Supplement 1

Protocol and statistical analysis plan

Effect of surgery vs functional bracing on functional outcome among patients with closed displaced humeral shaft fractures: The FISH Randomized Clinical Trial

Lasse Rämö, Bakir O. Sumrein, Vesa Lepola, Tuomas Lähdeoja, Jonas Ranstam, Mika Paavola, Teppo Järvinen and Simo Taimela on behalf of Finnish Shaft of the Humerus (FISH) Investigators

This supplement contains the following items:

1. PROTOCOL

1.1. Original Protocol

Original protocol submitted to Clinical Trials.gov on October 30, 2012 can be accessed at:
https://clinicaltrials.gov/ct2/history/NCT01719887?V_1=View#StudyPageTop

This same protocol with relevant information is copied below:
Study Start: October 2012 (first patient recruited November 4, 2012)
First Submitted: October 28, 2012 (at clinicaltrials.gov)

Brief Summary

Humeral shaft fractures represent 1-3% of all fractures and 20% of the humeral fractures. These fractures have historically been treated mainly conservatively with good results. Recent development in fracture treatment and findings that certain fracture types are more prone to non-union and bracing-related functional problems of adjacent joints are somewhat common have caused increasing interest in treating these fractures surgically. Return to activities is also considered to be quicker among surgically treated patients.

The purpose of this study is to evaluate effectiveness and cost-effectiveness of surgical treatment of humeral shaft fractures. Patients with a unilateral humeral shaft fracture who are willing to participate in the study after informed consent are randomly assigned to two different treatment methods:

Surgical treatment with an open reduction and internal fixation with a 4,5mm locking plate.
Conservative treatment with functional bracing

The randomization is done using blocked randomization (block sizes are not known by the enrolling or assigning physician) and stratification is done according to fracture type (AO-OTA type A vs. type B/C) and radial nerve status (total/subtotal motor palsy vs. no palsy).

Standard follow-up visits at 6 weeks, 3, 6 and 12 months are arranged. Later follow-up visits are arranged at 2, 5 and 10 years for the study purpose. Patients fill evaluation forms and clinical and radiological assessments are made. The physiotherapist doing objective functional measurements is blinded to treatment method. Both study groups receive physiotherapy after the initial treatment.

Study Design

Study Type: Interventional
Interventional Study Model: Parallel Assignment
Number of Arms: 2
Masking: Single Outcomes Assessor
Allocation: Randomized
Enrollment: 100 [Anticipated]

Arms and Interventions

Active Comparator: Conservative treatment
Conservative treatment with functional brace and physiotherapy.
Device: Conservative treatment
Conservative treatment with functional brace applied after 7 days of initial treatment with prefabricated cork splint.
Physiotherapy
Physiotherapy is arranged to both groups at 3 and 9 wks.
Experimental: Operative treatment

Operative treatment with open reduction and internal fixation with 4,5mm locking compression plate. Physiotherapy at 3 and 9 wks.

Procedure: Operative treatment

Operative treatment with open reduction and internal fixation using 4,5mm locking compression plate.

Physiotherapy

Physiotherapy is arranged to both groups at 3 and 9 wks.

Outcome Measures

Primary Outcome Measures:

1. Pain at rest and in activity, Change in Numerical Rating Scale (NRS) 0-10
   at 6 wks, 3, 6, 12 mo, 2, 5, 10 years
2. Change in The Disabilities of the Arm, Shoulder and Hand Score (DASH)
   at 6 wks, 3, 6, 12 mo, 2, 5, 10 years

Secondary Outcome Measures:

3. Subjective assessment of the function of the upper extremity
   Numerical Rating Scale (NRS) 0-10 Subjective assessment of the function of the upper extremity
4. Constant Score
5. Elbow ROM
6. Health-related quality of life (15D)
7. Complications
   Incidence of re-fracture, reoperation, infection and iatrogenic radial palsy is recorded and compared between study groups.
8. Union
   Time to union, non-union, malunion Union
9. Cost-effectiveness
   Quality-adjusted life years/months measured as a change in 15D tool, pain-NRS and other outcome measures. Cost-effectiveness
10. Subjective assessment of the function of the upper extremity
    Likert Scale 1-7 Subjective assessment of the function of the upper extremity
11. Subjective assessment of the function of the elbow
    Numerical Rating Scale (NRS) 0-10 Subjective assessment of the function of the elbow

Eligibility

Inclusion Criteria:

- Over 18 years old patient who agrees to the consent to participation in this study
- Unilateral dislocated humeral shaft fracture (dislocation over thickness of the bone cortex, fracture below the level of insertion of pectoralis major muscle and 5 cm above the olecranon fossa)
- Randomization can be done within 10 days and operation within 14 days after the initial trauma
- Patient is willing to participate all follow-up visits

Exclusion Criteria:

- Bilateral humeral shaft fracture
- A significant concomitant trauma of the same upper extremity that warrants operative treatment (fracture, tendon injury, soft tissue trauma)
- Other fracture or abdominal/thoracic trauma that warrants operative treatment
- Open fracture

- Pathological fracture
- Multi-trauma patient
- Vascular injury
- Plexus injury
- Previous trauma in the same upper extremity that causes functional deficit
- Trauma or condition that warrants use of walking aid (crutches, wheelchair etc.)
- Disease that affects significantly general condition of the patient
- Significantly impaired ability to co-operate for any reason (substance abuse, mental disorder, dementia)
- Unwilling to accept both treatment methods

1.2. Final Protocol – Amended Sections Only

(the final protocol was published in its entirety in Rämö et al[1])
The final protocol submitted to Clinical Trials.gov can be accessed at:
https://clinicaltrials.gov/ct2/show/NCT01719887

Enrollment: ~~100 [Anticipated]~~ 82 [Actual]

Outcome Measures

Primary Outcome Measures:

1. ~~Pain at rest and in activity, Change in Numerical Rating Scale (NRS) 0-10 at 6 wks, 3, 6, 12 mo, 2, 5, 10 years~~
2. ~~Change in The Disabilities of the Arm, Shoulder and Hand Score (DASH) at 6 wks, 3, 6, 12 mo, 2, 5, 10 years months~~
1. The Disabilities of the Arm, Shoulder and Hand Score (DASH) at 12 months

Secondary Outcome Measures:

7. Complications
   Incidence of complications (i.e. non-union, malunion, re-fracture, reoperation, infection and iatrogenic radial palsy) is recorded and compared between study groups.
11. The Disabilities of the Arm, Shoulder and Hand Score (DASH)
    at 6 wks, 3, 6 mo, 2, 5, 10 years
12. Pain at rest and in activity, Numerical Rating Scale (NRS) 0-10
    at 6 wks, 3, 6 mo, 12 mo, 2, 5, 10 years
13. Percentage of patients with acceptable symptom state (PASS)

1.3. Summary of Amendments

Primary and secondary outcomes

- Pain at rest and activities downgraded as secondary outcomes
- DASH at 12 months specified as the single primary outcome and other time points downgraded to secondary outcomes

When we registered the trial in ClinicalTrials.gov, our primary outcome measures were the pain at rest and activities at 6 weeks, 3 months, 6 months and 12 months as well as change in DASH at 6 weeks, 3 months, 6 months and 12 months. The secondary outcomes were as listed above in the original protocol. After discussing within the study group about the complexity of having several outcome measures at different time points we first decided to downgrade other time points than 12 months to secondary outcomes (the change was sent to clinicaltrials.gov on January 23,

2013) and later on we made a decision to have only one primary outcome, DASH at 12 months, since this instrument contains also questions regarding pain at rest and at activities. The change was made to clinicaltrials.gov on August 19, 2016.

- Percentage of patients with acceptable symptom state (PASS)

We added this secondary outcome when preparing our protocol publication in the spring 2017 and it was added to clinicaltrials.gov on May 28, 2017. We felt it would add value to our list of secondary outcomes if we define PASS of DASH score in our study population and define which part of the study group has achieved this at different time points.

Enrollment

- Enrollment from 100 [anticipated] to 82 [actual]

When we first registered the study, we reported the enrollment to be 100 patients. We had done the power analysis which showed 35 patients per group and we decided to have 12,5% lost to follow-up reservation. When we sent our study protocol to the ethical board of Helsinki and Uusimaa Hospital District, we put the correct value of 80 patients to the target field. We first registered the enrollment target to 100 patients and after noticing this mistake we made the correction to clinicaltrials.gov on May 28, 2017 when we unified the registered protocol between clinicaltrials.gov and the accepted protocol paper[1]. The number of enrolled patients became 82 since the enrollment took place in two separate units and we were unable to stop the recruiting exactly at 80 patients. After noticing we had achieved the target, we stopped the enrollment on January 2018.

Be it noticed here that all the above noted amendments to the original protocol were made prior to completion of the trial and before doing any data analysis and prior revealing the allocations of the study groups.

2. STATISTICAL ANALYSIS PLAN

2.1. Original Statistical Plan

A description of our original statistical analysis plan was published[1] as follows:

The data will be analyzed using IBM SPSS Statistics V.23 or higher. The results will be reported following the Consolidated Standards of Reporting Trials statement.

The baseline characteristics of the participants will be summarized by group, reported as a mean (SD) or median (first quartile, third quartile) for continuous variables, and count (%) for categorical variables.

Primary statistical analyses will be performed using intention-to-treat basis. For the primary analysis, a mixed-effects model (MM) analysis will be performed using the data set without multiple imputation to compare the mean DASH scores. Treatment group and visits will be included as fixed factors and patient as a random factor. The model will include interactions between treatment and visit. Randomization stratification factors and baseline value will be included as covariates. The treatment effect will be quantified with an absolute difference between the groups in the DASH score with the associated 95% CI and p value at 12 months post-randomization.

The MM model will also be used to analyze secondary outcomes where applicable (pain-NRS at rest and during activities, 15D, CS). For categorical response variables, effects will be analyzed by logistic regression analysis with treatment as the fixed-factor covariate. These secondary outcomes will only be supportive, explanatory or hypothesis-generating (or both), which is why multiplicity is not considered to be a problem.

The adverse events of the study arms will be reported descriptively. If the number of events is large enough, an analysis between study arms will be performed.

All scale variables will be tested for normality with the Kolmogorov-Smirnov test. Variance of homogeneity will be tested using Levene's test. We consider a two-sided p value of 0.05 to indicate statistical significance.

We will perform secondary statistical analyses to identify potential effect-modifying and mediating factors. Potential effect-modifying factors to be tested with regression analyses are age, gender, body mass index, physical activity, smoking, level of education, fracture of dominant/non-dominant arm and position of the fracture. The absence of adverse effects and treatment attendance as intended will be analyzed as a potential effect-mediating factor.

We will also perform an on-treatment analysis if there are patients treated with a non-allocated method because patients declined the allocated treatment after the randomization, thus causing crossover in study arms. A medical reason to change treatment method, practically from conservative treatment to ORIF because of non-union or fracture threatening skin integrity in the early phase of treatment, will not be considered as a crossover. However, we will analyze such patients in a separate subgroup.

2.2. Blinded Data Interpretation Protocol

We used blinded data interpretation in analyzing the results of this trial.[2] The blinded data interpretation protocol was published in our protocol paper[1] as follows:

Before accessing the primary outcome data, the Writing Committee will record a 'Background assumptions' statement, which will contain our definition of MID of the outcome measures and a brief summary of the key statistical analysis used in the evaluation of the outcome data. The document will be signed by the members of the Writing Committee and published as an appendix to the primary publication. After this, the Writing Committee will write two interpretations of the trial results on the basis of a blinded review of the primary outcome data (treatment A compared with treatment B), with the assumption that A is the ORIF group and another assuming that A is the conservatively treated group. Decisions regarding the key analyses and presentation format for the primary publication before data analysis will also be decided in a meeting of the Writing Committee. The minutes of this meeting will be recorded as a statement of interpretation document, which will be signed by all members of the Writing Committee before the unsealing of the randomization.

2.3. Final Statistical Analysis Plan - Amendments

The statistician doing the data analysis is using Stata version 15.1 (StataCorp LLC, Texas, USA) instead of IBM SPSS Statistics. We consider this a minor technical detail which does not affect the interpretation of our results.

Instead of Kolmogorov-Smirnov test for normality and Levene's test for homogeneity, we will use other techniques, e.g., graphical evaluation.

All P values larger than 0.01 are be reported to two decimal places, and those between 0.01 and 0.001 to three decimal places; P values smaller than 0.001 are be reported as P<0.001. We made this amendment since we did not state this in our protocol paper.

Primary analysis –Amendments

The primary comparison on the effectiveness of the treatment will be performed as a between-group comparison using a mixed-model repeated-measures analysis of variance (MMRM ANOVA). In the original analysis plan we used a term 'MM model' but changed the term to 'MMRM ANOVA' as it is more widely used term. We consider this only a terminological issue not affecting the analysis.

Study group and time of assessment (baseline, 6 weeks, 3, 6 and 12 months) were included as fixed factors, patient as a random factor. The model included interactions between study group and time of assessment. Change from baseline was estimated with baseline value as covariate. An unstructured covariance structure will be assumed. If the model cannot be fitted, compound symmetry will be assumed instead. The number of degrees of freedom will be assessed using Satterthwaite's method. The MMRM model will be used to quantify the treatment effect as the absolute difference between the groups in DASH score with the associated 95% confidence interval (CI) and p-value at 12 months post-randomization.

2.4. Implementation of Analysis Plan

This SAP will be used as a work description for the statistician performing the analyses. All analyses will be performed by the same statistician and none of the investigators involved in this trial will perform any of the statistical analyses.

The implementation of the SAP will be as follows:

1. A 'data collection form' will be outlined in a collaboration between the database manager (Leena Caravitis) and principal investigator (Lasse Rämö) and senior author (Simo Taimela).
2. The database manager will code each treatment arm into 'Group A' and 'Group B', thus leaving all others blinded to group assignment during the analyses.
3. Blinded data will be delivered to the statistician according to the 'data collection form'.
4. Primary, secondary and exploratory endpoint analyses will be made blinded to group assignment.

Results will be presented to the trial Writing Committee, any uncertainties will be clarified and blinded interpretations of the primary endpoint results will be conducted prior to unblinding of data.

A detailed description of the execution of the statistical analysis can be found in our "Blinded Data Analyses Statement of Interpretation"-document (Supplementary Appendix of our submission).

Be it reiterated here that the entire statistical analysis was carried out blinded and the randomization code was broken only after the main findings/interpretation of the results were mutually agreed on (and documented) by the entire manuscript writing committee.

References

1. Rämö L, Taimela S, Lepola V, Malmivaara A, Lähdeoja T, Paavola M. Open reduction and internal fixation of humeral shaft fractures versus conservative treatment with a functional brace: a study protocol of a randomised controlled trial embedded in a cohort. *BMJ Open*. 2017;7(7):e014076. doi:10.1136/bmjopen-2016-014076.

2. Järvinen TLN, Sihvonen R, Bhandari M, et al. Blinded interpretation of study results can feasibly and effectively diminish interpretation bias. *Journal of Clinical Epidemiology*. 2014;67(7):769-772. doi:10.1016/j.jclinepi.2013.11.011.