UNIVERSITY OF CALGARY
CUMMING SCHOOL OF MEDICINE

Alberta Health Services

McCAIG INSTITUTE FOR BONE AND JOINT HEALTH

The PrEvention of post-traumatic contractuRes with Ketotifen 2 (PERK 2) Trial

Statistical Analysis Plan

Version 1.0

December 7th, 2024

NCT03582176

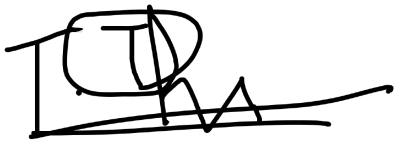
Approvals

Signature:  Date: February 18, 2025

Principal Investigator
Kevin A. Hildebrand, MD, FRCSC, FACS, CCPE
Cumming School of Medicine,
University of Calgary, Alberta

Signature:  Date: February 18, 2025

Research Coordinator
Gerardo Duque, M.A., CCRP
Cumming School of Medicine,
University of Calgary

Signature:  Date:

Lead Biostatistician
Tolulope Sajobi, PhD PStat®
Cumming School of Medicine,
University of Calgary

Signature:  Date: Feb 13, 2024

Data Manager
Ayoola Ademola, PhD
Cancer Advanced Analytics,
Alberta Health Services

Version History

Version Number	Version Date	Summary of Revisions Made
1.0	07 December 2024	

Table of Contents

1. Introduction

This document describes the statistical analysis plan for the PrEvention of post-traumatic contractuRes with Ketotifen 2 (PERK 2) trial. Joint contractures induced by trauma, arthritis, or reconstructive surgery may cause decreased elbow motion. Reduced elbow motion impairs the ability to perform daily activities such as dressing, writing, eating, and gripping objects. Also, reduced elbow motion could lower patients' overall quality of life[1-3]. Ketotifen Fumarate, a mast cell stabilizer inhibiting profibrotic growth, has been shown to minimize contracture severity in a preclinical rabbit model[4, 5].

The PrEvention of post-traumatic contractuRes with Ketotifen I (PERK-I) trial, a Phase II single-center randomized clinical trial (RCT) that compared oral Ketotifen Fumarate to a lactose placebo showed that Ketotifen Fumarate was neutral in terms of improving elbow range of motion (ROM)[6]. However, a subgroup of the trial, i.e., those who underwent surgery, indicated that Ketotifen Fumarate improved their range of motion. Therefore, the PrEvention of post-traumatic contractuRes with Ketotifen 2 (PERK-2) trial will assess the effectiveness of Ketotifen Fumarate in increasing the range of motion among surgical patients.

2. Study Objectives

The PERK-2 study aims to demonstrate that Ketotifen Fumarate (2 mg or 5 mg) administered within 10 days of injury is more effective than a placebo in reducing post-traumatic elbow joint contractures in adult participants with elbow fractures and/or dislocations. The efficacy of Ketotifen Fumarate over placebo will be assessed by comparing the difference in elbow range

of motion (ROM) across groups at different time points. Other secondary efficacy objectives will examine whether Ketotifen Fumarate can improve a patient's quality of life as evaluated by Patient-Reported Outcome Measures (PROM). Lastly, the study will assess the safety of Ketotifen Fumarate through the reporting of adverse events (AE) and serious adverse events (SAE), fracture healing, and heterotopic ossification (HO) formation.

3. Study design

The PERK-2 trial is a phase III randomized controlled multicentre trial with three parallel groups. Eligible participants were assigned to receive Ketotifen Fumarate 2 mg, Ketotifen Fumarate 5 mg, or a lactose placebo in a 1:1:1 ratio. Eligible participants will be recruited at fifteen centers across North America.

3.1. Treatment Allocation/Randomization

Randomization will be centralized, secured, and concealed via a computer-generated randomization scheme to prevent allocation bias. The randomization criteria will be reviewed and approved by an independent Biostatistician. Screening, randomization, and enrolment will be organized through the Epidemiology Coordinating and Research Centre (EPICORE) at the University of Alberta. Eligible participants will be assigned to Ketotifen Fumarate 2 mg, Ketotifen Fumarate 5 mg, or lactose placebo in a 1:1:1 ratio. A computer generated randomization scheme stratified by site with block sizes of 3:6:9 will be used to assign eligible participants. The Investigators will have access to the randomizer (on REDCap) over the Internet via a desktop computer or a web-enabled smartphone.

3.2. Blinding/masking

The participants, investigators, and assessors will be blinded to the randomization groups. Blinding will be accomplished by over-encapsulating each medication, making Ketotifen Fumarate and placebo capsules appear indistinguishable. TEVA Canada provided the treatment (ketotifen). Also, Bay Area Research Logistics (BARL) Hamilton, ON Canada, a company that specializes in clinical trial medication packaging and distribution, did the over encapsulation by placing the appropriate number of ketotifen tablets or lactose placebo with the methyl cellulose so that the investigational product in each treatment arm looked identical. However, unblinding might happen whenever serious suspected adverse reactions.

3.3. Concomitant Care

Participants will receive the following care while taking trial medication.

a) Surgery: The participant and surgeon will determine the type of operation. This management is not randomly assigned and is determined by injury characteristics. The operation can be performed before or after randomization; randomization must occur within 10 days of injury.

b) Medications: Analgesia can be provided by acetaminophen, opioids, and/or nonsteroidal anti-inflammatory drugs (NSAID) as required. The NSAIDs may have a confounding effect on the ROM measures by preventing HO following elbow injuries. Heterotopic ossification (HO) refers to bone formation in soft tissue. NSAIDs may be used in this trial as an analgesic at the discretion of the participant and surgeon because of the known analgesic benefits for

musculoskeletal injury and the ubiquitous accessibility to NSAIDs with non-prescription preparations.

c) Rehabilitation/Splints: If possible, at the site, a physiotherapist will direct a standardized home exercise program (HEP) for all groups. The HEP will consist of active ROM exercises for elbow extension-flexion arc and forearm pronation-supination arc performed 3x/day with 20 repetitions. Visits to the physiotherapist every 2 weeks will monitor progress and reinforce adherence to home therapy. Stretching splints will be used when physiotherapy alone is insufficient, which occurs at least 12 weeks post-randomization.

4. Sample Size Determination

The sample size was estimated using the Analysis of Covariance (ANCOVA) model using the mean difference of injured and contralateral elbow extension-flexion arc ROM at Week 12. The sample size was calculated based on a mean difference of 15^{o} in the extension-flexion range of motion and a standard deviation of 45^{o} between participants who received 2 mg or 5 mg Ketotifen Fumarate and those who received a placebo, using 95% power and a two-sided test at the 0.05 α-level. The sample size was recalculated due to COVID-19 (which temporarily impacted enrolment) and budgetary constraints. A total sample size of 381 participants (127 participants per arm) was determined to be adequate based on a mean change of 10^{o} extension-flexion arc ROM from baseline to week 12 and a standard deviation of 25^{o}, assuming an overall dropout rate of 11%. The standard deviation used was obtained from PERK-1 trial data. Following permission from Health Canada, the sample size was increased to 395.

5. Outcome measures

The primary endpoint examines whether at least one of the treatment arms (Ketotifen Fumarate 2 mg or 5 mg) is superior to placebo by comparing the mean difference in injured and contralateral elbow flexion-extension arc ROM at Week 12 (Table 1). The secondary endpoints will compare the mean change for multiple range of motion elements from Week 2 to other time points (Week 6, Week 12, Week 24, and Week 52) to establish Ketotifen Fumarate's efficacy over placebo. Other secondary endpoints will compare the mean difference in patients' reported outcome measures (PROMs) at multiple time points. Also, we will combined the two Ketotifen groups and compare it with placebo for the secondary endpoints. The safety assessment will compare the proportion of patients who experienced adverse effects (AEs), serious adverse effects (SAEs), suspected unexpected serious adverse reaction (SUSAR), non-union and heterotopic ossification, and reoperation in each arm. Table 1 presents a summary of the study's objectives.

Table 1: Objectives and Endpoints

Study Objectives	Endpoints
Primary	
Improvement in elbow range of motion.	The mean difference injured and contralateral elbow flexion-extension arc range of motion at Week 12.
Secondary	
1. Improvement in elbow range of motion (injured vs. contralateral elbow) at Week 12	The mean difference injured and contralateral elbow at Week 12 for the following range of motion elements: i. Flexion ii. Pronation iii. Extension iv. Supination
2. Improvement in elbow range of motion from Week 2 to other timepoints	The mean changes from Week 2 to another timepoint (Week 6, Week 12, Week 24, and Week 52) for the following range of motion elements of the injured elbow: i. Flexion-extension arc ii. Pronation- supination arc iii. Flexion iv. Pronation v. Extension vi. Supination
3. Improvement in health-related quality of life at Week 12	The mean difference for the following health-related quality of life measures at Week 12 i. The Disability of the Arm, Shoulder, and Hand (DASH) ii. The Pain Catastrophizing Scale (PCS)) iii. The Oxford Elbow Score (OES)
4. Improvement in health-related quality of life from Week 2 to other timepoints	Mean changes from Week 2 to another timepoint (Week 6, Week 12, Week 24, and Week 52) for the following health-related quality of life measures:

Study Objectives	Endpoints
	i. The Disability of the Arm, Shoulder, and Hand (DASH) ii. The Pain Catastrophizing Scale (PCS)) iii. The Oxford Elbow Score (OES)
5. Safety	▪ The proportion of SUSAR up till week 12 ▪ The proportion of fatal or life-threatening SAE up till week 12 ▪ The proportion of other SAE up till week 12 ▪ The proportion of composite reoperation for all elbow-related causes at week 12 ▪ The proportion of participants requiring reoperation for contracture at week 12 ▪ The proportion of SUSAR up till week 52 ▪ The proportion of fatal or life-threatening SAE up till week 52 ▪ The proportion of other SAE up till week 52 ▪ The proportion of composite reoperation for all elbow-related causes at week 52 ▪ The proportion of participants requiring reoperation for contracture at week 52

6. Analysis Set

a) **Intent-to-Treat (ITT) Population:** The primary endpoint analysis will be conducted on the ITT population, which includes all study participants and the treatment group to which they were assigned (i.e., the intended treatment). An ITT analysis will also be conducted for the secondary endpoints.

b) **Per Protocol (PP) Population:** The PP population includes all participants randomized and treated, with no major protocol deviations, including not meeting inclusion/exclusion criteria, not receiving the planned dose, receiving the incorrect study drug, or not obtaining consent (i.e., as treated). The primary analysis will be repeated on the PP population.

c) **Safety Population:** The safety population comprises all participants receiving any amount of study drug. Participants will be analyzed based on the intervention they receive.

7. Statistical Analysis

7.1. Descriptive Analysis

Continuous endpoints will be summarized using means, standard deviation, medians, and interquartile ranges. The differences among groups will be compared using Kruskal-Wallis's test. Categorical endpoints will be summarized using frequency and proportion. Chi-square tests will be used to compare categorical endpoints. Point estimates with two-sided 95% confidence intervals will be produced for group differences. These endpoints, baseline covariate, and stratification factors used during randomization will be summarized by treatment intervention

for the ITT population. Similar analyses will be conducted for the PP population and safety population.

7.2. Primary, secondary, and safety analyses

The linear regression model will be used to model the mean differences in the range of motion as a function of the intervention received. The unadjusted model will use a univariate linear regression model to assess whether the participants' means of range of motion received Ketotifen Fumarate (2 mg or 5 mg) differed from those who received lactose placebo. A multivariate regression model will be used to adjust for the effect of baseline covariates, i.e., sex, age, side of injury is the same as the dominant hand, type of elbow injury, concurrent injury and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. Table 2 summarizes the statistical analyses.

Table 2: Statistical Analysis of Primary, Secondary, and Safety Endpoints

Study Endpoint	Statistical Model	Estimate of treatment effects
Primary		
The mean difference between injured and contralateral elbow flexion-extension at Week 12	*Linear regression* This model will be used to obtain the unadjusted mean difference between the injured and contralateral elbow flexion-extension at Week 12.	The unadjusted mean difference and the corresponding 95% confidence intervals
Secondary		
The mean difference between injured and contralateral flexion at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference between injured and contralateral flexion at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference between injured and contralateral flexion at Week 12. The baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean difference between injured and contralateral elbow extension at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference between injured and contralateral extension at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference between injured and contralateral extension at Week 12. The baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis
The mean difference between injured and contralateral elbow pronation at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference between injured and contralateral pronation at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference between injured and contralateral pronation at Week 12. The baseline covariates sex, age, side of	The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis

Study Endpoint	Statistical Model	Estimate of treatment effects
	injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	
The mean difference between injured and contralateral elbow supination at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference between injured and contralateral supination at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference between injured and contralateral supination at Week 12. The baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in flexion-extension arc of the injuired elbow from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in flexion-extension arc between Week 2 and Week 52 among the treatment groups after adjusting for flexion-extension arc at Week 2 (baseline value). The model will have ▪ Dependent variable: The flexion-extension arc at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: Flexion-extension arc at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the flexion-extension arc recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in flexion of the injuired elbow from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in flexion arc between Week 2 and Week 52 among the treatment groups after adjusting for flexion arc at Week 2 (baseline value). The model will have ▪ Dependent variable: The flexion arc at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: Flexion arc at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the flexion arc recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in extension arc of the injuired elbow from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in extension arc between Week 2 and Week 52 among the treatment groups after adjusting for extension arc at Week 2 (baseline value). The model will have ▪ Dependent variable: The extension at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: Extension arc at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the extension arc recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in pronation of the injuired elbow from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in pronation between Week 2 and Week 52 among the treatment groups after adjusting for pronation arc at Week 2 (baseline value). The model will have ▪ Dependent variable: The pronation at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: pronation at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the pronation recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in supination of the injuired elbow from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in supination between Week 2 and Week 52 among the treatment groups after adjusting for supination at Week 2 (baseline value). The model will have ▪ Dependent variable: The supination at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: Supination arc at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the flexion-extension arc recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean difference in DASH score at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference in DASH score at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference in DASH score at Week 12. The baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean difference in PCS score at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference in PCS score at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference in PCS score at Week 12. The baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean difference in OES score at Week 12	The data will be analyzed with or without adjusting for specific baseline covariates. The statistical models are described below: *a) Linear regression model* This model will be used to obtain the unadjusted mean difference in OES score at Week 12. *b) The mixed-effects linear regression model* This model will be used to obtain the adjusted mean difference in OES score at Week 12. The baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ The mean difference and its corresponding 95% confidence intervals for both the unadjusted and adjusted analysis

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in DASH score from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in DASH score between Week 2 and Week 52 among the treatment groups after adjusting for DASH score at Week 2 (baseline value). The model will have ▪ Dependent variable: The DASH score at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: DASH score at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the DASH score recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in PCS score from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in PCS score between Week 2 and Week 52 among the treatment groups after adjusting for PCS score at Week 2 (baseline value). The model will have ▪ Dependent variable: The PCS score at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: PCS score at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the PCS score recorded at all time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
The mean change in OES score from Week 2 to Week 52	The data will be analyzed using a) *Analysis of Covariance (ANCOVA)* This model will be used to assess the mean difference in OES score between Week 2 and Week 52 among the treatment groups after adjusting for OES score at Week 2 (baseline value). The model will have ▪ Dependent variable: The OES score at Week 52 ▪ Independent variable: Treatment group ▪ Covariate: OES score at Week 2, sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. b) *Mixed-effects linear regression model* This model will be used to account for correlation among observations and estimate the effect of covariates on repeated outcomes. It will use the OES score recorded at the time points. Also, the baseline covariates sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site as a random variable. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna.	▪ ANCOVA: The p-value. ▪ Mixed-effect linear regression model: The mean difference and its corresponding 95% confidence intervals

Study Endpoint	Statistical Model	Estimate of treatment effects
Safety		
Suspected unexpected Serious Adverse Reaction (SUSAR) up to Week 12	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will be used to compare the proportion of SUSAR up to Week 12 in the treatment groups. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	Odds ratio and p-value for both the omnibus and pairwise differences models.
Fatal or life-threatening SAE up to Week 12	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* • The test will compare the proportion of Fatal or life-threatening SAEs in the treatment groups up to Week 12.	Odds ratio and p-value for both the omnibus and pairwise differences models.

Study Endpoint	Statistical Model	Estimate of treatment effects
	▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	
Other Serious adverse effect (SAE) up to Week 12	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will compare the proportion of other SAE in the treatment groups up to Week 12. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	Odds ratio and p-value for both the omnibus and pairwise differences models.
Composite reoperation for all elbow-related causes at Week 12	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test*	Odds ratio and p-value for both the omnibus and pairwise differences models.

Study Endpoint	Statistical Model	Estimate of treatment effects
	▪ The test will compare the proportion of composite reoperation for all elbow-related causes in the treatment groups at Week 12. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	
Participants requiring reoperation for contracture at Week 12	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will compare the proportion of participants requiring reoperation for contracture at Week 12 in the treatment groups. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	Odds ratio and p-value for both the omnibus and pairwise differences models.

Study Endpoint	Statistical Model	Estimate of treatment effects
Suspected unexpected Serious Adverse Reaction (SUSAR) up to Week 52	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will compare the proportion of SUSAR up to Week 52 in the treatment groups. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	Odds ratio and p-value for both the omnibus and pairwise differences models.
Fatal or life-threatening SAE up to Week 52	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will compare the fatal or life-threatening SAE proportion in the treatment groups up to Week 52. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ.	Odds ratio and p-value for both the omnibus and pairwise differences models.

Study Endpoint	Statistical Model	Estimate of treatment effects
	▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	
Other Serious adverse effect (SAE) up to Week 52	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test:* ▪ The test will compare the proportion of SAE up to Week 52 in the treatment groups. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	Odds ratio and p-value for both the omnibus and pairwise differences models.
Composite reoperation for all elbow-related causes at Week 52	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will compare the proportion of composite reoperation for all elbow-related causes in the treatment groups at Week 52.	Odds ratio and p-value for both the omnibus and pairwise differences models.

Study Endpoint	Statistical Model	Estimate of treatment effects
	▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	
Participants requiring reoperation for contracture at Week 52	The analyses will be unadjusted to avoid estimates with wide confidence intervals since we expected that the event rate might be low and that this is likely no event in some treatment groups. *Fisher's exact test* ▪ The test will compare the proportion of participants requiring reoperation for contracture in the treatment groups at Week 52. ▪ If the test produces a significant p-value at $\alpha = 0.05$, a post-hoc test will be performed to identify groups that differ. ▪ The p-values will be corrected using the Benjamin-Hochberg FDR method to ensure that the family-wise Type I error remains at 0.05.	Odds ratio and p-value for both the omnibus and pairwise differences models.

7.3. Subgroup Analyses

Subgroup analyses will be performed to evaluate differential treatment on the effects of patients' clinical and sociodemographic characteristics separately for each stratum. Heterogeneity in treatment effects will be explored via subgroup analyses of prespecified prognostic variables in the ITT population. The interactions between the treatment and subgroup variables will be tested by including a multiplicative interaction term (treatment*subgroup variable) in the model. Subgroup analyses will help determine the efficacy of any prespecified subgroup. Statistical significance for each subgroup analysis will be exploratory and conducted at α = 0.05. The subgroup analyses will include:

a. Age (<65 years vs. >= 65 years)
b. Sex (male vs. female),
c. The side of injury is the same as the dominant hand (no vs. yes)
d. Race (American Indian/Alaska Native vs. Asian vs. Black/African American vs. Native Hawaiian/Other Pacific Islander vs. White or Caucasian vs. Latino/Hispanic)
e. Type of injury
    - AO/OTA classification (distal humerus vs. proximal radius and ulna)
    - Distal humerus fracture type (Extraarticular 13A vs. Partial Articular 13B vs. Complete Articular 13C)
    - Radius, proximal end segment type 2R1 (Extraarticular fracture 2R1A vs. Partial articular fracture 2R1B vs. Complete articular fracture 2R1C)
    - Ulna, proximal end segment type 2U1 (Extraarticular fracture 2U1A vs. Partial articular fracture 2U1B vs. Complete articular fracture 2U1C)

f. Concurrent injury (no vs. yes)

g. Treatment adherence: The number of medications each participant has taken at Week 6 will be used to assess their treatment adherence. The proportion of medication (60%, 70%, and 80%) taken at Week 6 will be used to define three levels of adherence. The subgroup will be adherence vs. no adherence.

7.4. Handling of missing Data

Based on previous experience with the amount of missing data in the PERK-I trial, it is expected that there will be minimal missing data. Despite the sites' best efforts, some missing data may exist, mainly due to lost-to-follow-up. Efforts will be made to keep all missing data, particularly the primary outcome, to a minimum. Missing data will be assumed to be missing completely at random (MCAR), which means that the probability of missing values in one outcome is unaffected by the observed outcome value or any other variable values. We intend to use the multiple imputation method, which is generally regarded as the least biased since it incorporates the uncertainty of the imputed value. Missing values will be imputed for the primary outcome if the missingness is more than 5% of patients enrolled using the Multivariate Imputation by Chained Equations (MICE) approach. Ten sets of imputed data will be created and analyzed using the model described in Section 7.2. Important baseline data such as sex, age, side of injury is the same as the dominant hand, type of elbow injury, and concurrent injury will be regarded as fixed variables, and site will be included in the multiple imputation model. Type of elbow injury is defined as AO/OTA classification, i.e., distal humerus vs. proximal radius and ulna. Sensitivity analysis will be performed to assess the impact of missing data on the

primary outcome. We expect analyses using data from the multiple imputation methods to yield conclusions comparable to previous findings in Section 7.2, especially when there is minimal missing data.

8. Reference List

1. Jones V. Conservative management of the post-traumatic stiff elbow: A physiotherapist's perspective. Shoulder Elbow. 2016; 8:134-141

2. Oosterwijk AM, Nieuwenhuis MK, van der Schans CP, Mouton LJ. Shoulder and elbow ange of motion for the performance of activities of daily living: A systematic review. Physiother Theory Pract. 2018; 34:505-528

3. Siemensma MF, van der Windt AE, van Es EM, Colaris JW, Eygendaal D. Management of the stiff elbow: A literature review. EFORT Open Rev. 2023; 8:351-360

4. Monument MJ, Hart DA, Befus AD, Salo PT, Zhang M, Hildebrand KA. The mast cell stabilizer ketotifen fumarate lessens contracture severity and myofibroblast hyperplasia: A study of a rabbit model of posttraumatic joint contractures. J Bone Joint Surg Am. 2010;92:1468-1477

5. Monument MJ, Hart DA, Befus AD, Salo PT, Zhang M, Hildebrand KA. The mast cell stabilizer ketotifen reduces joint capsule fibrosis in a rabbit model of post-traumatic joint contractures. Inflamm Res. 2012; 61:285-292

6. Hildebrand KA, Schneider PS, Mohtadi NGH, Ademola A, White NJ, Garven A, et al. Prevention of posttraumatic contractures with ketotifen 1 (perk 1): A randomized clinical trial. J Orthop Trauma. 2020;34: e442-e448