accounted for by a pill count of remaining tablets upon participant completion of the trial, interim accountability is performed at each trial visit. The days of rescue medication use (if any) will be recorded as reported by the participant in the ePRO device after the first visit. Subjects will be asked at each site visit if the maximum dose was exceeded during any of the days since the last visit.

Administration of rescue medication is not allowed within 24 hours prior to a scheduled study visit.

6.5 Prior and concomitant Medications and Therapies

All concomitant medications taken by the participant during the trial from the date of signature of informed consent are to be recorded in the appropriate section of the eCRF, noting the name, dose, including route and frequency, start and end dates and indication of each drug, except for rescue medication dispensed in accordance with the protocol, which will be accounted for separately.

Nonpharmacological interventions and any changes to a concomitant medication or other intervention should also be recorded in the eCRF.

6.5.1 Permitted Medicines

Any medications that are considered necessary to protect participant welfare and that will not interfere with the trial medication may be given at the Investigator's discretion.

The following medication is considered permitted:

- Non-analgesic therapy for non-arthritic conditions
- Stable low-dose aspirin (up to 100 mg/day) for cardiovascular prophylaxis, if used regularly for at least 30 days prior to screening
- Stable therapy with sedative hypnotics (e.g., flurazepam, zolpidem), if used regularly for at least 14 days prior to screening
- Stable therapy with anti-depressants for treatment of depression, if used regularly for at least 14 days prior to screening, except for duloxetine and other SNRI
- The use of systemically delivered, low-dose corticosteroids (equivalent to ≤ 5 mg prednisolone/day) is allowed if used for shorter periods (< 14 days) during the trial
- Inhaled steroids for the treatment or prevention of asthma, seasonal allergy or other respiratory conditions
- Stable therapy for management of insomnia, if used regularly for at least 14 days prior to screening

The use of a walking cane, if used continuously for at least 14 days prior to screening, is allowed.

Oral supplements such as glucosamine, diacerein, chondroitin sulfate may be used during the trial if they have been taken at a stable dose over at least 4 weeks prior to the IMP application and are maintained at the same stable dose throughout the trial.

Rescue medications as dispensed at the trial clinic may be administered to address ineffective treatment of the treated knee(s) and may be used for other emergent painful conditions if deemed sufficient by the investigator or treating physician. The reason for use of rescue medication should be recorded.

6.5.2 Prohibited Medicines

Any medication or medical intervention likely to interfere significantly with the trial endpoints, apart from dispensed rescue medication, is prohibited during the treatment phase of the trial. Thus, the use of any non-IMP NSAIDs, opioids, COX-2 inhibitors or other analgesic medication including treatment for neuropathic pain, except paracetamol as rescue medication, is not permitted. Low-dose acetylsalicylic acid (aspirin) for cardiovascular protection is allowed. Duloxetine, pregabalin or gabapentin, regardless of the indication, is not allowed during the trial and within 4 weeks of the screening visit, and the use of opioids is not allowed within 6 weeks of the screening visit.

The use of corticosteroids, or hyaluronic acid in any form, administered IA into either knee is prohibited within six months prior to the screening visit, and the use is prohibited during the trial period.

Use of systemic corticosteroids in dose >5mg prednisolone/day or equivalent is not allowed during the trial.

Any systemic or topical analgesic used prior to enrollment in the trial should be discontinued at least 5 half-lives of the particular compound prior to the screening visit, with the exception of duloxetine, pregabalin, gabapentin and opioids requiring a longer wash-out period as discussed above. If the participant does not meet the criteria for analgesic wash-out at the time of the first screening visit (Visit 1a), a second screening visit (Visit 1b) may be planned at which the participant will meet wash out (5 half-lives) of analgesic medication but within 21 days after Visit 1a.

If the participant should experience intolerable pain during the screening period, the participant may use his/her usual analgesic treatment, provided that the medication is discontinued and adequately washed out prior to the screening and baseline visits.

Any topically delivered medicinal treatment at the site of the target knee other than the IMP is not allowed during the entire trial.

For the purpose of efficacy analysis, the impact of prohibited medication/non-permitted therapy use on efficacy evaluation will be assessed and classified to either low or high expected impact. In addition, duration of influence will also be assessed.

No interaction studies have been performed with AMZ001, it is unknown to which extent diclofenac from AMZ001 may interact with drugs which are observed in similar types of diclofenac-containing gels. As a result, the concomitant use of such drugs including, but not limited to *anticoagulants, aspirin, certain antihypertensives (angiotensin converting enzyme inhibitors, angiotensin receptor blocker, diuretics and beta-blockers), lithium, diuretics (furosemide and thiazides), pemetrexed and methotrexate*, should be used with caution.

6.5.3 Other Interventions

Other Permitted Therapies

If a participant is following an existing physiotherapy program at the time of screening, the program can be continued provided that the program is maintained at a stable level of intensity throughout the trial. Initiation of new non-pharmacological therapy such as physiotherapy, or initiation of use of device(s) to relieve pain (e.g., knee brace) from the time of screening is not allowed. Thus, patient education and physical or occupational therapy (e.g., assisting device for walking, diet for weight loss, muscle-strengthening exercises) are allowed if initiated prior to screening and kept stable throughout the trial.

Physical exercise is allowed if initiated prior to screening and stable during the trial. Participants currently regularly exercising will be encouraged not to alter frequency or intensity of the exercise.

Other Non-Permitted Therapies

The use of heating pads or bandages applied to any treated knee is not allowed during the trial period.

Interventional or diagnostic arthroscopy affecting the treated knee(s) is not permitted.

Use of electrotherapy (e.g., transcutaneous electrical stimulation) or acupuncture for OA affecting the treated knee(s) is not permitted.

6.6 Packaging and Labeling of the Investigational Medicinal Product

All IMPs will be packaged and labeled in accordance with all applicable regulatory requirements and Good Manufacturing Practice Guidelines.

The double-blinded IMP products (AMZ001 and placebo) will be identical in packaging and appearance. Each kit will cover at least 1 week of treatment, if both knees are treated, each knee will have its own dose-metered dispenser, which must be properly label to which knee it belongs to.

6.7 Preparation, Handling, and Storage of the Investigational Medicinal Product

All IMPs should be stored at 20°C (68°F) to 25°C (77°F) and should not be frozen. The storage temperature conditions at the investigational site must be registered in a log at least on a daily basis, using either a Continuous or a Min/Max Thermometer. The logs should be available for review as required by the Sponsor or its representatives.

Any temperature excursion out of the established range (15°C–30°C /59°F-86°F) should be immediately reported to the Sponsor through its representative. The Sponsor will determine the viability of use of the affected IMP.

Upon dispensation of IMPs, dose-metered containers must be primed by the study staff to ensure that the correct amount of gel is dispensed upon actuation. Priming is performed by pushing the actuator in the full range of motion a number of times, until some gel has been dispensed. The gel dispensed during priming is discarded. After priming, the kit is weighed as described below in Section 6.8.

6.8 Investigational Medicinal Product Accountability

The Investigator or designee is responsible for ensuring IMP accountability, including reconciliation of drugs and maintenance of records. The accountability of gel will be performed using a standardized scale to determine the weight of each dispensed and returned IMP, which will allow calculation of the quantity of gel used by the participant. The participants will be instructed not to discard/destroy any used kits, and that used kits must be returned to the site for evaluation of compliance.

Upon dispensation, each kit containing dose-metered dispensers will be primed by study staff, as described in Section 6.7. After priming, the weight of each dispenser is recorded.

The participants must be instructed to return any fully used/empty or partly used dispensers or tubes to the site at each visit as described in Section 7.1.

- Upon receipt of shipment of IMP, the responsible person will check for accurate delivery and that the product has been shipped within the defined temperature range;

the responsible person will acknowledge receipt by signing or initialing and dating the appropriate documentation and returning it to the location specified. A copy will be archived for the Investigator Site File.

- IMP dispensing will be recorded on the appropriate drug accountability forms so that accurate records will be available for verification at each monitoring visit.
- Trial site IMP accountability records will include the following:
  - Confirmation of IMP receipt, in good condition and in the defined temperature range.
  - The inventory of IMP provided for the clinical trial.
  - The disposition (including return, if applicable) of any unused IMP.
  - Dates, quantities, batch numbers, expiry dates, and the individual participant trial numbers.

The Investigator site should maintain records which adequately document that participants were provided the doses specified in this protocol, and that all IMPs provided were fully reconciled.

At each visit, the site staff will weigh each dispenser of the returned kit(s), regardless of whether new IMP is planned to be dispensed at the visit in question. The weight of each dispenser of the returned kits will be entered into the eCRF. Compliance will be monitored throughout the study. Site staff will be required to act upon findings of non-compliance with appropriate measures such as re-training of participants in the application procedure.

Unused IMP must not be discarded or used for any purpose other than the present trial.

6.9 Rescue Medication Accountability

During the treatment phase, the Investigator or designee is responsible for ensuring rescue medication accountability, including reconciliation of drugs and maintenance of records.

The quantity of rescue medication used will be assessed via a pill count and compared to the dispensed quantity to determine the used number of rescue medication tablets.

6.10 Assessment of Investigational Medicinal Product Compliance

The total quantity of IMP gel dispensed will be assessed with reference to the expected use of IMP during the double-blind placebo-controlled phase of the trial. Information regarding unilateral or bilateral application of gel during all or parts of the trial will be used to calculate the expected use of gel. The quantity of gel used will be calculated via the accountability assessments.

Repeated insufficient compliance with the protocol requirements regarding IMP application may lead to participant re-training or discontinuation from the trial, as determined by the

Investigator or the Sponsor. Significant non-compliance or insufficient compliance is defined at less than 50% or >150% of the per protocol use of IMP.

6.11 Blinding

The double-blind placebo-controlled trial phase consists of two treatment groups administering either AMZ001 once daily or placebo once daily, as shown above in Figure 2.

The double-blinded IMPs will be indistinguishable in appearance of the container, the label, as well as the gel appearance, smell and texture.

All direct or indirect breaks of the trial blind must be adequately documented, and the Sponsor and its representatives must be informed immediately.

6.12 Emergency Unblinding

The trial blind may be broken by the Investigator for an individual only if knowledge of the IMP is essential for clinical management of the participant. The Investigator is able to unblind the treatment assignment in the eCRF. The Investigator must report and explain the reason for any unblinding of an IMP to the Sponsor or delegate within 24 hours. This notification must be done without revealing the unblinded treatment details to any Sponsor employee, except the designated Pharmacovigilance representative. The Investigator must record the date of unblinding and the reason in the eCRF.

Under certain circumstances, Pharmacovigilance may be required to unblind the treatment assignment for an individual participant following a serious adverse event (SAE) or other serious event; for example, if an expedited regulatory report is required. See Section 7.6.1.1 for further details on expedited reporting and SAEs.

6.13 Treatment of Overdose

The safety profile of diclofenac sodium is well described. The risk of systemic diclofenac toxicity due to excessive topical application of gel is considered very low based on the reduced systemic bioavailability after topical administration, as well as the relatively wide therapeutic window of diclofenac sodium.

Excessive topical application of the IMP may lead to local skin reactions such as erythema, oedema and rash. A study in rabbits indicated that skin irritation caused by AMZ001 was comparable to that produced by commercially available Voltaren® (Diclofenac sodium) gel 1% in the USA. Data from the AMZ001-006 phase 2 trial did not indicate any substantial safety risks associated with use of AMZ001 (Bihlet *et al.*, 2020).

Even if it does not meet other criteria for an SAE, any significant overdose, as judged by investigator or the Sponsor, should be recorded in the eCRF and reported to Pharmacovigilance in an expedited manner using the SAE Report Form, and following the procedure in Section 7.6.1.4.

Management of adverse reactions related to IMP overdosing will follow routine clinical practice using medical judgment.

6.14 Medical Care of Participants after End of Trial

The Sponsor will not provide any additional care to participants after they leave the trial.

7 Trial Procedures and Assessments

An overview of trial procedures and assessments are presented in Table 1.

7.1 Schedule of Assessments

The trial will consist of a screening period, a treatment phase (which will begin with randomization and initiation of therapy at Day 1/Visit 2) and a follow-up period. (See Figure 2)

In case of inconsistencies between trial procedures and assessments in Table 1 and Section 7, Table 1 prevails.

Prior to performing any trial assessments that are not part of routine medical care for the participant, the Investigator will obtain written informed consent as described in Section 5.3.1.

7.1.1 General Instructions

At every visit, before any pain questionnaire is completed, the participants must receive instructions on reporting pain by means of administration of the Placebo Response Mitigation Script as described below in Section 7.1.1.1.

For the baseline visit the questionnaires WOMAC, CPSI, Patient Global Assessment (PGA), and EQ-5D-5L should be completed before any other visit procedures are performed, of which the WOMAC should be completed first. The skin tolerability assessment must be performed prior to application of the IMP. The proposed order is described in more detail below in Section 7.1.3.1.

For subsequent visits, the skin tolerability assessment and the weighing of used IMP containers should be performed before application of the IMP. The IMP should be applied before completing the questionnaires. The proposed order is described in more detail below in Sections 7.1.3.2 to 7.1.3.6.

After the baseline visit (V2), the participants should be instructed to apply the IMP at home in the morning in accordance with the procedures showed during V2, except for the days where they have site visit. All participants must be instructed to bring all available IMP to all treatment phase visits for evaluation of compliance by weighing of the kits. At every visit a new IMP container will be dispensed for each knee being treated.

Starting from the screening visit, the participant will be required to complete a single-question questionnaire once per day; the daily pain NRS using a suitable diary. Any use of rescue medication during the treatment phase will also be recorded in the diary (use and reason for using rescue medication). The site staff will aid the participant in setting up the required applications for completion of this diary, using instructions from the Sponsor or delegate. All

other questionnaires will be completed as electronic PROs, integrated into the eCRF system, at study visits.

The study staff may send reminders to the participants via phone at suitable times to remind them to bring their dispensed IMP kits to the site for evaluation of compliance.

7.1.1.1 Placebo Response Mitigation Script

The trial will include a standardized site staff-administered script, which is to be administered verbally to participants at all on-site visits, prior to any completion of patient-reported outcomes (excluding safety and tolerability assessments). The script is designed to educate participants on the main characteristics and purposes of blinded randomized clinical trials, with the purpose of reducing undue expectations to treatment benefit and improve the accuracy of pain reporting leading to reduced variability in the key efficacy outcomes.

7.1.2 Screening Period (Visits 1a and 1b)

The screening period will comprise a period of up to 28 days before the baseline visit, beginning at signature of the ICF, during which a participant's eligibility for the trial will be determined.

Eligibility will be confirmed at the baseline visit (Day 1/Visit 2) as described in Section 5.4.

The order of assessments is proposed below, with the following requirements:

Signature of the informed consent must be performed before any other procedure is performed. Before collection of pain questionnaires participants must be instructed on placebo response mitigation. WOMAC pain must be completed before blood sampling. X-ray evaluation should be performed as the last eligibility assessment.

Visit 1a (Days -28 to -7)

- Informed consent for the trial (see Section 9.2)
- Verbal administration of the placebo response mitigation script by site staff to the participant
- Collection of "typical OA pain when not taking analgesics"
- WOMAC pain sub-scale (5 questions). This procedure should be done after informed consent is obtained and before any other procedures or examinations are performed. The question should be answered for both knees and hips, since the target knee has not yet been identified. WOMAC should be completed only if the participant has undergone adequate wash-out of analgesic medication and participant reports that his/her typical OA pain in one or both knees when not using medication is ≥ 4 out of 10.

  NOTE: If the participant reports having taken analgesics recently (within 5 half-lives of the analgesic), the WOMAC pain sub-scale (5 questions) will not be taken until Visit 1b.
- Review of concomitant medications (see Section 6.5)

- Collection of demographic characteristics (see Section 7.2)
- Medical history (see Section 7.2)
- Physical examination,
- Collection of substance use
- Vital signs measurement, including height and weight
- 12-lead ECG
- Sample collection for hematology and biochemistry
- Serum pregnancy test for women of childbearing potential
  NOTE: In women of childbearing potential, a negative serum pregnancy test must be obtained before knee X-rays are taken.
- X-rays of both knees (see Section 7.3.1)
- Review of AEs
- Review of inclusion and exclusion criteria

Participants who report having taken analgesics recently (within 5 half-lives of the analgesic) will be asked to withhold all pain medications, including paracetamol (acetaminophen) for 5 half-lives, and Visit 1b will be scheduled accordingly to continue screening. Aspirin (acetylsalicylic acid) used for cardiovascular protection does not need to be withheld.

Participants not taking analgesic medication within at least 5 half-lives of the analgesic at the time of informed consent signature can have Visit 1a and1b procedures completed at the same visit.

Visit 1b (Days -13 to -7)

The following procedures are to be performed for all participants. Participants taking analgesic medications at the time of Visit 1a must undergo washout of all analgesic medications and schedule a follow-up visit (Visit 1b) as described above.

- Verbal administration of the placebo response mitigation script by site staff to the participant
- WOMAC pain sub-score (5 questions). This procedure should be done after informed consent is obtained and before any other procedures or examinations are performed. The question should be answered for both knees and hips, since the target knee has not yet been identified. NOTE: *If Visit 1a and 1b are being combined, do not repeat the WOMAC pain sub-score assessment.*
- Review of AEs and concomitant medications (see Sections 7.6.1 and 6.5)

- Review of inclusion and exclusion criteria

- (For participants who meet the WOMAC pain sub-score criterion at V1a/V1b): Electronic setup of, and instruction in timing of completion of the daily pain NRS questionnaire to be completed off-site.

For participants who are found eligible at Visit 1b, the daily pain NRS will be completed once daily, preferably in the evening, during the remainder of the screening phase (over a period of at least 7 days prior to baseline), starting from Visit 1b until Day 42, both days inclusive, if final eligibility is confirmed and the participant is randomized. For participants who are not eligible for randomization at the Baseline visit, no further data will be collected after this point.

Participants will again be asked to withhold all pain medications for 5 half-lives before the baseline questionnaires (Visit 2).

7.1.3 Treatment Phase (Visits 2 to 7)

The duration of the double-blind treatment phase will be 42 days.

Participant eligibility for the trial will be confirmed at the baseline visit (Visit 2). WOMAC pain sub-score (5 questions) for determination of eligibility is based on the score of the screening and baseline visits.

7.1.3.1 Baseline Visit (Visit 2)

The baseline visit must take place within 28 days after the beginning of screening, following the pain medication washout necessary for confirmation of trial eligibility (see Section 7.1.2).

The below order of assessments is proposed, with the following requirements:

The questionnaires WOMAC, Pain NRS, PGA, CPSI, and EQ-5D-5L should be completed before any other procedures are performed, of which the WOMAC should be completed first. Skin tolerability assessment must be performed prior to first application of the IMP.

After the skin tolerability assessment, the participant should perform the first application of the IMP while s/he is on-site, to allow resolution of questions regarding the application to the study staff, if needed.

The IMP should be applied on either one or both knees, depending on whether one or both knees fulfill the radiographic and the pain criteria for inclusion (see Section 5.3.1).

- Verbal administration of the placebo response mitigation script by site staff to the participant¨

- WOMAC (full questionnaire, 24 questions) on the target knee

- Pain NRS

- PGA

- EQ-5D-5L
- CPSI
- Review of AEs and concomitant medications
- Vital signs measurement
- Urine pregnancy test for women of child-bearing potential
- Review of inclusion and exclusion criteria
- Determination of unilateral or bilateral knee eligibility and choice of target knee
- Skin tolerability assessment
- Sample collection for hematology and biochemistry
- Randomization
- Dispensation, priming and weighing of IMP
- Dispensation of rescue medication and instruction to withhold intake of any rescue medication for at least 24 hours prior to attending the Week 1 visit (Visit 3).
- Instruction how to report use of rescue medication
- Instruction in use and application of the IMP (including "instructions for use" pamphlet) (see Section 6.2.1)
- Application of the IMP on-site under the supervision by the study staff, the time of the application must be recorded.
- Instruction to NOT apply the IMP in the morning of the subsequent visit (Visit 3), and to bring any IMP containers to the site for compliance assessment at Visit 3
- Instruction to withhold intake of any rescue medication for at least 24 hours prior to attending the Week 1 visit (Visit 3).

Following the first application of the IMP, participants will be asked to report their level of pain after the application, 1-hour, 4-hours, 12-hours and 24-hours using the 11-point pain NRS. Daily pain NRS will be collected daily until the last visit.

7.1.3.2 Week 1 (Visit 3)

The order of assessments is proposed below. The following procedures should be completed.

The skin tolerability assessment and the weighing of used IMP containers should be performed before application of the IMP. The IMP should be applied before completing the questionnaires.

The following procedures should be completed:

- Skin tolerability assessment
- Collection and weighing of used IMP for compliance assessment
- Dispensation, priming, and weighing of a new IMP kit. Participants with bilateral OA must be dispensed one bottle for each knee.
- Application of the IMP
- Assess rescue medication washout 24 hours prior to visit.
- Verbal administration of the placebo response mitigation script by site staff to the participant
- WOMAC (full questionnaire, 24 questions) on the target knee
- Review of AEs and concomitant medications
- Vital signs measurement
- Instruction to NOT apply the IMP in the morning of the subsequent visit (Visit 4), and to bring any IMP containers to the site for compliance assessment at Visit 4
- Collection/dispensation of rescue medication, if relevant.
- Instruction to withhold intake of any rescue medication for at least 24 hours prior to attending the Week 2 visit (Visit 4).

7.1.3.3 Week 2 (Visit 4)

The order of assessments is proposed below, with the following requirements:

The skin tolerability assessment and the weighing of used IMP containers should be performed before application of the IMP. The IMP should be applied before completing the questionnaires.

The following procedures should be completed:

- Skin tolerability assessment
- Collection and weighing of used IMP for compliance assessment
- Dispensation, priming, and weighing of a new IMP kit. Participants with bilateral OA must be dispensed one bottle for each knee.
- Application of the IMP
- Assess rescue medication washout 24 hours prior to visit.

- Verbal administration of the placebo response mitigation by site staff to the participant
- WOMAC (full questionnaire, 24 questions) on the target knee
- Review of AEs and concomitant medications
- Vital signs measurement
- Collection of used rescue medication packaging, if applicable (see Section 6.4)
- Dispensation of rescue medication (only if previous supply is depleted and packaging returned to the site)
- Instruction to NOT apply the IMP in the morning of the subsequent visit (Visit 5), and to bring any IMP containers to the site for compliance assessment at Visit 5Instruction to withhold intake of any rescue medication for at least 24 hours prior to attending the Week 3 visit (Visit 5).

7.1.3.4 Week 3 (Visit 5)

The order of assessments is proposed below, with the following requirements:

The skin tolerability assessment and the weighing of used IMP containers should be performed before application of the IMP. The IMP should be applied before completing the questionnaires.

The following procedures should be completed:

- Skin tolerability assessment
- Dispensation, priming, and weighing of new IMP kit. Participants with bilateral OA must be dispensed one bottle for each knee.
- Application of the IMP
- Assess rescue medication washout 24 hours prior to visit.
- Verbal administration of the placebo response mitigation script by site staff to the participant
- WOMAC (full questionnaire, 24 questions) on the target knee
- PGA
- EQ-5D-5L
- CPSI
- Review of AEs and concomitant medications

- Vital signs measurement
- Collection of used rescue medication packaging, if applicable (see Section 6.4)
- Dispensation of rescue medication (only if previous supply is depleted and packaging returned to the site)
- Instruction to NOT apply the IMP in the morning of the subsequent visit (Visit 6), and to bring any IMP containers to the site for compliance assessment at Visit 6
- Instruction to withhold intake of any analgesic medication for at least 24 hours prior to attending the Week 4 visit (Visit 6), if applicable

7.1.3.5 Week 4 (Visit 6)

The order of assessments is proposed below, with the following requirements:

The skin tolerability assessment and the weighing of used IMP containers should be performed before application of the IMP. The IMP should be applied before completing the questionnaires.

The following procedures should be completed:

- Skin tolerability assessment
- Dispensation, priming, and weighing of new IMP kit. Participants with bilateral OA must be dispensed one bottle for each knee. As the following visit will occur only in 2 weeks, two IMP kits must be dispensed for each treated knee. It must be instructed to participant to use one bottle/pump for the first week and the second bottle/pump for the second week even if the first pump is not empty.
- Application of the IMP
- Assess rescue medication washout 24 hours prior to visit.
- Verbal administration of the placebo response mitigation script by site staff to the participant
- WOMAC (full questionnaire, 24 questions) on the target knee
- Review of AEs and concomitant medications
- Vital signs measurement
- Collection of used rescue medication packaging, if applicable (see Section 6.4)
- Dispensation of rescue medication (only if previous supply is depleted and packaging returned to the site)

- Instruction to NOT apply the IMP in the morning of the subsequent visit (Visit 7), and to bring any IMP containers to the site for compliance assessment at Visit 7
- Instruction to withhold intake of any analgesic medication for at least 24 hours prior to attending the Week 6 visit (Visit 7), if applicable

7.1.3.6 Week 6 (Day 43, Visit 7)

The order of assessments is proposed below, with the following requirements:

The following procedures should be completed:

- Skin tolerability assessment
- Collection and weighing of used IMP for compliance assessment
- WOMAC (full questionnaire, 24 questions) on the target knee
- PGA
- EQ-5D-5L
- CPSI
- Review of AEs and concomitant medications
- Vital signs measurement, including weight
- 12-lead ECG
- Collection of used rescue medication packaging, if applicable (see Section 6.4)
- Collect ePRO device.
- Sample collection for hematology and biochemistry
- Urine pregnancy test for women of child-bearing potential

7.1.4 Follow-up period

The duration of the follow-up period will be 14 days.

7.1.4.1 Week 8 - Safety Follow-up Phone Visit (Day 57,Visit 8)

The participant will be contacted by phone:

- Review of AEs, and concomitant medication.

7.2 Demographic and Other Baseline Characteristics

The demographic variables to be documented will be the participant identifier (to be assigned by the eCRF), date of birth, sex, race and ethnicity, in addition to height and body weight.

Weight will also be measured at the final site visit (visit 7).

A complete medical history should be taken, covering all organ systems, diagnoses so far established, medications prescribed (including any vitamins and nutritional supplements taken), and surgical interventions performed. The medical history should also cover aspects such as allergies, consumption of alcohol, and smoking habits. Particular attention should be given to conditions potentially affecting trial eligibility.

Concerning concomitant treatments, particular attention should be given to the participant's use of NSAIDs, COXIBs, and analgesics in general. Information regarding these medications must be painstakingly recorded at screening, as this information is needed for determination of eligibility for the trial. It will also be important to carefully record concomitant medication use throughout the trial.

A physical exam should also be performed. Clinically significant findings of the physical examinations at screening will be documented as concomitant diseases in the medical history section of the eCRF. New findings or worsening of a finding after screening will be recorded as AEs.

Concerning the diagnosis of knee OA, it should be established for how long the disease has persisted. In the absence of an established diagnose, an estimate should be made by the investigator based on the reported symptomatology and radiographic evidence.

The target knee will be selected at the randomization visit (Visit 2) according to the radiographic and symptomatic criteria, more on the selection of the target knee in section 5.4.2.

7.3 Assessment of Eligibility

7.3.1 X-Ray Imaging

X-Ray Image Acquisition

Fixed flexion radiography will be used for assessment of the Kellgren-Lawrence score.

Details are provided in the Imaging Manual.

X-Ray Image Interpretation

X-rays will be read centrally, and an imaging report detailing the results of the reading will be returned to the site.

Kellgren-Lawrence Scoring

The Kellgren-Lawrence scale (Kellgren and Lawrence, 1957) assigns severity grades for knee OA based on radiological assessment of the femorotibial joint (it is important to note that only the femorotibial joint is considered). A number from 0 to 4 is assigned, as follows:

- Grade 0: No feature of osteoarthritis
- Grade 1: Doubtful narrowing of joint space and possible osteophytic lipping
- Grade 2: Definite osteophytes and possible narrowing of joint space
- Grade 3: Moderate multiple osteophytes, definite narrowing of joint space and some sclerosis and possible deformity of bone ends
- Grade 4: Large osteophytes, marked narrowing of joint space, severe sclerosis and definite deformity of bone ends

Kellgren-Lawrence scoring will be performed by the trial's imaging provider and a report will be issued to the site for filing.

7.3.2 Typical OA knee pain

In-clinic patient-reported score of typical OA knee pain when not using medication. Subjects will be asked separately for each knee:

"What is your typical level of knee OA pain in your left/right knee when you do not use pain medication?"

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
None	Mild			Moderate			Severe			Worst Pain Imaginable

7.4 Efficacy Assessments

The effects of treatment on pain, OA symptoms and functional limitation will be evaluated using the WOMAC questionnaire and the daily pain NRS. These assessments are described in the sections below. The OMERACT-OARSI responder rate (Pham *et al.*, 2004) (see Section 7.4.4) at defined points in time will also be determined as a secondary analysis, based on assessment of pain and function using the WOMAC and the PGA, as described in Sections 7.4.1 and 7.4.2.

Evaluation of treatment effects on health-related quality of life will be performed using the EQ-5D-5L and the CPSI.

7.4.1 WOMAC Questionnaire

The self-administered 24-question WOMAC Version 3.1 (Bellamy *et al.*, 1988) will be used to assess pain, physical function, and stiffness in the knee.

Participants will answer all of the 24 questions for themselves, using an 11-box numerical rating scale (with categories of 0 to 10), with reference to the past 24 hours. While the standard WOMAC is designed with a 48 hour recall-period, the index is considered sufficiently robust to tolerate these variations (48 to 24 hours) in the recall period timeframe (Bellamy 2005).

During the treatment phase of the trial, participants will use the WOMAC questionnaire specific to the target knee, regardless of whether both knees are being treated.

The WOMAC is widely used in clinical trials in hip and knee OA and has been extensively validated. Appropriate validated translations will be used in each participating country. The questionnaire addresses the degree of pain experienced with different activities or positions (5 questions), the degree and timing of joint stiffness (2 questions), and the degree of difficulty experienced in performing daily activities (17 questions). The maximum possible total score, indicating the worst possible OA symptoms, is 240 points. For convenience and ease of interpretation, all scores will be normalized on scales of 0 to 100 points in data analysis.

7.4.2 Patient Global Assessment

In the Patient Global Assessment (PGA) used in this trial, participants will be asked to answer the following question using an 11-box numerical rating scale:

"*Considering all the ways your osteoarthritis in your knee affects you, how are you doing today?"."*

7.4.3 The Daily Pain Numerical Rating Scale

Recent OA pain trials have utilized a simple unidimensional pain question taking into account the daily pain as opposed to less frequent, multidimensional measures such as the WOMAC (Chappell *et al.* 2009, Conaghan *et al.* 2018). The specific question to be addressed is:

On a scale of 0-10, please rate your target knee pain by indicating the number that best describes your pain in the last 24 hours with zero being "no pain" and ten being "pain as bad as you can imagine."

The performance in terms of sensitivity to change compared to the WOMAC is not fully established, but an increased resolution of the pain outcome facilitates analysis of the area under the curve (AUC), the weekly average of daily pain, as well assessments of time to onset of effect.

7.4.3.1 Pain NRS in first 24 hours after IMP application.

This question is a variation of the daily pain-NRS, assessed before the IMP has been applied to the target knee and after 1-hour, 4-hours, 12-hours and 24-hours. The question to be addressed is:

"*On a scale of 0-10, please rate your target knee pain by indicating the number that best describes your current level of pain with zero being "no pain" and ten being "pain as bad as you can imagine."*

7.4.4 OMERACT-OARSI Responder Criteria

The OMERACT-OARSI responder criteria involves changes that are deemed to be clinically relevant in the three domains: pain, function, and the PGA (Pham *et al.*, 2004). For each of these domains, ranges are defined for absolute and percent changes from baseline that correspond to "high improvement" and "moderate improvement". OMERACT-OARSI response is defined as either i) high improvement in pain or function or ii) moderate improvement in at least 2 of the 3 domains (pain, function, and PGA). In this trial, OMERACT-OARSI response will be determined using the WOMAC pain sub-score (5 questions), the WOMAC function sub-score (17 questions), and the PGA.

A participant is defined as a responder if they demonstrate either

a) High improvement in <u>at least 1</u> of the WOMAC pain or function sub-scores
or
b) Moderate improvement in <u>at least 2</u> of the WOMAC pain sub-score, WOMAC function sub-score or PGA.

where *high improvement* is defined as

- Relative improvement of ≥ 50% (i.e. relative change ≤ -50%) from baseline AND
- Absolute improvement of ≥ 20 (i.e. absolute change ≤ -20)) from baseline

and *moderate improvement* is defined as

- Relative improvement of ≥ 20% (i.e. relative change ≤ -20%)) from baseline AND
- Absolute improvement of ≥ 10 (i.e. absolute change ≤ -10)) from baseline.

7.5 Assessment of Quality of Life

7.5.1 EQ-5D-5L

The EQ-5D-5L is a standardized generic measure of health related quality of life (Drummond *et al.* 1998). The EQ-5D-5L descriptive questionnaire asks about overall health today in 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression with 5 levels of response (no problems, slight problems, moderate problems, severe problems, extreme problems/unable to). The health profiles are associated with utilities that are often derived from the general population and that can be used to calculate Quality-Adjusted-Life-Years.

The EQ visual analogue scale (EQ-VAS) records the patient's self-rated health on a vertical visual analogue scale, from 0 (the worst imaginable health) to 100 (the best imaginable health). The EQ-VAS can be used as a quantitative measure of health outcome that reflects the patient's own judgement.

7.5.2 Chronic Pain Sleep Inventory

The CPSI-3 is a self-report questionnaire that assesses sleep problems related to pain over a 1-month time interval. The CPSI consists of three items, measured on a 100-mm visual analog scale (VAS) assessing: trouble falling asleep due to pain (CPSI1), awakening by pain during the night (CPSI2) and in the morning (CPSI3), (Ayearst *et al.* 2012). All items have been shown to have good ability to discriminate and respond to changes in the presence and severity of sleep problems attributed to pain. The three-item CPSI questionnaire has been successfully used to document the effects of therapeutic interventions on various sleep characteristics in randomized clinical trials of OA of the hip or knee (Ayearst *et al.* 2012).

7.6 Assessment of Safety

The safety profile of the IMP will be assessed through the recording, reporting and analysis of baseline medical conditions, adverse events (AEs), physical examination findings including vital signs and laboratory tests. Further, an assessment of skin tolerability will be performed (see Section 7.6.3).

Comprehensive assessment of any apparent toxicity experienced by each participant will be performed from the time of giving informed consent and throughout the trial. The investigator will report any AEs, whether observed by the Investigator or reported by the participant (see Section 7.6.1). The reporting period for AEs is described in Section 7.6.1.3.

7.6.1 Adverse Events

7.6.1.1 Adverse Event Definitions

Adverse Event

An AE is any untoward medical occurrence in a participant or clinical investigation participant administered a pharmaceutical product, regardless of causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the use of the medicinal product.

For surgical or diagnostic procedures, the condition/illness leading to such a procedure is considered as the AE rather than the procedure itself.

The Investigator is required to grade the severity or toxicity of each AE.

Investigators must assess the severity of AEs according to the Qualitative Toxicity Scale, as follows:

Mild: The participant is aware of the event or symptom, but the event or symptom is easily tolerated.

Moderate: The participant experiences sufficient discomfort to interfere with or reduce his or her usual level of activity.

Severe: Significant impairment of functioning: the participant is unable to carry out his or her usual activities.

Investigators must also systematically assess the causal relationship of AEs to IMP(s), and rescue medication using the following definitions. Decisive factors for the assessment of causal relationship of an AE to the study treatment include, but may not be limited to, temporal relationship between the AE, the known side effects of diclofenac, medical history, concomitant medication, course of the underlying disease, trial procedures.

Unrelated: Not reasonably related to the study treatment. AE could not medically (pharmacologically/clinically) be attributed to the study treatment under study in this clinical trial protocol. A reasonable alternative explanation must be available.

Related: Reasonably related to the study treatment. AE could medically (pharmacologically/clinically) be attributed to the study treatment under study in this clinical trial protocol.

Abnormal Laboratory Findings and Other Abnormal Investigational Findings

Abnormal laboratory findings and other abnormal investigational findings (for example, on an ECG trace) should not be reported as AEs unless they are associated with clinical signs and symptoms, lead to treatment discontinuation or are considered otherwise medically important by the Investigator. If a laboratory abnormality fulfills these criteria, the identified medical condition (for example, anemia, increased ALT) must be reported as the AE rather than the abnormal value itself.

Serious Adverse Events

An SAE is any untoward medical occurrence that at any dose:

- Results in death.
- Is life-threatening. (Note: The term “life-threatening” refers to an event in which the participant is at risk of death at the time of the event, not an event that hypothetically might have caused death if it was more severe.)
- Requires inpatient hospitalization or prolongs an existing hospitalization.
- Results in persistent or significant disability or incapacity.
- Is a congenital anomaly or birth defect.
- Is otherwise considered to be medically important.

For the purposes of reporting, any suspected transmission of an infectious agent via an IMP is also considered an SAE.

Events that Do Not Meet the Definition of an SAE

Elective hospitalizations to administer, or to simplify trial treatment or trial procedures (for example, an overnight stay to facilitate diagnostic assessment) are not considered SAEs. However, all events leading to unplanned hospitalizations or unplanned prolongation of an elective hospitalization (for example, undesirable effects of any administered treatment) must be documented and reported as SAEs.

Events Not to Be Considered as AEs/SAEs

Medical conditions present at the initial trial visit that do not worsen in severity or frequency during the trial are defined as Baseline Medical Conditions and are not to be considered AEs.

7.6.1.2 Methods of Recording and Assessing Adverse Events

At each trial visit, the participant will be queried on changes in his or her condition. During the reporting period, any unfavorable changes in the participant's condition will be recorded as AEs, whether reported by the participant or observed by the investigator.

Complete, accurate and consistent data on all AEs experienced for the duration of the reporting period (defined below) will be reported on an ongoing basis in the appropriate section of the eCRF. All SAEs must be additionally documented and reported using the appropriate SAE Report Form.

It is important that each AE report include a description of the event, its duration (onset and resolution dates) and times. It is important to assess the time of AE onset relative to the recorded treatment administration time, its severity, its causal relationship with the trial treatment, any other potential causal factors, any treatment given, or other action taken, including dose modification or discontinuation of the IMP, and its outcome. In addition, serious cases should be identified, and the appropriate seriousness criteria documented.

Specific guidance can be found in the eCRF Completion Guide and SAE Reporting Guidelines provided by the Sponsor.

7.6.1.3 Definition of the Adverse Event Reporting Period

The AE reporting period for safety surveillance begins when the participant is initially included in the trial (date of first signature of informed consent/date of first signature of first informed consent) and continues until the final study visit, defined as the Safety Follow-up Phone Visit (V8). Findings made during the physical examination, on the ECG, the screening X-ray, or on laboratory analysis of samples taken on the first screening visit are regarded as existing at the time of signature of informed consent and should be recorded as medical history, not as AEs.

Any SAE assessed as related to study treatment must be reported whenever it occurs, irrespective of the time elapsed since the last administration of IMP.

7.6.1.4 Procedure for Reporting Serious Adverse Events

Serious Adverse Events

In the event of any new SAE occurring during the reporting period, the Investigator must immediately (within a maximum of 24 HOURS after becoming aware of the event) inform the Pharmacovigilance team at NBCD A/S (delegated by the Sponsor) in writing. All written reports should be transmitted using the SAE Report Form, which must be completed by the Investigator following specific completion instructions and sent to the following email address:

PV@nbcd.com

In exceptional circumstances, an SAE (or follow-up information) may be reported by telephone; in these cases, a written report must be sent immediately thereafter e-mail. Names, addresses, and telephone number for SAE reporting will be included in the trial specific SAE Report Form.

Relevant pages from the CRF may be provided in parallel (for example, medical history, concomitant drugs). Additional documents may be provided by the Investigator, if available (for example, laboratory results, hospital report, autopsy report). In all cases, the information provided on the SAE Report Form must be consistent with the data about the event recorded in the CRF.

The Investigator must respond to any request for follow-up information (for example, additional information, outcome, final evaluation, other records where needed) or to any question the Sponsor/designee may have on the AE within the same timelines as those noted above for initial reports. This is necessary to ensure prompt assessment of the event by the Sponsor or designee and (as applicable) to allow the Sponsor to meet strict regulatory timelines associated with expedited safety reporting obligations.

Requests for follow-up will usually be made via the pharmacovigilance department, who may contact the Investigator directly to obtain further information or to discuss the event.

7.6.1.5 Safety Reporting to Health Authorities, Independent Ethics Committees/ Institutional Review Boards and Investigators

The Sponsor or designee will send appropriate safety notifications to Health Authorities in accordance with applicable laws and regulations.

The Investigator must comply with any applicable site-specific requirements related to the reporting of SAEs (particularly deaths) involving trial participants to the IEC/IRB that approved the trial.

In accordance with ICH GCP, the Sponsor/designee will inform the Investigator of “findings that could adversely affect the safety of participants, impact the conduct of the trial or alter the IEC’s/IRB’s approval/favorable opinion to continue the trial”. In particular and in line with respective regulations, the Sponsor/designee will inform the Investigator of AEs that are both serious and unexpected and are considered to be related to the administered product (“suspected unexpected serious adverse reactions” or SUSARs). The Investigator should place copies of

Safety Reports in the Investigator Site File. National regulations with regard to Safety Report notifications to Investigators will be followed.

When specifically required by regulations and guidelines, the Sponsor/designee will provide appropriate Safety Reports directly to the concerned lead IEC/IRB and will maintain records of these notifications. When direct reporting is not clearly defined by national or site-specific regulations, the Investigator will be responsible for promptly notifying the concerned IEC/IRB of any Safety Reports provided by the Sponsor/designee and of filing copies of all related correspondence in the Investigator Site File.

For trials covered by the Clinical Trials Regulation No 536/2014, the Sponsor's responsibilities regarding the reporting of SAEs/SUSARs/Safety Issues will be carried out in accordance with that Regulation and with the related Detailed Guidance documents.

7.6.1.6 Monitoring of Participants with Adverse Events

AEs are recorded and assessed continuously throughout the trial (see Section 7.6.1.3) and are assessed for the final outcome at the final study visit. All SAEs ongoing at the final study visit must be monitored and followed up by the Investigator until stabilization or until the outcome is known, unless the participant is documented as "lost to follow-up". Reasonable attempts to obtain this information must be made and documented. It is also the responsibility of the Investigator to ensure that any necessary additional therapeutic measures and follow-up procedures are performed.

7.6.2 Pregnancy and In Utero Drug Exposure

Only pregnancies considered by the Investigator to be related to trial treatment (for example, resulting from a drug interaction with a contraceptive medication) are considered to be AEs. However, all pregnancies with an estimated conception date during the period defined in Section 7.6.1.3 must be recorded by convention in the AE page/section of the CRF. The same rule applies to pregnancies in female participants and to pregnancies in female partners of male participants. The Investigator must notify the Sponsor/designee in an expedited manner of any pregnancy using the Pregnancy Report Form, which must be transmitted according to the same process as described for SAE reporting in Section 7.6.1.4.

Investigators must actively follow up, document and report on the outcome of all these pregnancies, even if the participants are withdrawn from the trial.

The Investigator must notify the Sponsor/designee of these outcomes using the Pregnancy Report Form. If an abnormal outcome occurs, the SAE Report Form will be used if the participant sustains an event and the Parent-Child/Fetus Adverse Event Report Form if the child/fetus sustains an event.

Any abnormal outcome must be reported in an expedited manner as described in Section 7.6.1.4, while normal outcomes must be reported within 45 days after delivery.

In the event of a pregnancy in a participant occurring during the course of the trial, the participant must be discontinued from trial medication immediately. The Sponsor/designee must be notified without delay and the participant must be followed as mentioned above.

7.6.3 Skin Tolerability Assessment

Based on experience from pre-clinical and clinical trials with AMZ001, in addition to the current label information of approved diclofenac gels e.g., Voltaren Emulgel (diclofenac diethylamine) , skin irritation or other application site reactions may occur. In order to adequately evaluate the local tolerability and safety of AMZ001, the Investigator or delegate will complete an IMP application site evaluation at each study visit. To establish a pre-treatment reference, a skin assessment will be performed at the baseline visit prior to first IMP application.

The skin will be assessed with the Local skin responses (LSR) grading scale (Rosen *et al.* 2014). The scale measures presence and severity (0-4) of six clinical characteristics that are commonly observed LSRs from topical treatments. The grading of the application site will be performed in accordance with Figure 3 below. A reference guide with photographic examples of each grade will be available to the Investigator to ensure that a standardized evaluation is performed across study sites.

As the IMP may be applied to more than one knee, the application site assessment will be limited to the target knee. The result of each evaluation is entered into the eCRF on a dedicated page. Skin evaluations of grade 2 of the target knee or above should also be recorded as an adverse event.

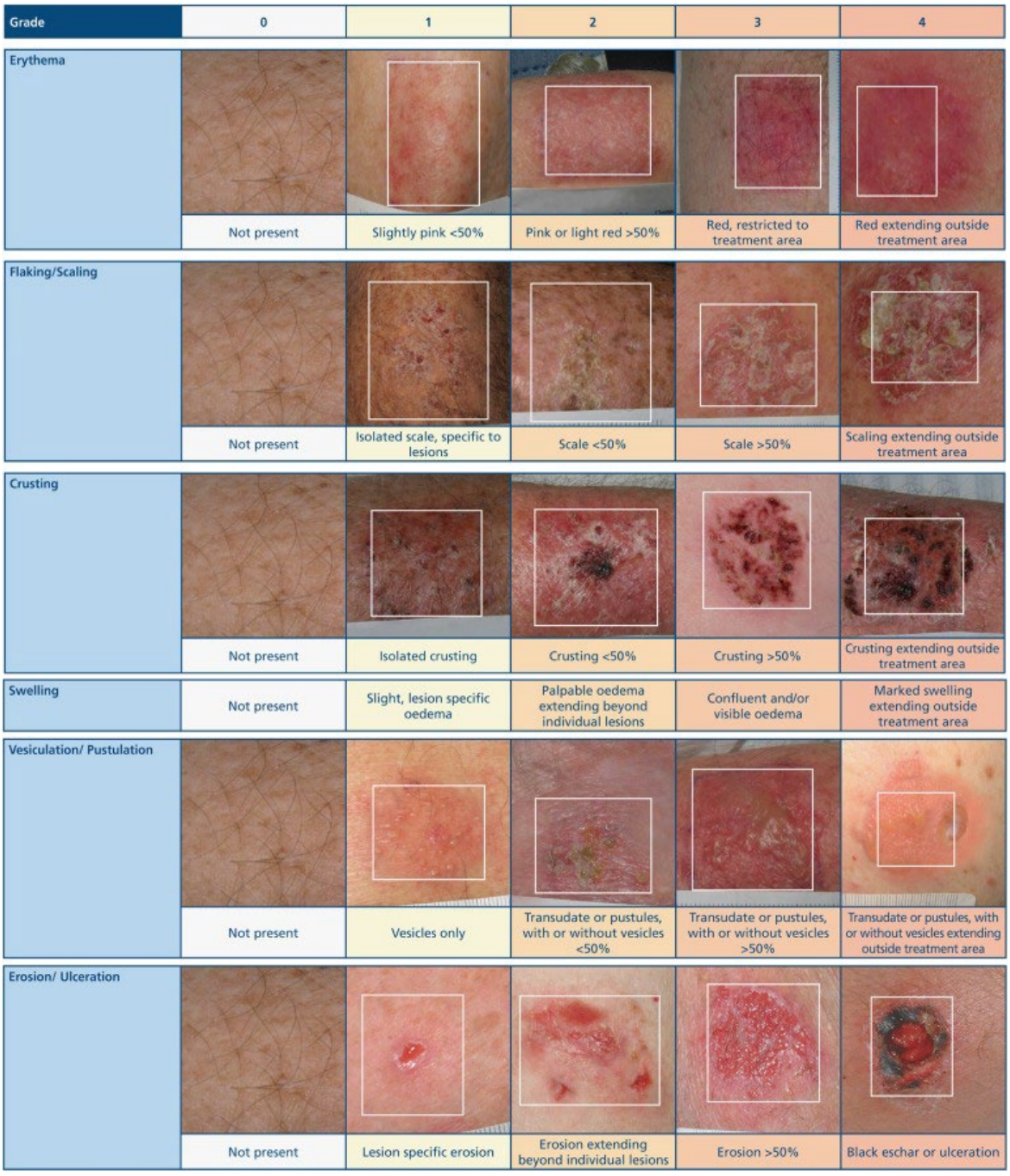


Figure 3: Local Skin Responses (LSR) grading scale (Rosen et al. 2014)

7.6.3.1 Skin Tolerability symptoms

At each site visit the investigator will inquire on self-perceived symptoms commonly reported in topical treatments. Pain, itch and burning sensation will be individually assessed using a numeric rating scale from 0-10. These symptoms will be assessed for each knee being treated.

The questions to be addressed are:

For burning: *On a scale of 0-10, please rate the burning sensation in the skin of your target knee by indicating the number that best describes the burning sensation on average during the last 7 days with zero being "no burning sensation" and ten being "burning sensation as bad as you can imagine"*

For itch: *On a scale of 0-10, please rate the itching sensation in the skin of your target knee by indicating the number that best describes the itching sensation on average during the last 7 days with zero being "no itching" and ten being "itching as bad as you can imagine"*

For pain: *On a scale of 0-10, please rate the pain in the skin of your target knee by indicating the number that best describes your pain on average during the last 7 days with zero being "no pain" and ten being "pain as bad as you can imagine".*

7.6.4 Clinical Laboratory Assessments

Blood samples will be collected for clinical laboratory tests as described in Appendix C, following the timing noted in the Schedule of Assessments. Participants are not required to be fasting at the time of sample acquisition.

The acquisition, handling, shipping and storage of samples is described in the study Laboratory Manual.

Pregnancy Tests

For women of childbearing potential, pregnancy tests will be performed at the time points shown in Table 1. A serum test will be performed during screening, and urine tests will be performed at subsequent time points.

Management of Participants with Laboratory Abnormalities

Laboratory abnormalities that Investigators consider to be clinically significant should be reported as AEs (see Section 7.6.1.2).

The best interest of the participating participant will always come before that of the trial. Abnormalities of any safety laboratory analysis considered to represent a significant danger to the participant will lead to immediate discontinuation of the drug, and the Medical Responsible must be informed immediately.

In the event of significant abnormalities considered not to represent a danger, continuation of the drug may be allowed after discussion with the Sponsor's designated Medical Responsible.

7.6.5 Vital Signs, Physical Examinations, and Other Assessments

Physical examinations and measurement of height, weight, and vital signs (sitting systolic and diastolic blood pressure, heart rate, and body temperature) will be performed at the time points specified in Table 1.

Standard 12-lead ECGs (reporting ventricular rate, PR, QRS, QT, and QTc intervals) will be obtained at the time points specified in Table 1. ECGs will be read locally by study personnel. ECGs can be repeated at the discretion of the Investigator or Sub-Investigator.

8 Statistics

8.1 Sample Size

The primary end point will be change from baseline in the WOMAC pain sub-score (questions 1 to 5) in the target knee as evaluated at week 2 using AMZ001 as compared to placebo. The treatment allocation will be balanced, i.e. 1:1, corresponding to applying AMZ001 and placebo, respectively.

The power calculation is based on data from a 4-week double-blind study of AMZ001 QD or BID vs. placebo in a study population highly similar to that proposed in the current trial.

The assumed difference between AMZ001 QD and placebo is based on a sub-group analysis of AMZ001-006, consisting of participants with WOMAC pain scores of ≥ 20 and ≤ 45 out of 50 at both screening AND baseline visits. In AMZ001-006 this subgroup comprised 89.4% of the total (mITT) population (Bihlet *et al.*, 2020).

It is assumed that approximately 20% of participants will discontinue the trial before the Week 6 time point. Assuming a common standard deviation of 19.1 (out of 100) in the WOMAC pain sub-score, analysis of 216 evaluable participants per treatment group will result in minimally detectable difference between placebo and active treatment of 5.9 (out of 100) with a two-sided 5% level of significance, with 90% power. As the proposed analysis methods will be based on all available data from all randomized participants, and not only study completers, this estimate is considered a conservative estimate.

The effect size for the power calculations (minimal detectable difference of 5.9 and standard deviation of 19.1) corresponds to a Cohen's d of 0.31, which is the ratio of the effect divided by standard deviation of the estimate. To examine the sensitivity of the assumed effect size on the resulting sample size, the table below shows required sample size with 90% power if the Cohen's d is varied between 0.20 and 0.55. The statistical software R version 4.3 (function power.t.test) has been used for sample size estimation.

Table 2: Sensitivity analysis showing how sample size changes with varying assumptions about the minimum detectable effect in terms of Cohen's d

(Effect/SD)	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**	**0.55**
Participants per group	527	338	235	173	133	105	86	71
Participants per group + 20% dropout	659	423	294	217	167	132	108	89
Total sample size	1318	846	588	434	334	264	216	178

8.2 Randomization

The screening/randomization procedure will be centrally managed through an electronic Interactive Web Response System (IWRS) integrated in the eCRF. To enroll and randomize a new participant, the Investigator/authorized personnel will have to access the eCRF by entering their username and password and filling in the requested data.

Eligible participants will be randomized in a 1:1 ratio to one of two treatment groups. Each of the two double-blinded treatment groups of AMZ001 and placebo will be randomized with approximately 270 participants for the primary objective of assessment of an analgesic effect of topical AMZ001 gel compared with placebo.

Randomization will be stratified by country to ensure balance of treatment groups with respect to intrinsic and extrinsic factors that may affect OA outcomes.

8.3 Endpoints

8.3.1 Primary Endpoint

The primary endpoint of this trial is the change from baseline in WOMAC pain sub-score (questions 1 to 5) in the target knee as evaluated at week 2.

8.3.2 Secondary Endpoints

For the purposes of confirmatory testing, the following are identified as Key Secondary Endpoints:

- Change from baseline in the WOMAC pain sub-score at week 1, in the target knee.
- Change from baseline in pain at Day 8, Day 7, Day 6, Day 5, Day 4, Day 3, Day 2, and Day 1 using an 11-point Numeric Rating Scale (NRS)

The remaining Secondary Endpoints as listed below are considered Supportive Secondary Endpoints

- Partial AUCs of the observed change from baseline based on pain recorded at 24 hours, 12 hours, 4 hours and 1 hour after the first drug application using an 11-point Numeric Rating Scale (NRS).
- Change from baseline in PGA at weeks 3 and 6
- Changes from baseline in the Chronic Pain Sleep Inventory at weeks 3 and 6
- Change from baseline in the weekly average of daily pain recorded using an 11-point pain NRS until week 6 in the target knee
- Change from baseline in the WOMAC pain sub-score during the trial in the target knee (weeks 3, 4 and 6)

- Change from baseline in the WOMAC function sub-score during the trial (weeks 1,2, 3,4 and 6).
- Change from baseline in the WOMAC total score during the trial (weeks 3 and 6).
- Changes from baseline in quality of life as assessed using the EQ-5D-5L at weeks 3 and 6
- Proportion of participants meeting the OMERACT-OARSI responder criteria at weeks 3 and 6
- Proportion of participants meeting at least a 30 % reduction from baseline in WOMAC pain sub-score at Week 3 and 6.
- Proportion of participants meeting at least a 50 % reduction from baseline in WOMAC pain sub-score at Week 3 and 6.
- Time (days) to first use of rescue medication since baseline
- Average daily amount of rescue medication taken at Weeks 1-6
- Weekly proportion of days used rescue medication at Weeks 1-6

8.3.3 Pre-defined subgroup analyses

Subgroup analysis includes high vs. low baseline pain scores, and Kellgren-Lawrence grade.

Subpopulation analyses (especially for the analysis of the primary endpoint) will be considered from the ITT Analysis Set excluding the participants with a BMI >40kg/m^2 and participants who have a baseline WOMAC pain sub-score >45.

Further subgroup analyses may be considered and defined in the SAP.

8.3.4 Safety Endpoints

Safety endpoints are:

- Nature, incidence, and severity of AEs
- Changes in laboratory safety parameters, vital signs, 12-lead ECG parameters, and weight
- Nature, incidence, and severity of skin reactions at the application site

8.4 Analysis Sets

Intention-to-Treat Analysis Set

The Intention-to-Treat (ITT) Analysis Set will include all participants randomly allocated to a treatment, based on the intention to treat “as randomized” principle (i.e., the planned treatment

regimen rather than the actual treatment given in case of any difference). The ITT set will be used for efficacy endpoints.

Safety Analysis Set

The Safety Analysis Set will include all participants who have taken at least one application of trial treatment. Participants will be analyzed according to the actual treatment they receive. The Safety Analysis Set will be used for safety endpoints.

Per Protocol Set is not defined as clinical efficacy of trial product will be assessed using the secondary estimand.

8.5 Description of Statistical Analyses

8.5.1 General Considerations

An outline of the analysis currently planned is presented in the following sections. A detailed Statistical Analysis Plan (SAP) will be written and finalized before database lock and unblinding. The SAP will provide full and final details of all planned data analyses and data display.

After the last participant's last visit of the treatment phase, the database will be locked, the treatment code will be unblinded, complete analyses will be performed, and a detailed Clinical Trial Report will be written based on the data collected up to the end of the treatment phase.

Results from statistical analyses will generally be accompanied by two-sided 95% confidence intervals and corresponding p-values. Superiority will be claimed if p-values are less than 5% and the estimated treatment contrast favors AMZ001.

The ITT set will be used for efficacy endpoints, the Safety Analysis Set for safety endpoints. For safety assessments, the last available data before dosing will be considered as baseline. For efficacy endpoints, the values collected at Baseline Visit (Day 1, Visit 2) will be considered as baseline. For onset of action using NRS at 1, 4, 12 and 24 hours, and for NRS key secondary endpoints, baseline will be computed as average of all available responses in the last 7 days prior to Baseline.

Summary Statistics

The summary statistics presented for quantitative variables will be the number of observations (n), mean, standard deviation (SD), median, first and third quartiles (Q1, Q3), and minimum (min), and maximum values (max).

The summary statistics presented for categorical data will be the number of observations (n), the count and percentage of participants in each category, taking the non-missing participant count for the denominator.

Adjustment for multiplicity

The primary endpoints and key secondary endpoints will be evaluated in hierarchical manner, testing hypotheses in sequential order, passing unused alpha from a successful hypothesis test

to the subsequent hypotheses at lower level. If at a certain step the tested hypothesis is not successful, the alpha associated with that hypothesis will be considered spent and will not be passed further down.

- Testing will start at the primary endpoint of change from baseline in WOMAC pain sub-score at week 2 at alpha = 0.05.
- If successful, change from baseline in WOMAC pain sub-score at week 1 will be tested at alpha = 0.05.
- If successful, the change from baseline in pain relief at Day 8 (NRS) will be tested at alpha = 0.05.
- This is then continued for days 7, 6, 5, 4, 3, 2 and 1, one at a time, in the order that they are listed.

See below for graphical representation of how the sequence of hypothesis testing and delegation of alpha will be executed.

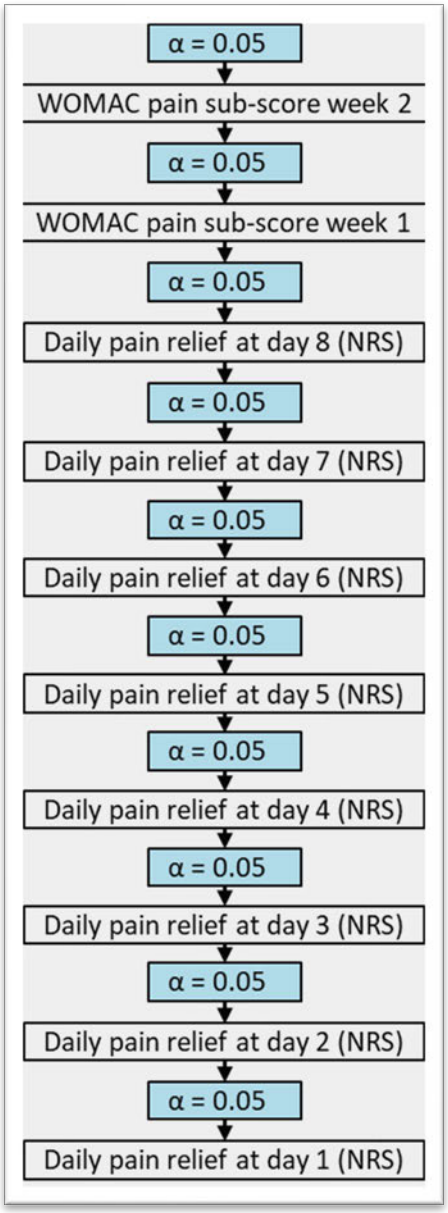


Figure 4: Sequence of hypothesis testing and delegation of alpha

Intercurrent events

- See Section 4.3 on estimands.

Missing data

Missing data will (potentially) be handled using two approaches. Imputation defined by the authors of questionnaires (if any) and imputation defined by the estimands/general missing data strategies. Imputation defined by the authors of the questionnaires is always carried out before the imputations defined by the estimands/general missing data strategy. Imputation techniques for each estimand, for primary and sensitivity analyses, will be presented below along with underlying imputation assumptions.

8.5.2 Analysis of Primary Endpoint

The primary endpoint is the change from baseline in WOMAC pain sub-score (questions 1 to 5) in the target knee as evaluated at week 2.

8.5.2.1 Primary estimand for the primary endpoint

The primary estimand for the primary endpoint will be the "treatment attributable" estimand defined in Section 4.3.1 Primary estimand for continuous endpoints. The analysis model will be ANCOVA of change from baseline at the corresponding visit with baseline as covariate and treatment group & pooled center as factors.

Data after ICEs of treatment discontinuation (for treatment related reasons) and data affected by use of prohibited medication/non-permitted therapies with high expected impact is discarded (hypothetical strategy) and imputed based on MNAR (return-to-baseline). Rescue medication use & prohibited medication with low expected impact is dealt with treatment policy strategy (use data as is).

Handling of missing data and analysis:

The described imputation procedure will impute values up to and including week 6 at visits during which the data was collected to be used for the secondary endpoints as well. This section only outlines the procedure, the final details will be given in the SAP.

Missing data arising from study withdrawals for possibly treatment related reasons will be imputed under MNAR, return to baseline. Otherwise, data will be assumed to be missing under MAR. Likewise, the general assumption is that intermittent missing data is missing under MAR. Handling of missing data after the occurrence of various ICEs is described above. Imputations will involve the following steps:

1) Within each treatment arm, the algorithm for multiple imputation (MI) will impute intermittent values, if any, using MCMC method to reach monotone pattern.

2) Within each treatment arm, a sequential series of univariate linear regression models will be used to sample regression parameters from posterior distributions , each model containing the baseline value and values from the previous post-baseline visit models as covariates, and pooled center (if feasible) as factor.

3) Within each treatment arm, using the parameters obtained in step 2), missing data will be imputed sequentially conditioning on values at previous visits, observed and imputed, using regression models defined by parameters drawn at step 2). Samples will be drawn from truncated normal distributions to ensure imputed values fall within the allowed interval.

4) The values assumed to be MNAR will then be replaced by return-to-baseline imputed values. Details of return-to-baseline MNAR imputation will be described in SAP, based on Wang et al., 2023.

5) Each of the above steps will be replicated at least 100 times.

For each imputed dataset, the endpoint (assessment at Week 2) will be analyzed using an ANCOVA model of change from baseline that contains baseline value as covariate, and treatment group and pooled center as factor.

The estimates and standard errors from each of the imputed dataset analyses will be combined using Rubin's rule to form a unique point estimate and standard error. The estimated treatment difference to placebo will be presented with a 95% confidence interval (CI). Significant p-value of treatment group regression coefficient with estimate that points in favor of Diclofenac Gel vs. Placebo is needed to claim success.

A sensitivity analysis that will be described in detail in SAP will handle missing data after study withdrawals where imputations are done using pooled data across the two arms.

Potential other sensitivity analyses exploring robustness of this analysis will be detailed in SAP.

8.5.2.2 Secondary estimand for the primary endpoint

The secondary estimand for the primary endpoint will be the "efficacy" estimand defined in section 4.3.3 Secondary estimands.

Data post treatment discontinuation and use of prohibited medication/non-permitted therapies with high expected impact will not be used in this analysis. Analysis will be done using mixed model for repeated measures (MMRM). The MMRM model will use change from baseline in WOMAC pain sub-score as outcome and the same factors and covariates as for the primary analyses, all nested within the post baseline visits at which WOMAC pain sub-score is collected. An unstructured covariance matrix for measurements within the same participant will be employed. For the primary endpoints, the model's estimates at week 2 will be reported.

8.5.3 Analysis of Secondary Endpoints

The continuous variables of the secondary endpoints will be analyzed in a similar manner as the primary endpoint using the primary estimand.

For weekly average of NRS daily scores, the minimum required number of non-missing scores if 4. If less than 4 daily NRS are available in a given week, that week's average NRS score will be considered missing. For each week, the weekly average will be computed based on all available weekly scores.

Change from baseline in quality-of-life VAS as assessed using the EQ-5D-5L questionnaire will be analyzed in a similar manner as the primary endpoint. The 5 descriptive domains will be summarized descriptively.

Each domain of CPSI is evaluated on a 100-mm VAS and will be analyzed separately, in a similar manner as the primary endpoint.

8.5.3.1 Analysis of OMERACT-OARSI responder criteria

For the OMERACT-OARSI responder rate, the estimand for binary endpoints will be used.

OMERACT-OARSI Responder Criteria

The OMERACT-OARSI responder criteria (Pham, et al., 2004) are based on the WOMAC pain and function sub-scores and the Patient Global Assessment (PGA) questionnaire.

A participant is defined as a responder if they demonstrate either

a) High improvement in at least 1 of the WOMAC pain or function sub-scores
or
b) Moderate improvement in at least 2 of the WOMAC pain sub-score, WOMAC function sub-score or PGA.

where *high improvement* is defined as

- Relative improvement of ≥ 50% (i.e. relative change ≤ -50%) from baseline AND
- Absolute improvement of ≥ 20 (i.e. absolute change ≤ -20)) from baseline

and *moderate improvement* is defined as

- Relative improvement of ≥ 20% (i.e. relative change ≤ -20%)) from baseline AND
- Absolute improvement of ≥ 10 (i.e. absolute change ≤ -10)) from baseline.

and

relative change = percentage change during the study
= ((final score – baseline score)/ baseline score) * 100

absolute change = absolute change during the study
= final score (normalized) – baseline score (normalized)

In the situation that at least one of the three endpoints (i.e. WOMAC pain and function sub-scores and PGA) required to assess the responder criteria is missing:

- If the available endpoints satisfy the responder criteria, then the participant will be classified as a responder. (e.g. If a participant has high improvement on the WOMAC pain sub-score but the WOMAC function sub-score and PGA are missing then the participant is a responder).
- If the available endpoints do not satisfy the responder criteria AND there is no possibility that they could have been a responder when considering possible values for the missing score(s), then the response will be set to non-response. (e.g. If a participant does not achieve moderate response on either the WOMAC pain sub-score or the PGA and the WOMAC function sub-score missing then the participant is a non-responder).

If the available endpoints do NOT satisfy the responder criteria BUT the participant could have been a responder depending on the possible values for the missing score(s)), then the response will be set to missing (e.g. If either of the WOMAC pain or function sub-scores are missing then the response rate will be set to missing).

Data after treatment discontinuation (for treatment related reasons) and data affected by use of prohibited medication/non-permitted therapies with high expected impact are viewed as non-

response. Rescue medication use & prohibited medication with low expected impact is dealt with treatment policy strategy (use data as is).

Imputation of missing data for OMERACT-OARSI responder criteria endpoint

Missing data arising due to study withdrawals for possibly treatment related reasons will be considered non-response. Otherwise, data will be assumed to be missing under MAR. Likewise, the general assumption is that intermittent missing data is missing under MAR. Handling of missing data after the occurrence of various ICEs is described above.

First, for the intermittent missing values, the approach proposed in (Multiple Imputation and its Applications, Carpenter et al., 2023) will be used, where exactly the same MCMC imputation algorithm as described for continuous endpoints will be used, treating the binary response as continuous. After imputation, imputed values are rounded to nearest value of the binary variable. That is, in order for Y to take on values of either 0 or 1, values < 0.5 are rounded to 0 and values > 0.5 are rounded to 1. Once monotone pattern is achieved, the same steps will be followed as for continuous endpoints, replacing the linear regression model with logistic regression model containing responder status at previous post-baseline visits and pooled center (if feasible) as factors and baseline WOMAC pain and function sub-scores as covariates, producing at least 100 simulated datasets that will each be analyzed using logistic regression with treatment group and pooled center as factors and baseline WOMAC pain and function sub-scores as covariates.

The estimates and standard errors from each of the imputed dataset analyses will be combined using Rubin's rule to form a unique point estimate and standard error. The estimated odds ratio versus placebo will be presented with a 95% CI.

Additionally, the average adjusted risk difference may be presented together with a plot of individual estimated adjusted risk under both treatments.

For the secondary estimand, if done, data post treatment discontinuation or due to use of prohibited medication/non-permitted therapies with high expected impact will instead be deemed missing and imputed assuming MAR, i.e., as for unlikely treatment related discontinuations, and analyzed as described above.

8.5.3.2 WOMAC pain responder endpoints

Unlike OMERACT-OARSI composite responder endpoint, missing data for 30% and 50% WOMAC pain responder endpoints will be derived directly from the imputed WOMAC pain scores. The analysis model will remain the logistic regression with baseline WOMAC pain as covariate and pooled center and treatment as factors.

8.5.3.3 Rescue medication endpoints

No estimands will be applied for these endpoints.

Time to first use of rescue medication due to pain in target knee will be summarized using statistics for time-to-event data, e.g., median time to event, 25^{th} & 75^{th} percentile time to event, number at risk, number experienced event. Comparison between groups will be done using

Cox proportional hazard model with baseline WOMAC pain as covariate and pooled center and treatment as factors.

Weekly proportion of days used rescue medication will be summarized at each week. For a given week, a subject must have logged whether they have used rescue medication or not at least 4 out of 7 days, otherwise that week's data is considered missing. Comparison between groups at each week will be done using logistic regression model, where response is binomial number of days used rescue medication out of all days logged use in the corresponding week with baseline WOMAC pain as covariate and pooled center and treatment as factors.

The average daily amount of rescue medication used will be summarized at each visit. Comparison between groups at each visit will be done using linear regression model, where response is average daily amount of rescue medication taken between visits with WOMAC pain as covariate and pooled center and treatment as factors. If the linear model shows clear signs of lack of fit Wilcoxon rank-sum test will be used instead.

8.5.4 Analysis of Exploratory Endpoints

N/A

8.5.5 Analysis of Safety

Safety data will be summarized descriptively overall and by treatment group, using the Safety Analysis Set.

AEs will be coded using the latest available version at study start of the Medical Dictionary for Regulatory Activities (MedDRA).

All AEs will be reported overall, by severity, and by relationship to trial drug, and summaries will be prepared of events leading to discontinuations.

Laboratory safety parameters, vital signs, ECG parameters, and changes from baseline in each of these parameters will be descriptively summarized as appropriate.

Skin reactions on the application site will be descriptively summarized as appropriate.

8.5.6 Participant Demographics, Baseline Characteristics and Study Treatments

The following domains will be described overall and by treatment group using summary statistics:

- Disposition of participants and discontinuations
- Protocol deviations
- Demographics, medical history, and other baseline characteristics
- Past and concomitant medications with particular attention to pain medications

- Treatment exposure and compliance, as assessed by duration of exposure, total exposure, average daily exposure, and compliance.

9 Ethical and Regulatory Aspects

9.1 Responsibilities of the Investigator

The Investigator is responsible for the conduct of the trial at the site and will ensure that the trial is performed in accordance with this protocol, the ethical principles outlined in the Declaration of Helsinki, ICH GCP and any other applicable regulations. The Investigator must ensure that only participants who have given informed consent are included in the trial.

According to United States Code of Federal Regulations Part 54.2 (e), for trials conducted in any country that could result in a product submission to the United States Food and Drug Administration for marketing approval and could contribute significantly to the demonstration of efficacy and safety of an IMP (which are considered “covered clinical trials” by the FDA), the Investigator and all sub-investigators are obliged to disclose any financial interest which they, their spouses or their dependent children may have in the Sponsor or the Sponsor’s product under study. This information is required during the trial and for 12 months following completion of the trial.

9.2 Participant Information and Informed Consent

An unconditional prerequisite for each participant prior to participation in the trial is written informed consent, which must be given before any trial-related activities are carried out. Adequate information must therefore be given to the participant by the Investigator or an appropriate designee (if local regulations permit) before informed consent is obtained.

A participant information sheet must be prepared in the local language in accordance with ICH GCP and will be provided by the Sponsor for the purpose of obtaining informed consent. In addition to providing this written information to a potential participant, the Investigator or a designate will inform the participant verbally of all pertinent aspects of the trial, using language chosen so that the information can be fully and readily understood by laypersons. The participant will be given sufficient time to read the information and the opportunity to ask questions and to request additional information and clarification.

If permitted by national regulations, a person other than the Investigator may inform the participant about the trial and sign the Informed Consent Form, as above.

After the information is provided by the Investigator, the Informed Consent Form must be signed and dated by the participant and the Investigator.

The signed and dated declaration of informed consent will remain at the Investigator’s site and must be safely archived so that the forms can be retrieved at any time for monitoring, auditing and inspection purposes. A copy of the signed and dated information and Informed Consent Form should be provided to the participant prior to participation.

Whenever important new information becomes available that may be relevant to informed consent, the Investigator will revise the participant information sheet and any other written information to be provided to the participants and submit them to the IRB/IEC for review and opinion. Using the approved revised participant information sheet and other written information, the Investigator will explain the changes to the previous version to each trial participant and obtain new written consent for continued participation in the trial. The participant will be given sufficient time to read the information and the opportunity to ask questions and to request additional information and clarification about the changes.

9.3 Participant Identification and Privacy

A unique screening number will be assigned to each participant immediately after informed consent has been obtained. This number will serve as the participant's identifier in the trial as well as in the clinical trial database. All participant data collected in the trial will be stored under the appropriate subject ID. Only the Investigator will be able to link trial data to an individual participant via an identification list kept at the site. For each participant, original medical data will be accessible for the purposes of source data verification by the Monitor, audits and regulatory inspections, but patient confidentiality will be strictly maintained.

Data protection and privacy regulations will be observed in capturing, forwarding, processing, and storing participant data. Participants will be informed accordingly and will be requested to give their consent on data handling procedures in accordance with national regulations.

9.4 Emergency Medical Support and Participant Card

Participants will be provided with Emergency Medical Support cards supplied by the Sponsor for use during trial participation in order to provide clinical trial participants with a way of identifying themselves as participating in a clinical trial and to give health care providers access to any information about this participation that may be needed to determine the course of medical treatment for the participant.

The first point of contact for all emergencies will be the clinical trial Investigator caring for the affected participant. The Investigator agrees to provide his or her emergency contact information on the card for this purpose. If the Investigator is available when an event occurs, they will answer any questions. Any subsequent action (for example, unblinding) will follow the standard process established for Investigators.

9.5 Clinical Trial Insurance and Compensation to Participants

Insurance coverage will be provided for each country participating to the trial. Insurance conditions shall meet good local standards, as applicable.

9.6 Independent Ethics Committee or Institutional Review Board

Prior to commencement of the trial at a given site, this clinical trial protocol will be submitted together with its associated documents to the responsible IEC or IRB for its favorable opinion or approval, which will be filed in the Investigator Site File. A copy will be filed in the Sponsor Trial Master File.

Written evidence of favorable opinion or approval that clearly identifies the trial, the clinical trial protocol version and the Participant Information and Informed Consent Form version reviewed should be provided. Where possible, copies of the meeting minutes should be obtained.

Amendments to this clinical trial protocol will also be submitted to the concerned IEC or IRB, before implementation of substantial changes. Relevant safety information will be submitted to the IEC or IRB during the course of the trial in accordance with national regulations and requirements.

Plans for any substantial amendments to the clinical trial will also be submitted to the concerned IRB before they are implemented. Relevant safety information will be submitted to the IRB during the course of the trial in accordance with national regulations and requirements.

9.7 Health Authorities

The clinical trial protocol and any applicable documentation (for example, Investigational Medicinal Product Dossier, Participant Information and Informed Consent Form) will be submitted or notified to the Health Authorities in accordance with all local and national regulations for each site. For European sites, any serious breach will be notified via CTIS.

10 Trial Management

10.1 Case Report Form Handling

The main purpose of the eCRF is to obtain data required by the clinical trial protocol in a complete, accurate, legible and timely manner. The data in the eCRF should be consistent with the relevant source documents.

The Investigator or designee is responsible for ensuring that the data collected in the course of this trial is accurate and documented. They will then be processed, evaluated, and stored in anonymous form in accordance with applicable data protection regulations. The Investigator must ensure that the eCRFs and any other associated documents forwarded to Sponsor, or its designated organization contain no mention of any participant names or other patient-identifiable information.

For patient-reported outcome data such as QoL and pain assessments, ePRO will be used.

The data will be entered into a validated database. The Sponsor or its designee will be responsible for data processing, in accordance with the Sponsor’s data management procedures. Database lock will occur once quality control and quality assurance procedures have been completed. PDF files of the eCRFs will be provided to the Investigators at the completion of the trial.

10.2 Source Data and Participant Files

The Investigator must keep a file (medical file, original medical records) on paper or electronically for every participant in the trial. It must be possible to identify each participant

by using this participant file. This file will contain the demographic and medical information for the participant listed below and should be as complete as possible. A non-comprehensive list of the content includes, but is not limited, to:

- Participant's full name, date of birth, sex, height, weight
- Medical history and concomitant diseases
- Prior and concomitant therapies (including changes during the trial)
- Trial identification, that is, the Sponsor trial number for this clinical trial, and participant number
- Dates for entry into the trial (informed consent) and visits to the site
- Any medical examinations and clinical findings predefined in this clinical trial protocol
- All AEs
- Date that the participant left the trial including any reason for early withdrawal from the trial or study treatment

All documents containing source data must be filed, including, but not limited to, ECG recordings, and laboratory results. Such documents must bear the subject number and the date of the procedure. If possible, this information should be printed by the instrument used to perform the assessment or measurement. As necessary, medical evaluation of such records should be performed; all evaluations should be documented, signed, and dated by the Investigator.

10.3 Investigator Site File and Archiving

Upon initiation of the trial, the Investigator will be provided with an Investigator Site File containing all necessary trial documents, which will be completed throughout the trial and updated as necessary. The file must be available for review by any Sponsor representative (e.g. the trial monitor), during Sponsor audits and for inspection by Health Authorities during and after the trial and must be safely archived for at least 25 years (or longer, per local requirements or as otherwise notified by the Sponsor) after the end of the trial. The documents to be archived include the Participant Identification List and the signed participant Informed Consent Forms. If archiving of the Investigator Site File is no longer possible at the site, the Investigator must notify the Sponsor/designee.

All original participant files (medical records) must be stored at the site (hospital, research institute, or practice) for the longest possible time permitted by the applicable regulations, and/or as per ICH GCP guidelines, whichever is longer. In any case, the Investigator should ensure that no destruction of medical records is performed without the written approval of the Sponsor.

10.4 Monitoring, Quality Assurance and Inspection by Health Authorities

This trial will be monitored in accordance with the ICH GCP, the Japanese ministerial ordinance on GCP and any other applicable regulations. The designated site Monitor will perform visits to the trial site at regular intervals.

The clinical trial protocol, each step of the data capture procedure, and the handling of the data, including the final clinical trial report, will be subject to independent Quality Assurance activities. Audits may be conducted at any time during or after the trial to ensure the validity and integrity of the trial data. Representatives of the Quality Assurance unit from the Sponsor or a designated organization, as well as Health Authorities, must be permitted to access all trial documents and other materials at the site, including the Investigator Site File, the completed CRFs, all IMP and IMP accountability records, and the original medical records or files for each participant. Quality Tolerance Limits (QTLs) will be described in the Study Management Plan.

10.5 Changes to the Clinical Trial Protocol

Changes to the clinical trial protocol will be documented in writing. Substantive amendments will usually require submission to the Health Authorities and to the relevant IEC/IRB for approval or favorable opinion. In such cases, the amendment will be implemented only after approval or favorable opinion has been obtained.

Minor (non-substantial) protocol amendments, including administrative changes, will be filed by the Sponsor and at the site. They will be submitted to the relevant IEC/IRB or to Health Authorities only as requested by pertinent regulations. Any amendment that could affect the participant's agreement to participate in the trial requires additional informed consent prior to implementation following the process as described in Section 9.2.

10.6 Clinical Trial Report and Publication Policy

10.6.1 Clinical Trial Report

After completion of the trial, a clinical trial report will be written by the Sponsor in consultation with the Coordinating Investigator following the guidance in ICH Topic E3.

10.6.2 Publication

The first publication will include the results of the analysis of the primary endpoint and will include data from all trial sites that provided evaluable data. Any publications and presentations of the results (abstracts in journals or newspapers, oral presentations, etc.), either in whole or in part, by Investigators or their representatives will require review by the Sponsor before submission. The Sponsor will not suppress publication but maintains the right to delay publication in order to protect intellectual property rights.

Posting of data on Clinicaltrials.gov and EU CTR public database is planned and will occur at the latest 12 months after the last clinic visit of the final trial participant or another appropriate date to meet applicable requirements.

11 References Cited in the Text

Abramson, S. B., Attur, M. and Yazici, Y. (2006). "Prospects for disease modification in osteoarthritis." Nat Clin Pract Rheumatol 2(6): 304-312.

Altman, R., Asch, E., Bloch, D., Bole, G., Borenstein, D., Brandt, K., Christy, W., Cooke, T. D., Greenwald, R., Hochberg, M. and et al. (1986). "Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association." Arthritis Rheum 29(8): 1039-1049.

Ayearst, L., Harsanyi, Z., & Michalko, K. J. (2012). The Pain and Sleep Questionnaire three-item index (PSQ-3): a reliable and valid measure of the impact of pain on sleep in chronic nonmalignant pain of various etiologies. *Pain research & management*, *17*(4), 281–290. https://doi.org/10.1155/2012/635967

Baraf, H. S., Gold, M. S., Clark, M. B. and Altman, R. D. (2010). "Safety and efficacy of topical diclofenac sodium 1% gel in knee osteoarthritis: a randomized controlled trial." Phys Sportsmed 38(2): 19-28.

Barthel, H. R., Haselwood, D., Longley, S., 3rd, Gold, M. S. and Altman, R. D. (2009). "Randomized controlled trial of diclofenac sodium gel in knee osteoarthritis." Semin Arthritis Rheum 39(3): 203-212.

Bellamy, N. (2005). "The WOMAC Knee and Hip Osteoarthritis Indices: development, validation, globalization and influence on the development of the AUSCAN Hand Osteoarthritis Indices." Clin Exp Rheumatol 23(5 Suppl 39): S148-153.

Bellamy, N., Buchanan, W. W., Goldsmith, C. H., Campbell, J. and Stitt, L. W. (1988). "Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee." J Rheumatol 15(12): 1833-1840.

Bihlet, A. R., Byrjalsen, I., Simon, L. S., Carrara, D., Delpy, L. and Derne, C. (2020). "A novel diclofenac gel (AMZ001) applied once or twice daily in subjects with painful knee osteoarthritis: A randomized, placebo-controlled clinical trial." Semin Arthritis Rheum 50(6): 1203-1213.

Bijlsma, J. W., Berenbaum, F. and Lafeber, F. P. (2011). "Osteoarthritis: an update with relevance for clinical practice." Lancet 377(9783): 2115-2126.

Bookman, A. A., Williams, K. S. and Shainhouse, J. Z. (2004). "Effect of a topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial." CMAJ 171(4): 333-338.

Brunner, M., Dehghanyar, P., Seigfried, B., Martin, W., Menke, G. and Muller, M. (2005). "Favourable dermal penetration of diclofenac after administration to the skin using a novel spray gel formulation." Br J Clin Pharmacol 60(5): 573-577.

Carpenter, J. R., Bartlett, J., Kenward, M. G., Morris, T., Quartagno, M., & Wood, A. (2023). *Multiple imputation and its application / James R. Carpenter, Jonathan Bartlett, Tim Morris, Angela Wood, Matteo Quartagno, Michael G. Kenward.* (Second edition.). John Wiley & Sons, Inc.

Chappell, A. S., Ossanna, M. J., Liu-Seifert, H., Iyengar, S., Skljarevski, V., Li, L. C., Bennett, R. M. and Collins, H. (2009). "Duloxetine, a centrally acting analgesic, in the treatment of patients with osteoarthritis knee pain: a 13-week, randomized, placebo-controlled trial." Pain 146(3): 253-260.

Conaghan, P. G., Hunter, D. J., Cohen, S. B., Kraus, V. B., Berenbaum, F., Lieberman, J. R., Jones, D. G., Spitzer, A. I., Jevsevar, D. S., Katz, N. P., Burgess, D. J., Lufkin, J., Johnson, J. R., Bodick, N. and Investigators, F. X. P. (2018). "Effects of a Single Intra-Articular Injection of a Microsphere Formulation of Triamcinolone Acetonide on Knee Osteoarthritis Pain: A Double-Blinded, Randomized, Placebo-Controlled, Multinational Study." J Bone Joint Surg Am 100(8): 666-677.

Cordero, J. A., Alarcon, L., Escribano, E., Obach, R. and Domenech, J. (1997). "A comparative study of the transdermal penetration of a series of nonsteroidal antiinflammatory drugs." J Pharm Sci 86(4): 503-508.

Cordero, J. A., Camacho, M., Obach, R., Domenech, J. and Vila, L. (2001). "In vitro based index of topical anti-inflammatory activity to compare a series of NSAIDs." Eur J Pharm Biopharm 51(2): 135-142.

Dakin, P., DiMartino, S. J., Gao, H., Maloney, J., Kivitz, A. J., Schnitzer, T. J., Stahl, N., Yancopoulos, G. D. and Geba, G. P. (2019). "The Efficacy, Tolerability, and Joint Safety of Fasinumab in Osteoarthritis Pain: A Phase IIb/III Double-Blind, Placebo-Controlled, Randomized Clinical Trial." Arthritis Rheumatol 71(11): 1824-1834.

Dieppe, P. A. and Lohmander, L. S. (2005). "Pathogenesis and management of pain in osteoarthritis." Lancet 365(9463): 965-973.

Drummond, M. F., O'Brien, B., Stoddard, G. L. and Torrance, G. W. (1998). Methods for the Economic Evaluation of Health Care Programmes. Second edition. Oxford, U.K.: Oxford University Press, 1997.

FDA. (1999). "Guidance for Industry: Clinical Development Programs for Drugs, Devices, and Biological Products Intended for the Treatment of Osteoarthritis (OA)."

FDA. (2009). "Diclofenac 1% Gel: Prescribing Information."

FDA. (2023). "Guidance for Industry: Adjusting for Covaiates in Randomized Clinical Trials for Drugs and Biological Products."

Frenk, S. M., Porter, K. S. and Paulozzi, L. J. (2015). "Prescription opioid analgesic use among adults: United States, 1999-2012." NCHS Data Brief(189): 1-8.

ICH E6 (R2) (2017) "Guideline for good clinical practice".

ICH E9 (R1) (2020) "Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials."

Kellgren, J. H. and Lawrence, J. S. (1957). "Radiological assessment of osteo-arthrosis." Ann Rheum Dis 16(4): 494-502.

Lane, N. E., Schnitzer, T. J., Birbara, C. A., Mokhtarani, M., Shelton, D. L., Smith, M. D., & Brown, M. T. (2010). Tanezumab for the treatment of pain from osteoarthritis of the knee. The New England journal of medicine, 363(16), 1521–1531. https://doi.org/10.1056/NEJMoa0901510

McAlindon, T. E., Bannuru, R. R., Sullivan, M. C., Arden, N. K., Berenbaum, F., Bierma-Zeinstra, S. M., Hawker, G. A., Henrotin, Y., Hunter, D. J., Kawaguchi, H., Kwoh, K., Lohmander, S., Rannou, F., Roos, E. M. and Underwood, M. (2014). "OARSI guidelines for the non-surgical management of knee osteoarthritis." Osteoarthritis Cartilage 22(3): 363-388.

McAlindon, T. E., LaValley, M. P., Harvey, W. F., Price, L. L., Driban, J. B., Zhang, M. and Ward, R. J. (2017). "Effect of Intra-articular Triamcinolone vs Saline on Knee Cartilage Volume and Pain in Patients With Knee Osteoarthritis: A Randomized Clinical Trial." JAMA 317(19): 1967-1975.

Rosen, R., Marmur, E., Anderson, L., Welburn, P., & Katsamas, J. (2014). A new, objective, quantitative scale for measuring local skin responses following topical actinic keratosis therapy with ingenol mebutate. Dermatology and therapy, 4(2), 207–219.

Pham, T., van der Heijde, D., Altman, R. D., Anderson, J. J., Bellamy, N., Hochberg, M., Simon, L., Strand, V., Woodworth, T. and Dougados, M. (2004). "OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited." Osteoarthritis Cartilage 12(5): 389-399.

Schnitzer, T. J., Easton, R., Pang, S., Levinson, D. J., Pixton, G., Viktrup, L., Davignon, I., Brown, M. T., West, C. R. and Verburg, K. M. (2019). "Effect of Tanezumab on Joint Pain, Physical Function, and Patient Global Assessment of Osteoarthritis Among Patients With Osteoarthritis of the Hip or Knee: A Randomized Clinical Trial." JAMA 322(1): 37-48.

Stevens, R. M., Ervin, J., Nezzer, J., Nieves, Y., Guedes, K., Burges, R., Hanson, P. D., & Campbell, J. N. (2019). Randomized, Double-Blind, Placebo-Controlled Trial of Intraarticular Trans-Capsaicin for Pain Associated With Osteoarthritis of the Knee. Arthritis and Rheumatology, *71*(9), 1524-1533.

Thurston J. *Ionizing radiation exposure of the population of the United States*. vol NCRP No. 160. National Council on Radiation Protection and Measurements. 2009

Wang, Y., Tu, W., Kim, Y., Sinks, S., He, J., Cambon, A., Crackel, R., Hamilton, K., Kettermann, A., & Clark, J. (2023). Statistical methods for handling missing data to align with treatment policy strategy. Pharmaceutical Statistics. 22(4):650-670

Whelton, A. (2000). "Renal and related cardiovascular effects of conventional and COX-2-specific NSAIDs and non-NSAID analgesics." Am J Ther 7(2): 63-74.

Yazici, Y., McAlindon, T. E., Gibofsky, A., Lane, N. E., Lattermann, C., Skrepnik, N., Swearingen, C. J., Simsek, I., Ghandehari, H., DiFrancesco, A., Gibbs, J., Tambiah, J. R. S., & Hochberg, M. C. (2021). A Phase 2b randomized trial of lorecivivint, a novel intra-articular CLK2/DYRK1A inhibitor and Wnt pathway modulator for knee osteoarthritis. *Osteoarthritis and Cartilage*, *29*(5), 654–666.

Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Vol. 26, Clinics in Geriatric Medicine. 2010. p. 355–69.

12 Appendices

12.1 Appendix A – American College of Rheumatology (ACR) Clinical and Radiographic Criteria Knee OA

Clinical and Laboratory	Clinical and Radiographic	Clinical†
Knee pain + at least 5 of 9: • Age > 50 years • Stiffness < 30 minutes • Crepitus • Bony tenderness • Bony enlargement • No palpable warmth • ESR < 40 mm/hour • RF < 1:40 • SF OA	Knee pain + at least 1 of 3: • Age > 50 years • Stiffness < 30 minutes • Crepitus • Osteophytes	Knee pain + at least 3 of 6: • Age > 50 years • Stiffness < 30 minutes • Crepitus • Bony Tenderness • Bony enlargement • No palpable warmth
92% sensitive	91% sensitive	95% sensitive
75% specific	86% specific	69% specific

ESR = erythrocyte sedimentation rate (Westergren); RF = rheumatoid factor; SF OA = synovial fluid signs of OA (clear, viscous, or white blood cell count < 2,000/mm^3).

† Alternative for the clinical category would be 4 of 6, which is 84% sensitive and 89% specific.

12.2 Appendix B – Contraceptive Guidance and Woman of Childbearing Potential

Definitions

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile. Permanent sterilization methods include hysterectomy, bilateral salpingectomy, and bilateral oophorectomy.

Women in the following categories are not considered WOCBP

1. Premenopausal female with 1 of the following:

    - Documented hysterectomy
    - Documented bilateral salpingectomy
    - Documented bilateral oophorectomy

    Note: Documentation can come from the site personnel's review of participant's medical records, medical examination, or medical history interview.

2. Premenarchal

3. Postmenopausal female

    - Females who are postmenopausal (age-related amenorrhea ≥ 12 consecutive, or who have undergone hysterectomy or bilateral oophorectomy are exempt from pregnancy testing.
    - Females on HRT and whose menopausal status is in doubt will be required to use one of the non-hormonal highly effective contraception methods if they wish to continue their HRT during the study. Otherwise, they must discontinue HRT to allow confirmation of postmenopausal status before study enrollment.

Contraceptive Guidance

Highly Effective Contraceptive Methods That Are User Dependent
Failure rate of <1% per year when used consistently and correctly.

• Combined (estrogen- and progestogen-containing) hormonal contraception associated with inhibition of ovulation [a]
  • oral
  • intravaginal
  • transdermal

• Progestogen-only hormonal contraception associated with inhibition of ovulation [a]
  • oral
  • injectable

Highly Effective Methods That Are User Independent
• Implantable progestogen-only hormonal contraception associated with inhibition of ovulation [a] • Intrauterine device (IUD) • Intrauterine hormone-releasing system (IUS) • bilateral tubal occlusion
• Vasectomized partner (A vasectomized partner is a highly effective contraception method provided that the partner is the sole male sexual partner of the WOCBP, and the absence of sperm has been confirmed. If not, an additional highly effective method of contraception should be used.
• Sexual abstinence (Sexual abstinence is considered a highly effective method only if defined as refraining from heterosexual intercourse during the entire period of risk associated with the study drug. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study and the preferred and usual lifestyle of the participant.)
NOTES: a) Typical use failure rates may differ from those when used consistently and correctly. Use should be consistent with local regulations regarding the use of contraceptive methods for participants participating in clinical studies.

12.3 Appendix C – Laboratory Parameters to be analyzed

Hematology

- White blood cell (WBC) count
- Lymphocytes (absolute count)
- Monocytes (absolute count)
- Neutrophils (absolute count)
- Eosinophils (absolute count)
- Basophils (absolute count)
- Red blood cells
- Hematocrit
- Hemoglobin
- Platelets

Coagulation (screening assessment only)

- International Normalized Ratio (INR)

Clinical biochemistry

- Aspartate aminotransferase (AST)
- Alanine aminotransferase (ALT)
- Alkaline phosphatase (ALP)
- Total protein
- Total bilirubin
- Creatine kinase (CK)
- Creatinine
- Calcium
- Phosphorus
- Sodium
- Potassium
- Magnesium
- Blood urea nitrogen (BUN)
- Glucose
- Chloride
- Albumin
- High-sensitivity C-reactive protein (hsCRP)

Urinalysis

Pregnancy tests:

- Serum during screening; urine at other time points.

12.4 Appendix D – Signature Page – Principal Investigator

Trial Title	A Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Trial to Evaluate the Efficacy and Safety of Once Daily Diclofenac Gel AMZ001 [redacted] in the Treatment of Pain and Symptoms of Knee Osteoarthritis
EU CT Number	2024-517404-11-00
Clinical Trial Protocol Date / Version	09May2025 Version 2.0

I, the undersigned, approve the design of the clinical trial. I am responsible for the conduct of the trial at this site and affirm that I understand and will conduct the trial according to the clinical trial protocol, any approved protocol amendments, International Council of Harmonization Good Clinical Practice (Topic E6), and all applicable Health Authority requirements and national laws.

_____ _____

Signature Date of Signature

Name, academic degree:

Function / Title: Principal Investigator

Institution:

Address:

Telephone number:

Fax number:

E-mail address:

12.5 Appendix E – Signature Page – Protocol Lead –Sponsor

Trial Title:	A Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Trial to Evaluate the Efficacy and Safety of Once Daily Diclofenac Gel AMZ001 [redacted] in the Treatment of Pain and Symptoms of Knee Osteoarthritis
IND Number:	116375
EU CT Number:	2024-517404-11-00
Clinical Trial Protocol Date / Version:	09May2025 Version 2.0

Protocol Lead:

I approve the design of the clinical trial:

_____ _____

Signature Date of Signature

Name, academic degree:	[redacted]
Function / Title:	Head of R&D
Institution:	Amzell BV
Address:	Siriusdreef 31, 2132 WT Hoofddorp, The Netherlands
Telephone number:	[redacted]
E-mail address:	[redacted]

12.6 Appendix F – Signature Page- Protocol Lead- CRO

Trial Title:	A Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Trial to Evaluate the Efficacy and Safety of Once Daily Diclofenac Gel AMZ001 [redacted] in the Treatment of Pain and Symptoms of Knee Osteoarthritis
IND Number:	116375
EU CT Number:	2024-517404-11-00
Clinical Trial Protocol Date / Version:	09May2025 Version 2.0

Protocol Lead:

I approve the design of the clinical trial:

_____ _____

Signature Date of Signature

Name, academic degree:	[redacted]
Function / Title:	Chief Scientific Officer
Institution:	NBCD A/S
Address:	Telefonvej 8D, 2860 Søborg, Denmark
Telephone number:	[redacted]
E-mail address:	[redacted]

12.7 Appendix G- Signature Page – Statistical Responsible

Trial Title:	A Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Trial to Evaluate the Efficacy and Safety of Once Daily Diclofenac Gel AMZ001 [redacted] in the Treatment of Pain and Symptoms of Knee Osteoarthritis
IND Number:	116375
EU CT Number:	2024-517404-11-00
Clinical Trial Protocol Date / Version:	09May2025 Version 2.0

Statistician:

I approve the design of the clinical trial:

_____ _____

Signature Date of Signature

Name, academic degree:	[redacted]
Function / Title:	Chief Statistician
Institution:	Omicron ApS
Address:	Telefonvej 8D, 2860 Søborg, Denmark
Telephone number:	[redacted]
E-mail address:	[redacted]

12.8 Appendix H -Protocol Amendments and List of Changes

Type	Version Number	Notes
Original protocol	1.0	
Protocol version Version EU1.1	1.1	Addition of Visit 8- safety follow-up phone call visit after a follow-up 2 week period. Addition of section describing the benefits and risks for the participants that will be enrolled in the protocol. Clarification about access to RM analgesics for placebo arm. Inclusion #9 has been modified to reflect the need to use only highly effective methods of birth control. Clarification added for physical activity. Addition of drug interactions observed in similar diclofenac gels, and the drugs where the use should be with caution. Other minor Editorial changes.
Protocol version Version 2.0	2.0	Modification of the intercurrent events (ICE) , the clinical question of interest, handling strategy for missing data and imputation. In #6 was reworded to clarify better the timepoint where the WOMAC is evaluated and that the contralateral knee must not exceed the target knee. Excl 15 WOMAC >45 was removed Excl 17 BMI >45kg/m^2 was removed. Excl 21 reworded to clarify previous conflicting wording regarding use of opioid medication. Allowed dose of Aspirin has been adjusted to reflect low dose cardiovascular prophylaxis up to 100mg/day. For participants who discontinue the study, it has been added that the site should take all possible actions to obtain the reason for discontinuation. Removal of use of prohibited concomitant medication or therapy as reason for IMP discontinuation. It has been added that at each site visits subjects will be asked if they exceeded the maximum dose of RM any day since the last visit. Allowed dose of Aspirin has been adjusted to reflect low dose cardiovascular prophylaxis up to 100mg/day. Mention that use of prohibited medication will be assessed and classified as low or high expected to impact the study endpoints. Addition of secondary endpoints for WOMAC 30% reduction, 50% reduction and use of rescue medication was described as three different secondary endpoints, the days to first use of RM, average daily use and weekly proportion. Addition of Appendix H Other minor Editorial changes..