Surgery or Cast for Injuries of the EpicoNdyle in Children's Elbows (SCIENCE) Study

Health Economics Analysis Plan (HEAP)

Version 1.0 – 20 August 2024

Based on Protocol version 7.0 – 20 June 2024

Trial registration: ISRCTN16619778

Health Economist:  date: 20Aug2024

Dr Kandiyali signs on behalf of both herself and Emeritus Professor James Mason, who played a key role in the development of this plan and approved an earlier version.

Chief Investigator:  date: 23Aug2024

CONTENTS

ADMINISTRATIVE INFORMATION

This document describes the planned analysis of economic data within the SCIENCE trial. This Health Economics Analysis Plan (HEAP) should be read in conjunction with the SCIENCE Trial Statistical Analysis Plan and Trial Protocol which provide in detail: trial design and methods, amendments, documentation, oversight, roles and responsibilities, and the statistical plan of analysis of clinical and patient outcome measures.

BACKGROUND

The management of fractures of the medial epicondyle is a controversial issue in paediatric fracture care [1]. These fractures typically occur in children around 10-12 years old [2], with or without dislocation of the elbow joint. The debate for clinicians is whether to realign and hold the bone fragments with operative fixation, or whether to allow the fragments to heal in their current position without surgery by resting the elbow in a cast.

Observational studies have demonstrated support for both operative and non-operative treatment strategies, generating uncertainty amongst surgeons. Two published systematic reviews have demonstrated disagreement in the management of this injury [2, 3]. One systematic review concluded that nonsurgical treatment offers excellent functional results equivalent to surgical treatment [3]. The other concludes that surgical fixation should be strongly considered to achieve union of the bone fragments thereby maximising elbow stability for an increasingly athletic child population [2]. To add further to the debate, a widely used 'evidence-based review' textbook has recently advocated against surgery, citing increased long-term pain and stiffness compared to non-operative treatment [4].

Much of the controversy has arisen because there have been no prospective studies evaluating the treatment of these fractures. Current literature has serious methodological limitations, especially with regard to inconsistent follow-up, no standardisation to treatment approaches, infrequent use of patient reported outcomes, and selection bias for operative fixation [4]. There has also been a lack of agreement on how to record successful outcomes; with radiographic union of the fracture fragments being the most commonly used outcome, and pain or function being infrequently recorded, although there is known to be poor correlation between radiographic union and functional outcomes [3]. The uncertainty within the literature has resulted in considerable variation in clinical practice. There is an increasing tendency toward surgery, which has been particularly driven by US literature identifying the athletic demands of children and adolescents, and the expectations of early mobilisation and return to sport [1, 5].

An audit of surgical practice amongst 30 centres in the UK was conducted, and identified 520 medial epicondyle fractures over a 3-year period. Overall 225 (43%) were treated with surgical fixation, with the remaining 295 (57%) treated non-operatively. 39 children (8%) had an incarcerated fragment, which is an absolute requirement for surgery. The decision to offer 'operative intervention' is highly dependent upon the surgeon. Surgical fixation is thought to improve the likelihood of 'bony union' of the fracture. However this is balanced against the small but definite risks from the surgery including infection, nerve damage around the elbow, broken and retained metalwork, and the risks associated with general anaesthesia. It is unclear whether bony union has any bearing on functional recovery, including return to sports.

There is a clear and pressing need to inform patients about the benefits or otherwise of operative fixation versus non-operative treatment, and the need to inform commissioners regarding the costs of the different

treatment strategies to the NHS and society. The SCIENCE study is a randomised superiority trial of operative fixation versus non-operative treatment for medial epicondyle fractures of the humerus in children.

TRIAL DESIGN

SCIENCE is a multi-centre, prospective, randomised, superiority trial using a two-arm parallel group design. Patients will be randomised in a 1:1 ratio to either receive operative fixation, or non-operative treatment. The nature of the treatment means that neither the patients and their parents/guardians, nor the treating clinician can be blinded to the treatment received; however, the treating clinical team will take no part in the follow-up assessments.

The trial includes an internal pilot (phase 1) which will confirm the expected rate of recruitment in a large-scale, multi-centre randomised controlled trial. This pilot will take place at 20 centres over 12 months. The Data Safety and Monitoring Committee (DSMC) will make a recommendation to the Trial Steering Committee (TSC) regarding trial continuation in the event that the recruitment target for the internal pilot is not met. Otherwise the trial will continue into the main phase (phase 2), and patients from the internal pilot will be included in the final analysis. The main trial phase will be recruiting from a minimum of 35 centres treating children's fractures across the UK, stratified by centre to account for centre specific effects, and by dislocation status of the elbow. Follow-up will be made electronically (sent by e-mail or text message) for the Patient Reported Outcomes at 6 weeks, 3 months, 6 months and 12 months (Primary outcome time-point).

All children aged 7-15 years old presenting at the trial centres with a medial epicondyle fracture of the humerus are potentially eligible to take part in the trial. After consent has been gained, a local research associate will collect baseline demographic data, the Patient Reported Outcomes Measurement Information System (PROMIS) Upper Extremity Score for Children Computer Adaptive Test, DASH S/PA Module, Wong Baker Faces Pain Score, and health-related quality of life using the EuroQoL EQ-5D-Y.

Event	**Date**
Grant activation	01 Oct 2018
Trial Open	20 May 2019
First DSMC meeting	15 May 2019
Expected end of recruitment	Sept 2023 as of 22/03/2023
Date expected end follow-up/start of data cleaning	Sept 2024
Expected start of final analysis:	Nov 2024 as of 12/06/2024
Expected start of long-term analysis (detailed in Appendix B of SAP)	Jul 2027 tbc
End of Grant	Feb 2030

OBJECTIVE

The health economic objective is to evaluate the cost-effectiveness of operative fixation versus non-operative treatment for displaced medial epicondyle fractures of the elbow in children using resource use and quality of

life data from baseline to 12 months follow-up. Analysis is by intention-to-treat, presenting resource use, cost and quality of life findings by trial arm. Attention will be paid to completeness of data, identifying issues and potential remedies.

ECONOMIC EVALUATION

A prospectively planned economic evaluation will be conducted from an NHS and personal social services perspective, according to the recommendations of the NICE reference case[29]. Operative fixation will be compared with immobilisation of a displaced medial epicondyle of the humerus in children. The trial is recruiting internationally, however in order to provide generalisable information to the NHS the economic analysis will be limited to UK recruiting sites and follow intention-to-treat principles. Participants are aged 7-15 years, so it is possible that the responses to the child completed version of the EQ-5D-Y may be assisted by parent/carers, especially in younger children. Resource use questions will be primarily completed or assisted by parents/carers.

RESOURCE USE AND COSTS

Healthcare resource use will be costed using most recently available published national reference costs, related to a common year (as a guide this is likely to be the most recent year for which a schedule of NHS trust costs) is available. Index hospital procedures and any sequelae procedures will be costed using the diagnosis and procedure codes applied to this procedure and then mapped to a HRG (healthcare resource group) code. The relevant HRG is HT64 (Intermediate Elbow Procedures for Trauma), specifically either HT64C if the participant is 6-18 years or HT64D if ≤5 years old. Some patients may need revision or have metal work removed requiring a further surgery; we will record OPCS codes and documentation flags (where available) in order to include these additional costs. Children without fixation will just be seen in fracture outpatient clinic, and this will be costed as a Paediatric outpatient consultation (service code) 171. Participants' health service contacts, made in connection with their elbow, will be recorded at 3, 6 and 12 months via parent/carer completed resource use questionnaire and costed using the schedule of NHS trustcosts[32]. Personal expenses, parent/carer time from work and time from school will also be recorded as part of a broader societal perspective (using the same questionnaire).

OUTCOMES

Health-related utility score measurement in children and particularly adolescents is difficult due to a lack of validated measures. While the EQ-5D is a widely used measure for calculating quality adjusted life years (QALYs), there is presently no validated UK value set for use in both children/adolescents. The youth version of the EQ-5D (the EQ-5D-Y) (**https://euroqol.org/eq-5d-instruments/eq-5d-y-about/**) categorises health on each of five domains but there have been unexpected delays in the dissemination of a UK value set. The EQ-5D-Y also includes a visual analogue scale (VAS), which asks respondents to rate their health on a thermometer-like scale from zero (worst) to 100 (best) health. Another instrument, the CHU-9D, has a value set for calculating utility, [1] but was originally developed in primary school aged children in the UK. This has more recently been validated in a population of older aged children and adolescents (11–17-year-olds) in Australia.

In the absence of a validated measure that spans primary and secondary school-aged children, we propose to assess health related quality of life using the version of EQ-5D-Y at baseline, 6 weeks, 3, 6, and 12 months. EQ-

5D-Y scores will be converted to health utility scores using the most appropriate tariff available at the time of analysis; note that the reporting of these scores, as secondary outcomes, is described in the Statistical Analysis Plan (SAP, Secondary Outcomes). In order to facilitate the cost-utility analysis, patient-level QALY estimates will be derived using the trapezoidal rule to measure the area under the curve (AUC). Since AUC estimates are predicted to correlate with baseline scores (and thus potential baseline imbalances), AUC estimates will be adjusted for baseline scores within regression analyses. A further analysis repeating the above but using the VAS scores will be performed if the UK value set for the youth version remains unavailable at the time of analysis.

ANALYSIS

The base case analysis will present the within trial incremental cost and QALY quality-adjusted life years (QALYs) gained, adjusted for (baseline EQ-5D-Y score and elbow dislocation status, with additional prognostic factors to be included pending a decision by our statistical colleagues). If overall data missingness exceeds 5% the primary analysis will include multiple imputation using the MI framework in Stata. Should this occur, a complete case analysis will also be included as a sensitivity analysis and findings compared.

Multiple imputation provides unbiased estimates of treatment effect if data are missing at random: this assumption will be explored in the data, for example by using logistic regression for missingness of costs and QALYs against baseline variables [2]. A regression model will be used to generate multiple imputed datasets (or 'draws') for individual treatment groups, where missing values are predicted. Outcome measures and costs (at each time point) will contribute as predictors and imputed variables. Each draw provides a complete dataset, which reflects the distributions and correlations between variables. Predictive mean matching drawn from the five nearest neighbours (knn=5) will be used to enhance the plausibility and robustness of imputed values, as normality may not be assumed. The imputation model will use fully conditional (MCMC) methods (multiple imputation by chained equations), which are appropriate when missing and correlated data occur in more than one variable. Each draw will be analysed independently using bivariate/mixed effects regression (see below) and the estimates obtained will be pooled to generate mean and variance estimates of costs and QALYs using Rubin's rule – a method that captures within and between variances for imputed samples [3]. To minimise the information loss of finite imputation sampling, 20 draws will be taken. The distribution of imputed and observed values will be compared visually and statistically to establish the consequences of estimation.

As there are a large number of centres participating in the trial, and the number of participants in some centres will be low, we anticipate some clustering by site. The regression approach therefore needs to consider whether such clustering is important. The regression approach will be informed by visual inspection and univariate testing. For the univariate testing, we will run separate regression equations for Cost and QALYs with and without site as a cluster variable. Should these be unimportant (as determined by significance of the random effect variable, viewed in conjunction with results on the Akaike Information criteria and Log rank tests), we will proceed to a bivariate SUR and bootstrapping to produce the CE plane. Otherwise, we will use a mixed model which explicitly models clustering in the data (individuals at level one and sites at level two).

Both approaches consider skew and the joint correlation between costs and effects while allowing for adjustment for a set of covariates, which can be explored and which improve precision [4]; however only the

mixed model will allow for cluster-specific random effects. Baseline QOL scores will be included within all models to allow for potential baseline imbalances [5].

The incremental cost-effectiveness ratio (ICER) will be estimated as the difference between treatments in average total costs divided by the difference in average total QALYs. Value-for-money is determined by comparing the ICER with a threshold value, typically the NICE threshold for British studies, of £20k-30k/QALY [6]. This represents the willingness to pay for an additional QALY, and lower values than the threshold could be considered cost-effective for use in the NHS. Base case assumptions will be explored using a range of supportive sensitivity analyses, providing an assessment of the robustness of findings.

The net monetary benefit (NMB) of changing treatment will be reported as a recalculation of the ICER at a range of thresholds of willingness to pay for an additional QALY. The NMB succinctly describes the resource gain (or loss) when investing in a new treatment when resources can be used elsewhere at (upto) the same threshold. NMB estimates will be used to generate cost-effectiveness acceptability curves (CEACs). The CEAC compares the likelihood that treatments are cost-effective as the willingness to pay threshold varies [13].

The expected value of perfect information (EVPI) is the upper limit of the value to a healthcare system of further research to eliminate uncertainty [7]. Findings from cost-effectiveness analyses remain uncertain because of the imperfect information they use. If a wrong adoption decision (to make a treatment available) is made this will bring with it costs in terms of health benefit forgone: the NMB framework allows this expected cost of uncertainty to be determined and guide whether further research should be conducted to eliminate uncertainty.

Analyses and modelling will be undertaken in Stata 17 SE (or later release if available). Reporting will follow the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement [8]

Should costs and quality-of-life not converge within 12 months, more extensive economic modelling using decision-analytic methods may be considered to extend the target population, time horizon and decision context, drawing on best available information from the literature and stakeholder consultations to supplement the trial data. Parameter uncertainty in the decision-analytic model will be explored using probabilistic sensitivity analysis. If longer term decision modelling is to be undertaken, then costs and outcomes will be discounted at 3.5% after the first year of randomisation in line with NICE reference case [6]. A decision as to the necessity of building a decision analytic model and its specification will be made following discussion between the health economists and the trial team following preliminary analysis of the data. This will be informed by considerations such as the conclusiveness and direction of within trial results. For example, if the control dominates the intervention and extrapolation would only increase the strength of this result then there is little need to extrapolate further as the intervention should be rejected.

Findings will be visualised in the cost-effectiveness plane, as cost-effectiveness acceptability curves, net monetary benefit and value of information analysis.

If overall data missingness exceeds 5% the primary analysis will include multiple imputation using the MI framework in Stata. Mechanisms of missingness of data will be explored and multiple imputation methods will be applied to impute missing data, following best practice[35,36,37]. Imputation sets will be used in subsequent analysis of costs and QALYs to generate incremental cost per QALY estimates and confidence intervals. If the level of missingness is low, then a complete case analysis will be conducted without imputation.

Findings will be analysed and visualised in the cost-effectiveness plane, as cost-effectiveness acceptability curves, net monetary benefit and value of information analysis. A within-trial analysis will use the first 12 months of data, to correspond to the primary analysis. If incremental costs and benefits are non-convergent within the trial follow-up then extrapolated modelling will be considered.

DUMMY TABLES

In accordance with the analysis plan, planned tables and figures are described below.

1.1 Completeness of data by follow-up visit

	Control[1] n (%, N)	Intervention[2] n (%, N)	Total n (%, N)
Health status[3]			
EQ-5D-Y Baseline			
EQ-5D 6 weeks			
EQ-5D 3 months			
EQ-5D 6 months			
EQ-5D 12 months			
EQ-5D All visits			
Resource use[4]			
Inpatient			
Outpatient			
Community healthcare			
Personal expenses			
Parent/carer time off work			
School absence			

1 Single advice session only
2 Single advice session and at least one physiotherapy session
3 EQ-5D-Y descriptive system complete for all 5 domains
4 Range shown, lowest to highest completion at measurement points

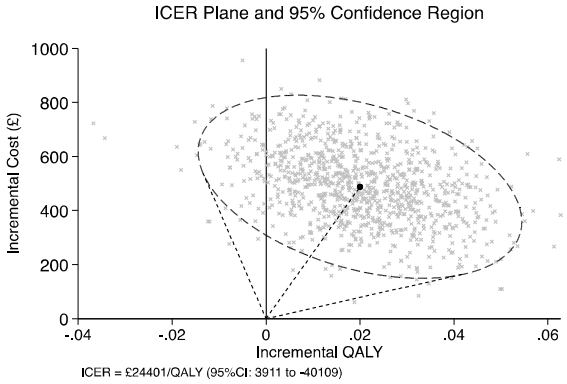
1.2 Health status, resource use and cost

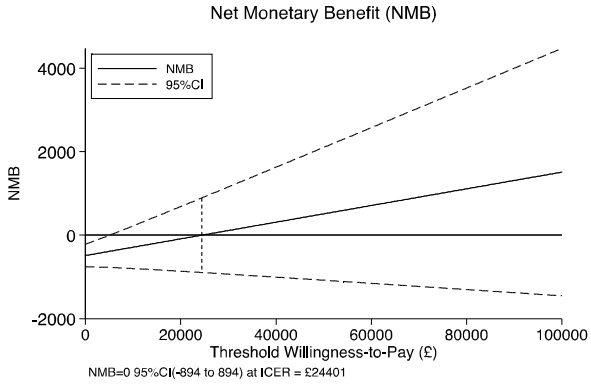
	Control		Intervention	
	mean	(SD)	mean	(SD)
Health status[1]				
EQ-5D Baseline				
EQ-5D 6 weeks				
EQ-5D 3 months				
EQ-5D 6 months				
EQ-5D 12 months				
EQ-5D AUC				
Resource use (all visits)				
Inpatient days				
Outpatient visits (rehabilitation)				
Community healthcare				
Personal expenses				
Parent/ Carer work absence (days)				
Cost				
Inpatient days				
Outpatient visits (rehabilitation)				
Community healthcare				
Personal expenses				
Parent/ Carer work absence (days)				

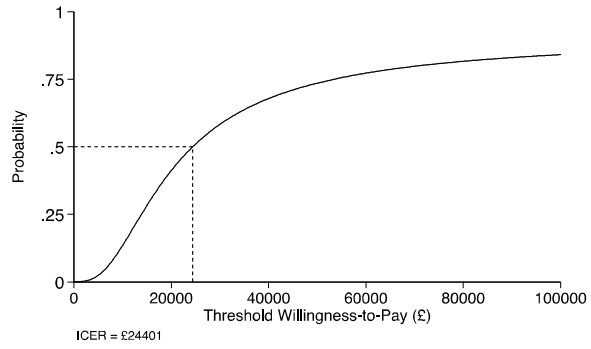
1 EQ-5D-Y else EQ-5D-3L index score, pending availability of EQ-5D-Y utility weights for the UK.

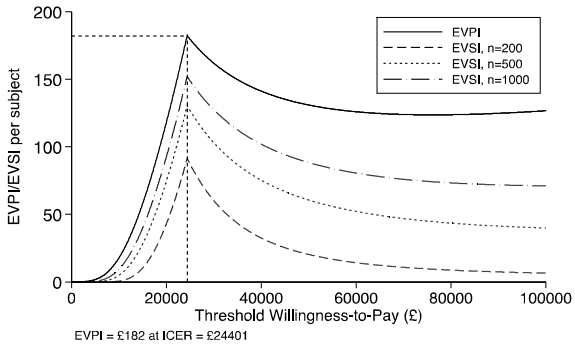
1.3 Cost-effectiveness in terms of total and incremental QALYs and net monetary benefits

1.4 Example figures









REFERENCES

1. Stevens, K., *Valuation of the Child Health Utility 9D Index.* Pharmacoeconomics, 2012. **30**(8): p. 729-47.
2. Faria, R., et al., *A Guide to Handling Missing Data in Cost-Effectiveness Analysis Conducted Within Randomised Controlled Trials.* PharmacoEconomics, 2014. **32**(12): p. 1157-1170.
3. White, I.R., P. Royston, and A.M. Wood, *Multiple imputation using chained equations: Issues and guidance for practice.* Statistics in Medicine, 2011. **30**(4): p. 377-399.
4. Willan, A.R., A.H. Briggs, and J.S. Hoch, *Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data.* Health economics, 2004. **13**(5): p. 461-475.
5. Manca, A., N. Hawkins, and M.J. Sculpher, *Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility.* Health economics, 2005. **14**(5): p. 487-496.
6. NICE, *Guide to the methods of technology appraisal 2013.* 2013.
7. Claxton, K., M. Sculpher, and M. Drummond, *A rational framework for decision making by the National Institute for Clinical Excellence (NICE).* The Lancet, 2002. **360**(9334): p. 711-715.
8. Husereau, D., et al., *Consolidated health economic evaluation reporting standards (CHEERS) statement.* Cost Effectiveness and Resource Allocation, 2013. **11**(1): p. 6.