Statistical Analysis Plan

Chief Investigator: Dawn Carnes

Senior Trial Statistician: Paul Vaucher

Data Manager and Statistician: Kevin Brownhill

Version and Date V1.4 17.08.22

History of versions

Version (date)	Changes from prior version	Origin of demand
V0.1 (09/06/22)	Changes from the initial statistical plan included in the study plan: • Use regression analysis at a daily entry level instead of overall mean crying time. • Primary analysis is to adjust for pre-defined potential effect modifiers. • Even if diaries have missing data, they are to be considered for analysis if at least 7/14 days are filled in. • Specification on assumptions to be tested. • Worse and best-case scenarios to account for missing diaries.	The TSC asked the DMC to consider improving the statistical plan and use an analysis accounting for reported individual daily crying times.
V0.2 (24/06/22)	• Improvements in describing datasets • Agreed on co-factors that are to be entered in main regression analysis: baseline crying time, infant age, days within trial and expected outcomes. • Sensitivity analysis added to account for potential performance differences between groups. • Dummy tables added.	Decisions made by the statistical team to improve the plan.
V0.3 (07/07/22)	• Improved details on blinding of statistician. • Improved description of primary outcome and way to calculate daily crying time from diaries. • Detailed description of way to test assumptions.	Final plan written by the statistical team for validation by the TMG.
V1.0 (17/07/22)	• Improving description of regression analysis.	Validated version by the TMG.
V1.2 (28/07/22)	• Improved description for datasets and metadata files.	Statistical team
V1.3 (04/08/22)	• Improving description of way to test assumptions.	Statistical team
V1.4 (17/08/22)	• Description of how to manage missing data within diaries. • Description of how to impute values for diaries missing crying time in minutes during crying hours. • Sensitivity analysis of results without missing diary imputation.	Validation by TSC.

1. Description of study

Background:

Infants who excessively cry, are distressed and unsettled can have a marked impact on family life. Around 1 in 6 families are affected, it is associated with maternal depression, anxiety and loss of parenting confidence. These infant behaviors are usually self-limiting (subsiding around 12 weeks after birth) but during this difficult period many parents look for additional support. There is limited research and therefore much debate about the effectiveness of manual therapy and osteopathic care for these infants.

Aim:

To evaluate the effectiveness and healthcare cost of osteopathic care for excessively crying, unsettled and distressed infants (< 10 weeks old).
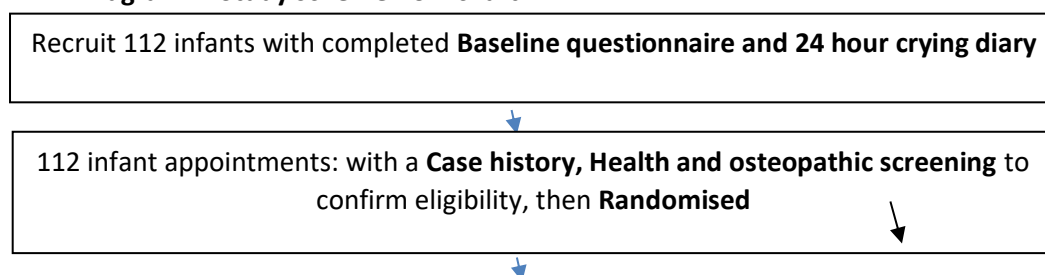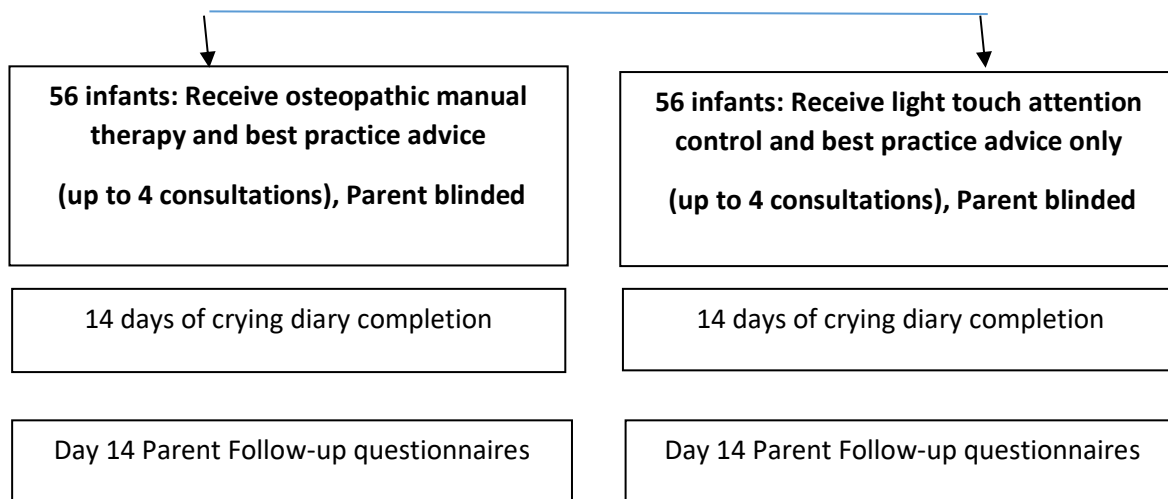
Study design:

A two-arm pragmatic randomised controlled trial to assess the effectiveness of osteopathic manual therapy on infant crying time. We will need to randomize 112 parent(s)/carers and their infants to either: i) Osteopathic manual therapy with best usual care advice and support or, ii) Light non-specific touch attention control (equal time spent with infant) with best usual care advice and support. Parents will be blinded to whether their infant receives the manual therapy care or not. The cost of the delivery of the intervention will be determined and compared with data about the cost of other care.

Population: Healthy infants under 10 weeks old, reported by their parent(s)/carers as excessively crying, unsettled and distressed and difficult to console using the Rome IV criteria (> 3 hours of crying per day, for 3 days for 1 week or more). Infants with diagnosed health conditions for which they are receiving medical treatment, or who would not normally be suitable for osteopathic treatment will be excluded from the study.

Outcomes: The primary outcome is reduced average daily infant crying time over 14 days, collected via parent reported diaries. Secondary outcomes are: i) Parental self-efficacy ii) Parent perceived global improvement iii) Satisfaction and experience with treatment and iv) Adverse events.

Diagram 1. Study scheme flow chart

Recruit 112 infants with completed **Baseline questionnaire and 24 hour crying diary**

112 infant appointments: with a **Case history, Health and osteopathic screening** to confirm eligibility, then **Randomised**

56 infants: Receive osteopathic manual therapy and best practice advice (up to 4 consultations), Parent blinded	56 infants: Receive light touch attention control and best practice advice only (up to 4 consultations), Parent blinded
14 days of crying diary completion	14 days of crying diary completion
Day 14 Parent Follow-up questionnaires	Day 14 Parent Follow-up questionnaires

2. Source documents

Crying diary (participant self-report)

Corresponds to day-to-day reporting entry for crying time reported in minutes. Parents are asked to record their infants crying time to the nearest 5 minutes over 24 hours for 14 days. Each day starts at 6AM and finished the next day at the same time. Data is entered for each time slot separately. Each daily total is then compared to an independent day to day data entry (pseudo-double data entry). If daily sums do not match, the data is verified again, and errors are corrected. Data cleaning includes removing entries that would account for more than 24h of daily crying, reporting crying days from the day the first treatment was provided, transforming data when necessary (ex. 5 minutes time slots reported, graphical representation of crying time), replacing negative values or non-explicit string entries by missing values. Values are entered as minutes. Missing data is entered as blanks.

If some parents only report the daily total, these are to be recorded on a separate file (diary2.csv).

CastorEDC entries

Data is exported from the CastorEDC entry system from three separate forms:

1. *Baseline questionnaire (participant self report ; 24 questions)*
2. *Follow up questionnaire at 14 days post randomisation (participant self report, 27 questions)*
3. *Consultation information (osteopath reporting after each encounter)*

Data cleaning will be performed by the trial statistician consultation with the data manager

3. Datasets for data analysis

Six separate datasets are to be prepared.

1. Merged data from Castor EDC (one line per participant)

Dataset name: castor.csv

Source: CastrorEDC Baseline questionnaire (bq), follow-up questionnaire (fq) and consultation information (oq)

Format: csv

Details for data format

Variable	Source	Description	Type	Levels
id	bq, fq, oq	UID	Numerical	Nominal
pract	bq, oq	Practitioner	Numerical	Nominal
date_b	bq1.1	Date of entry	Date	
date_end	fq1.1	Date of last visit	Date	
parent_rel1	bq1.2	Relation to child for person answering 1st questionnaire	Nominal	1= mother, 2= father, 3=legal guardian, 4= other
parent_rel2	fq1.2	Relation to child for person answering 2nd questionnaire	Nominal	1= mother, 2= father, 3=legal guardian, 4= other
age	bq1.3	Infant age (days)	Continuous	Days
bw	oq3.1	Birth weight (Kg)	Continuous	Kg
weight_b	oq3.2	Weight current (Kg)	Continuous	Kg
weight_end	fq1.3	Weight at end of follow-up	Continuous	Kg
ct_b	bq1.4/1.5	Reported crying time prior 24h (min)	Continuous	Minutes
sibling	bq1.6	Siblings	Ordinal	0=0 , 1=1, 2=2, 3=3, 4=4, 5=5 or more
parent_age	bq1.7	Age of parent (years) completing the questionnaire	Ordinal	1= 18-20 years, 2=21-25 years, 3=26-30 years, 4=31-35 years, 5=36-40 years, 6= 41-45 years, 7= 46 – 50 years, 8= 50 years or older
parent_type	bq1.8	Parenting type	Dichotomous	1= co-parenting 2= parenting alone
gender	bq1.9	Gender	Nominal	0= female, 1= male, 2= other/undetermined/not willing to respond
expectation	bq1.10	Expectation on treatment outcome	Nominal	1 "Very well" 2 "Well" 3 "Unsure" 4 "Not very well" 5 "Not well" 6 "Missing"
pcs_b1	bq2.1	Parenting confidence at baseline Q1	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time; 4= not applicable
pcs_b2	bq2.2	Parenting confidence at baseline Q2	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b3	bq2.3	Parenting confidence at baseline Q3	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b4	bq2.4	Parenting confidence at baseline Q4	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b5	bq2.5	Parenting confidence at baseline Q5	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b6	bq2.6	Parenting confidence at baseline Q6	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b7	bq2.7	Parenting confidence at baseline Q7	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b8	bq2.8	Parenting confidence at baseline Q8	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b9	bq2.9	Parenting confidence at baseline Q9	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time; 4= not applicable
pcs_b10	bq2.10	Parenting confidence at baseline Q10	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b11	bq2.11	Parenting confidence at baseline Q11	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b12	bq2.12	Parenting confidence at baseline Q12	Ordinal	3=No hardly ever; 2=No, not very often; 1= Yes, some of the time; 0= Yes, most of the time
pcs_b13	bq2.13	Parenting confidence at baseline Q13	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b14	bq2.14	Parenting confidence at baseline Q14	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_b15	bq2.15	Parenting confidence at baseline Q15	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e1	fq2.1	Parenting confidence at 14 days Q1	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time; 4= not applicable
pcs_e2	fq2.2	Parenting confidence at 14 days Q2	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e3	fq2.3	Parenting confidence at 14 days Q3	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time

pcs_e4	fq2.4	Parenting confidence at 14 days Q4	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e5	fq2.5	Parenting confidence at 14 days Q5	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e6	fq2.6	Parenting confidence at 14 days Q6	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e7	fq2.7	Parenting confidence at 14 days Q7	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e8	fq2.8	Parenting confidence at 14 days Q8	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e9	fq2.9	Parenting confidence at 14 days Q9	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time; 4= not applicable
pcs_e10	fq2.10	Parenting confidence at 14 days Q10	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e11	fq2.11	Parenting confidence at 14 days Q11	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e12	fq2.12	Parenting confidence at 14 days Q12	Ordinal	3=No hardly ever; 2=No, not very often; 1= Yes, some of the time; 0= Yes, most of the time
pcs_e13	fq2.13	Parenting confidence at 14 days Q13	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e14	fq2.14	Parenting confidence at 14 days Q14	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
pcs_e15	fq2.15	Parenting confidence at 14 days Q15	Ordinal	0=No hardly ever; 1=No, not very often; 2= Yes, some of the time; 3= Yes, most of the time
gc_end	fq3.1	Global change at day 14	Ordinal	1=completely recovered, 2= much improved, 3= slightly improved, 4= No change, 5= Slightly worse, 6= Much worse, 7= vastly worse
ur_reported	fq3.2	Concerns	Bi-nomial	0=no, 1=yes
ur_typeX	fq3.3	Concern type	Bi-nomial (7 variables)	1- more distress 0= no, 1= yes 2- more crying 0= no, 1= yes 3- more unsettled 0= no, 1= yes 4- more vomiting 0= no, 1= yes 5- difficulties feeding 0= no, 1= yes 6- difficulties sleeping 0= no, 1= yes 7- other 0= no, 1= yes
ur_type7_txt	fq3.3.1	Other concerns	Open text	String
other_careX	fq3.5	Other care	Bi-nominal (8 variables)	1- GP 0= no, 1= yes 2- A&E 0= no, 1= yes 3- unplanned hospital admission 0= no, 1= yes 4- other osteopath 0= no, 1= yes 5- paediatrician 0= no, 1= yes 6- pharmacist 0= no, 1= yes 7- lactation consultant or midwife 0= no, 1= yes 8- other 0= no, 1= yes
other_care8_txt	fq3.5.1	Other care	Open text	String
pharmX	fq3.4	Pharm & Supplements	Bi-nominal (7 variables)	1- Prescribed medicine 0= no, 1= yes 2- Anti-gas drops 0= no, 1= yes 3- Herbal supplements 0= no, 1= yes 4- Probiotics 0= no, 1= yes 5- Non-cow or ant-allergenic formula milk 0= no, 1= yes 6- Omeprazole 0= no, 1= yes 7- Other 0= no, 1= yes
pharm7_txt	fq3.4.1	Pharm & Supplements	Open text	String
help_belief_end	fq3.6	Help belief at day 14	Dichotomous	0=no, 1=yes
experience_end	fq4.1	Experience at day 14	Ordinal	1= very good, 2=fairly good, 3=neither good nor bad, 4=fairly poor, 5=very poor
satisfaction_end	fq4.2	Satisfaction	Ordinal	1=very satisfied, 2=fairly satisfied, 3=Neither satisfied nor dissatisfied, 4=fairly dissatisfied,5=very dissatisfied
allocation_guess	fq4.3	Participant's guess about allocation	Nominal	0=Control treatment (TTR), 1=Test treatment (GTR), 2=Don't know/unsure

2. Diary and group allocation entries

Dataset name: group.csv (*one line per participant*)
Source: Castor & Main investigator
Format: csv

- Group allocation is named A and B without the statistician knowing which group is which until after the analysis is complete.
- Per protocol analysis is completed from the detailed description of each deviation.

Details for data format

Variable	Source	Description	Type	Levels
id	diary	UID	Numerical	integer
pilot	Data manager	Origin of the data	Dichotomous	0=Main trial, 1=Pilot phase
group	Castor ECD	Group allocation	Dichotomous	Blinded to statistician
perprotocol	Participant log	Management was in accordance to protocol	Bi-nomial	0=No, 1=Yes
deviation_txt	Principle investigator	Text description of protocol deviation	String	

Dataset name: diary.csv (*one line per day per participant*)
Source: Crying diaries, crude data reported hourly
Format: csv

Details for data format

Variable	Source	Description	Type	Levels
id	diary	UID	Numerical	integer
h6	diary	Crying time from 6AM to 7AM (mins)	Continuous	minutes
h7	diary	Crying time from 7AM to 8AM (mins)	Continuous	minutes
h8	diary	Crying time from 8AM to 9AM (mins)	Continuous	minutes
h9	diary	Crying time from 9AM to 10AM (mins)	Continuous	minutes
h10	diary	Crying time from 10AM to 11AM (mins)	Continuous	minutes
h11	diary	Crying time from 11AM to 7AM (mins)	Continuous	minutes
h12	diary	Crying time from 12AM to 7AM (mins)	Continuous	minutes
h13	diary	Crying time from 1PM to 2PM (mins)	Continuous	minutes
h14	diary	Crying time from 2PM to 3PM (mins)	Continuous	minutes
h15	diary	Crying time from 3PM to 4PM (mins)	Continuous	minutes
h16	diary	Crying time from 4PM to 5PM (mins)	Continuous	minutes
h17	diary	Crying time from 5PM to 6PM (mins)	Continuous	minutes
h18	diary	Crying time from 6PM to 7PM (mins)	Continuous	minutes
h19	diary	Crying time from 7PM to 8PM (mins)	Continuous	minutes
h20	diary	Crying time from 8PM to 9PM (mins)	Continuous	minutes
h21	diary	Crying time from 9PM to 10PM (mins)	Continuous	minutes
h22	diary	Crying time from 10PM to 11PM (mins)	Continuous	minutes
h23	diary	Crying time from 11PM to midnight (mins)	Continuous	minutes
h24	diary	Crying time from midnight to 1AM (mins)	Continuous	minutes
d14	diary	Crying time from 1AM to 2AM (mins)	Continuous	minutes
d13	diary	Crying time from 2AM to 3AM (mins)	Continuous	minutes
d14	diary	Crying time from 3AM to 4AM (mins)	Continuous	Minutes
d13	diary	Crying time from 4AM to 5AM (mins)	Continuous	minutes
missing	diary	Number of missing time slots	Continuous	Count

Dataset name: diary2.csv (*one line per participant*)
Source: Crying diaries with only summary daily crying time
Format: csv

Details for data format

Variable	Source	Description	Type	Levels
id	diary	UID	Numerical	integer
d0	diary	Crying time day before treatment (mins)	Continuous	minutes
d1	diary	Crying time on day of treatment (mins)	Continuous	minutes
d2	diary	Crying time day 1 after treatment (mins)	Continuous	minutes
d3	diary	Crying time day 2 after treatment (mins)	Continuous	minutes
d4	diary	Crying time day 3 after treatment (mins)	Continuous	minutes
d5	diary	Crying time day 4 after treatment (mins)	Continuous	minutes
d6	diary	Crying time day 5 after treatment (mins)	Continuous	minutes
d7	diary	Crying time day 6 after treatment (mins)	Continuous	minutes
d8	diary	Crying time day 7 after treatment (mins)	Continuous	minutes
d9	diary	Crying time day 8 after treatment (mins)	Continuous	minutes
d10	diary	Crying time day 9 after treatment (mins)	Continuous	minutes
d11	diary	Crying time day 10 after treatment (mins)	Continuous	minutes
d12	diary	Crying time day 11 after treatment (mins)	Continuous	minutes
d13	diary	Crying time day 12 after treatment (mins)	Continuous	minutes
d14	diary	Crying time day 13 after treatment (mins)	Continuous	Minutes

3. Serious Adverse Events (one line for each event)

Dataset name: sae.csv
Source: consultation information, open text from follow-up questionnaire
Format: csv
Details for data format: One line for each reported adverse event

Variable	Source	Description	Type	Levels
id	SAE log	UID	Numerical	
Date	SAE log	Date of the event	Date	
sae_ser	SAE log	Level of seriousness	Nominal	0=Non-serious adverse event 1=Unexpected hospital admission 2=Death 3=Irreversible morbidity
sae_rel	SAE log	Type of relationship with intervention	Nominal	0=Not related, 1=Unlikely, 2=Possibly, 3= Probably, 4= Definitely

4. Treatment (one line per treatment session)

Dataset name: treatment1.csv | treatment2.csv
Source: consultation information, follow-up questionnaire
Format: csv
Transformed data

- "treatment1" holds all the specific data from the test group (TTR). For the file to remain anonymous, the version that is sent to the statistician will have alternative IDs making impossible to match the data to individual data from the trial. The final archived dataset should contain true IDs.

- "treatment2" holds all the data on advice with ID data making it possible to link data to each participant and compare them between allocated groups.

- Text fields are uniquely for data cleaning purposes. This makes it possible to make sure each treatment is correctly categorised.

Details for data format "treatment1.csv"

Variable	Source	Description	Type	Levels
id	Generated from study ID	Newly generated IDs to make it impossible for statistician to know which group is which	Numerical	
ttt_blt	os4.2 | 7.1 | 10.1 | 13.1	BLT	Dichotomous	0=no, 1=yes
ttt_btt_c	os4.3 | 7.2 | 10.2 | 13.2	BTT(cranium)	Dichotomous	0=no, 1=yes
ttt_btt_tpl	os4.4 | …	BTT(trunk, pelvis, limb)	Dichotomous	0=no, 1=yes
ttt_btt_v	os4.5	BTT(viscera)	Dichotomous	0=no, 1=yes
ttt_btt_0	os4.6	BTT(other)	Dichotomous	0=no, 1=yes
ttt_cff	os4.7	Central fluid fluctuation (CFF)	Dichotomous	0=no, 1=yes
ttt_pff	os4.8	Peripheral fluid fluctuation (PFF)	Dichotomous	0=no, 1=yes
ttt_count	os4.9	Counterstrain/Facilitated Positional Release	Dichotomous	0=no, 1=yes
ttt_lfm	os4.10	Lymphatic fluid management (pumping and drainage)	Dichotomous	0=no, 1=yes
ttt_jart	os4.11	Joint articulation	Dichotomous	0=no, 1=yes
ttt_ind	os4.12	Indirect / functional	Dichotomous	0=no, 1=yes
ttt_myoR	os4.13	Myofascial release	Dichotomous	0=no, 1=yes
ttt_visc	os4.14	Visceral manipulation	Dichotomous	0=no, 1=yes
ttt_st	os4.15	Soft tissue techniques	Dichotomous	0=no, 1=yes
ttt_other	os4.16	Other	Dichotomous	0=no, 1=yes

Details for data format "treatment2.csv"

Variable	Source	Description	Type	Levels
id	Castor	UID	Numerical	numerical code
consult		Number of consultations	Ordinal	
ad_rout		Advice on routines	Dichotomous	0=no, 1=yes
ad_sleep	os5.1 | 8.1 | 11.1 | 14.1	Advice on sleep	Dichotomous	0=no, 1=yes
ad_feeding	os5.2 | 8.2 | 11.2 | 14.2	Advice on feeding	Dichotomous	0=no, 1=yes
ad_diet	os5.3 | …	Advice on diet		
ad_handling	os5.4	Advice on handling	Dichotomous	0=no, 1=yes
ad_env	os5.5	Advice on environment	Dichotomous	0=no, 1=yes
ad_beh	os5.6	Advice on parenting behaviour	Dichotomous	0=no, 1=yes
ad_reas	os5.7	Reassurance	Dichotomous	0=no, 1=yes
ad_other	os5.8	Advice other	Dichotomous	0=no, 1=yes

5. Description of osteopaths

Dataset name: ost.csv
Source: Specific questionnaire provided once osteopaths are eligible for entering participants
Format: csv

Only osteopaths that recruited participants are to be described.

Variable	Description	Type	Levels
ost	Code for osteopath	Nominal	numerical code
country	Country of practice	Nominal	1=UK, 2=Australia, 3=Switzerland
ost_age	Osteopath age (years)	Ordinal	Numerical
ost_gender	Osteopath gender	Dichotomous	0=male, 1=female
ost_training	Formal training in pediatric osteopathic care	Binominal	0= no, 1= yes
years_pract	Years in practice	Ordinal	Numerical

4. Blinding of statistician

Statistical blinding will be assured to generate summary variables and set the final analysis for between group differences. Once the final statistical coding is defined and validated by the TSC, the true group identity will be revealed and additional descriptive information on treatment will be provided.

- Coding allocation to random A, B to blind first step of statistical analysis
- Treatment description included in separate dataset to be coded with different IDS than main dataset.

5. Outcome measures

Primary Outcome measure

Crying time diary

The main outcome is daily crying time in minutes. This can be measured for crying times from the first day after treatment to the thirteenth day after treatment. Daily crying time is obtained by adding hourly reported crying times from 6AM to 5:59AM the next day. D-1 is computed from the hourly reported crying time on the day prior to the treatment if the entire day is completed. When incomplete, it is replaced by the parent reported 24h baseline crying time (castor.csv). Daily crying time is considered as missing if any of the time hour slots has missing values within the day.

The primary outcome is the between group average difference in crying time from day 1 after the treatment to day 13 after the treatment. Repeated measures random linear mixed effect adjusting for baseline crying time, expectations for treatment outcome and infant's age is used to estimate the between group difference.

If diaries only report the presence of crying during each hour slot, then each time slot is replaced by the corresponding median reported crying time for infants that were crying during that hour on that specific day of follow-up.

Secondary Outcome measure

Parent Confidence

Changes of parent confidence will be measured using the Parent Confidence Score at baseline and 14 day follow up. This score relies on 15 questions about parenting confidence with a likert scale of 4 choices (0 to 45 points). Changes over time will be measured at an individual level by adding the baseline value as a dependant variable in the mixed effect model. Movement of 6 points or more is to be considered as meaningful.

Global change

This question asks about overall change in the infant's symptoms, from completely recovered to vastly worse. (Completely recovered, much improved, slightly improved, no change, slightly worse, much worse, vastly worse). These levels are coded numerically in an ordinal manner and are to be compared between groups using non-parametrical approaches.

Experience and satisfaction

Experience was measured using a 5-point likert scale from Very good, fairly good, neither good not bad, fairly poor and very poor.

Satisfaction was measured using a 5-point likert scale from Very satisfied, fairly satisfied, neither satisfied nor dissatisfied, fairly dissatisfied and very dissatisfied.

We also asked if the parent thought osteopathic care helped their baby using a yes, no or unsure response.

Unexpected and or Adverse events

We asked parents about additional care they may have administered to their infants and additional health care consultations they sought during the trial period as a proxy indicator of adverse events, and we also asked directly about any changes in symptoms that cause the parent concern. These included more distress, more crying, more unsettled, more vomiting, increased difficulty in feeding, sleeping and anything else (other).

Serious adverse events are reported by parents and managing osteopaths. These are assessed by the trial managing group and coded as for the reason for being a SAE and the estimated likelihood these are related to the treatment.

6. Randomisation

Block randomisation with variable block size of 4 and 6 generated by CastorEDC.

7. Sample size

From data collected during the pilot phase (n=13), we confirmed crying times as expected by Wolke et al. We estimate the average daily crying time over the two-week follow-up period to be of 120 minutes with a SD of 45 minutes. To detect a minimum of 30 minutes additional reduction in crying time between the intervention and control group with 90% power and a two-sided 5% significance level we would need 48 infants in each arm of the trial. If we allow for a 15% drop out we would need to randomise 112 infants, 66 in each arm of the trial. This sample size calculation is conservative because it disregards the fact that there are repeated measures in each infant and that the model will be corrected for the baseline crying time value. Thus, the actual power will be higher.

Power	MCID (daily reduction in crying time in minutes)			
	10	15	30	60
0.8	636	284	72	18
0.9	852	380	**96**	24
0.95	1054	468	118	30

Table 4 :   Sample size estimates without loss to follow up.

8. Statistical Analysis

All analysis will be intent-to-treat except for a secondary per protocol analysis on the primary outcome. All *P* values will be two sided, and the significance level is set at 5%. We will use STATA15 to precode the entire analysis prior to data collection (.do file). Deviations from the initial statistical plan will be recorded on the statistical output report.

Computing new variables

Protocol deviation

A new variable will be computed to identify deviations from protocol. Any participant who dropped-out, missed out a planned treatment, received external osteopathic care during the follow-up, or did not complete the diary daily are considered as participants who deviated from the protocol.

Unexpected reactions and adverse events

Serious adverse events, sae, will be reported as a total number for each group and as a risk ratio.

Serious adverse events will be monitored throughout the trial and reported by either the osteopaths or the participant parent.

Detailed reporting of unexpected reactions from follow-up and from the follow-up questionnaire are to be summarised to describe the presence of any unexpected reactions during the follow-up period.

Treatment description

A new set of variables is to be merged from detailed description of treatments during each visit and summarize whether each subset of care was provided or not.

Parent confidence scale

Parent confidence scale at baseline (pcs_b) and at 14 days (pcs_e) are computed from the 15 individual questions. Likert scales are scored 0 to 3 points. Scores for Q13 are inverted. Values for not applicable in Q1 and Q9 are given the score of 2. The points are then added up.

For descriptive purposes, the parent confidence scale at baseline can be compared with the following ranges:

- Non-clinical range 40 or more
- Mild clinical range 36-39
- Moderate clinical range 31-35
- Severe clinical range 31 or less

Missing data

Where participants withdraw, we will compare the characteristics of those withdrawing against those who remain in the study. Furthermore, we will do an:

- overall descriptive analysis of missing data searching for patterns,
- compare between group differences in missing data,
- Analyse potential correlations between missing data in diary and overall summary measures.

Missing data for hourly entries will not be replaced (e.i. not missing total at random) unless parents did not report minutes in an explicit was. Graphical representations are replaced by estimates of hourly crying time. Reported crying within a hourly time slot without specification on the duration is replaced by the median observed crying time from other reported crying for that specific time slot (i.e. hour of the day and day of the trial).

Given the nature of the analysis, missing data within the diary will only reduce the power but should not affect the overall analysis.

For a diary to be considered, at least 7/14 days need to be filled in. Furthermore, best-case, and worst-case scenario will be simulated for the primary outcome implementing lowest and highest values for missing data (see sensitivity analysis).

Descriptive analysis

Descriptive statistics will be used to summarise the characteristics of participants in each arm of the trial. This will include baseline values, treatment schedules, compliance and follow-up. We will show number and percentage or mean and standard deviation for categorical and normally distributed continuous variables respectively. For data severely deviating from a normal distribution, we will present median and interquartile ranges. Probability of observed between groups difference will be provided using either Fisher's exact test, Chi2 test, and or Student's t-test.

Primary outcome analysis

We will analyse the primary outcome of crying time (e.i. individual daily measures of mean crying time in minutes) using random-effects multilevel linear regression. Crying time will be entered as the response variable (dependant) and the link function is identity. Therefore, lack of independence for measures from the same participant will be accounted for. To improve statistical power, the following explanatory variables are to be added to improve the modelling of the between-group differences
- baseline crying time (continuous),
- expectations from treatment (ordinal)
- and age of infants at entry (continuous).

To account for missing data, the data can be categorized and treated as nominal.

Results will be reported as between group difference in crying time in minutes with a 95% confidence interval. Statistical inference for significance of effects of the intervention will be calculated using Wald Chi^2 test.

Checking assumptions for primary outcome

a) Assumptions for repeated measures random linear mixed effect

The following assumptions for repeated measures random linear mixed effect need to be tested:
- The explanatory variables are related linearly to the response.
- The errors have constant variance.
- The errors are independent.
- The errors are normally distributed.

A graphical representation will plot the residuals against group allocation. This can indicate if the model fits the data appropriately. The residuals will also be plotted in the order of days after treatment to see any indications of visible trend over time. The residuals will be plotted against the fitted values to test whether there is any non-constant error variance. Finally, a histogram with a normal probability plot and Wilk-Shapiro test will be used to test the assumption of normality for crying-time distribution in each group.

If the distribution of crying time is not Gaussian, then it will be categorised, and the distribution of the response variable will be considered as ordinal, and the link function is logit. When the assumptions for distribution of the residuals are not met, transformation is to be used to make the analysis possible.

b) Testing and accounting for trends over time

Three methods are used to investigate trend over time:

a. Visual and statistical assessment of trend over time is to be made using fractional polynomials in regression analysis separately for each group.
b. Repeated measures random linear mixed model is to be used to test interactions between group allocation and days of follow-up. This makes it possible to assess that group differences are constant through time. Explanatory variables are group allocation, interaction between group allocation and number of days of follow-up, baseline crying time day before treatment, infant age at entry, treatment expectations. Linearity of response for the explanatory variable is tested as trends over time using mixed regression analysis.
c. Day to day means in each group will be calculated using Delta-method linear prediction with fixed portion. Furthermore, linear trend over time will be tested in each group by modelling day to day residuals within the model.

Assumptions are to be verified on the linearity of the progression through time. Transformation is to be used if linearity is not met.

Secondary outcomes analysis

Mixed effect linear regression will be used to test continuous secondary outcomes. Categorical response variables using Likert scales (i.e. global change, parent satisfaction) will be considered as ordinal and non-parametrical tests (Mann Whitney U test) will be used to test between group differences. When possible (results from fractional polynomials), the outcome is to be dichotomised and the link function will be logit. Between group changes in parenting confidence score will compare changes from baseline by entering baseline score as an explanatory variable.

Proportion of participants who reported at least one adverse event will be compared between groups using a Fisher exact test. The same approach will be used to compare unexpected reactions (distress, crying, unsettled, vomiting, difficulties feeding, difficulties sleeping, other).

For cost of treatment, we will determine the mean number of treatments per infants in the trial and per group, the mean number of treatments will be multiplied by the average cost of an osteopathic treatment in UK pounds, determined by the latest data from the Institute of Osteopathy census.

Descriptive exploratory analysis

- Analyse trends over time using fractional polynomials.
- Adjust for osteopath level variables if between osteopath heterogeneity.

- Description of eventual correlation between years of experience and treatment effects (i.e. analyse of trend and comparison for cut-off at 3 or less years versus 4 or more.
- Investigate eventual association between post-graduate training and treatment outcome (interaction term).

Complementary confirmatory analysis

- A sensitivity analysis without imputed median values for crying times in time slots were parents reported crying will be run.
- A sensitivity analysis to adjust for recruiting site (ESO, UCO, Private practice, Swiss and Australian sites).
- A sensitivity analysis is planned to verify that the observed results hold in the worst- and best-case scenarios for missing data. The 25th and 75th percentile are to be used to replace missing data (25th for one group and 75th for the other) before running the primary analysis.

- If baseline imbalance (p-value<0.10) is observed, a secondary analysis with adjustment for these variables will be run to confirm results.
- A sensitivity analysis adjusting for number of sessions, counselling and adjunctive treatment will be done to verify that the observed differences are not due to imbalance in patient management other than the tested intervention.
- Per protocol analysis.

9. Results analysis

Provisional tables

Table 1 Baseline characteristics for each group. GTR=, TTR=, SD=standard deviation

Characteristics	Test group (GTR) N=xx	Control group (TTR) N= xx	Group imbalance (p-values)
Females; n (%)			
Age of infant; n (%)			
1–14 days (1–2 weeks)			
15–28 days (3–4 weeks)			
29–42 days (5–6 weeks)			
43–56 days (7–8 weeks)			
57–70 days (9–10 weeks)			
71–84 days (11–12 weeks)			
Infant weight in kg; mean (SD)			
At birth			
At baseline			
Excessive crying, distress or unsettlement prior 24h; mean (SD)			
Age of mother in years; mean (SD)			
Siblings; n (%)			
None			
1			
2			
3 or more			
Co-parenting; n (%)			
Parenting Confidence Score; n (%)			
Non-clinical range (40 or more)			
Mild clinical range (36-39)			
Moderate clinical range (32-35)			
Severe clinical range (31 or less)			
Expected response to osteopathic care; n (%)			
Very well			
Well			
Unsure			
Not very well			
Not well			
Missing			

* Significant group imbalance at baseline

Table 2 Effects of treatment on crying time in minutes

Outcomes	Test group (TTR) N=xx	Control group (GTR) N= xx	Adjusted * between group difference	Unadjusted between group difference
	Mean (Std.Err)	Mean (Std.Err)	Mean (CI 95%; P-value)	Mean (CI 95%; P-value)
Mean daily crying time (minutes)				
From day 1 to day 13				
From day 1 to day 6				
From day 7 to day 13				
Daily crying times (minutes)				
Day before treatment				
Treatment day				
Days after treatment				
Day 1				
Day 2				
Day 3				
Day 4				
Day 5				
Day 6				
Day 7				
Day 8				
Day 9				
Day 10				
Day 11				
Day 12				
Day 13				

    * Adjusted for baseline crying time, infant age and prior expectations for osteopathic care

Table 3 Effects of treatment on other outcomes including undesirable events

Outcomes	Test group (TTR)	Control group (GTR)	Between group difference*
	N=xx	N= xx	(CI 95%; P-value)
Parenting Confidence Score [0–45]; mean (SD)			
Perceived changes in symptoms; n(%)			
Completely recovered			
Much improved			
Slightly improved			
No change			
Slightly worse			
Much worse			
Vastly worse			
Weight changes over 14 days (kg); mean (SD)			
Satisfaction with received care; n (%)			
Very satisfied			
Fairly satisfied			
Neither satisfied nor dissatisfied			
Fairly dissatisfied			
Very dissatisfied			
Unexpected reactions; n(%)			
More distress			
Crying more			
More unsettled			
Vomiting more			
Increased feeding difficulties			
Increased difficulties sleeping			
Other			
Serious adverse events; n(%)			
Death			
Hospitalisation			
Long lasting morbidity			
Any			

* P-values measured using likelihood ratio test in linear/logistic regression, Mann-Whitney U test for ordinal outcomes, Chi2 for categorical

Table 4 Number of sessions, advice given, additional treatment and effectiveness of blinding

	Test group (TTR) N=xx	Control group (GTR) N= xx	Between group difference*
	Mean (SD)	Mean (SD)	(CI 95%; P-value)
Number of sessions; mean (SD)			
Advices; n(%)			
Managing the baby's sleep pattern			
Managing the baby's feeding pattern			
Mother's diet			
Handling the baby			
Managing the baby's environment			
Parenting behaviour			
Reassure parent/guardian			
Other			
Medication and remedies; n(%)			
Prescribed medication			
Anti-gas drops			
Herbal supplements			
Probiotics			
Non-cow or anti-allergenic formula milk			
Omeprazol			
Other			
Additional care; n(%)			
General practitioner			
Accident and Emergency			
Unplanned hospital admission			
Another osteopath			
Paediatrician			
Pharmacist			
Lactation consultation or midwife			
Other			
Blinding; n(%)			
Thinks the baby was in test group (TTR)			
Thinks the baby was in the control group (GTR)			
Doesn't know / unsure			

* Adjusted for baseline crying time. P-values measured using likelihood ratio test in liner/logistic regression

Provisional Figures

Figure 1 Flow chart of participant inclusion and follow-up

Figure 2 Average daily crying time during follow-up between groups

Figure 3 Types of manual care received over the treatment period