EL4BS

EARLY LASER FOR BURNS SCARS

STATISTICAL AND HEALTH ECONOMIC ANALYSIS PLAN (SHEAP)

Short Title: ELABS - Early Laser for Burn Scars

Title: Early Laser for Burn Scars (ELABS) - A multi-centre randomised controlled trial of the effectiveness and cost-effectiveness of the treatment of hypertrophic burn scars with Pulsed Dye Laser and standard care compared to standard care alone

ISRCTN Number: 14392301

Funding Body: NIHR Research for Patient Benefit (NIHR201080)

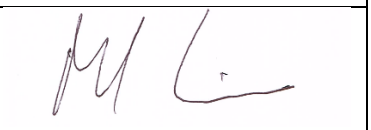
IRAS and HRA Number: 283345

Sponsor: Salisbury NHS Foundation Trust (Ref: RNZ)

Chief Investigator: Dr Mark Brewin, Salisbury NHS Foundation Trust

Current stage of SHEAP: First Draft

Statistical and Health Economic Analysis Plan Final Sign-Off:

	Name	Date	Signature
Author	Sharon Docherty	28/04/2023	Sharon Docherty
Author (economics)	Katie Breheny	10/05/2023	Katie Breheny
Author (statistics)	Obioha Ukoumunne	28/04/2023	Obioha Ukoumunne
Chief Investigator	Mark Brewin	08/06/2023	
Independent statistician	Lizzie Pitt	5/06/2023	Lizzi Pitt

Amendments:

Amendment Number	Date	Sign-off

Abbreviations and definitions

Abbreviation	Meaning
A&E	Accident and Emergency
BNF	British National Formulary
CACE	Complier Average Causal Effect
CARe	Centre for Appearance Research burn scales
CEA	Cost-effectiveness analysis
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
CUA	Cost-Utility Analysis
ELABS	Early Laser for Burn Scars
GP	General Practice
ICER	Incremental cost-effectiveness ratio
ISRCTN	International Standard Randomised Controlled Trial Number
MCID	Minimal Clinically Important Difference
NMB	Net Monetary Benefit
PDL	Pulsed Dye Laser
PPI	Patient and Public Involvement
POSAS	Patient Observer Scar Assessment Scale
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALYs	Quality-Adjusted Life Years
QoL	Quality of Life
SD	Standard deviation
SF-6D	Method for deriving the quality of life score from the SF-12
SF-12	12-item Short Form survey

SHEAP	Statistical and Health Economic Analysis Plan
TBSA	Total Burn Surface Area
TMG	Trial Management Group

Contents

1 Study summary

Study title	Early Laser for Burn Scars (ELABS) - A multi-centre randomised controlled trial of the effectiveness and cost-effectiveness of the treatment of hypertrophic burn scars with pulsed dye laser (PDL) and standard care compared to standard care alone.
Short study title:	ELABS - Early Laser for Burn Scars **(ELABS)**
Study design	This study is a two-arm, definitive superiority pragmatic parallel group multi-centre randomised controlled trial with qualitative component and health economic evaluation comparing pulsed dye laser treatment with standard care (intervention) to standard care alone (control).
Study participants	**Inclusion Criteria** NHS patients with >1% Total Body Surface Area burn injury who: • Have delayed healing of >2 weeks • Have potential for hypertrophic scarring • Are suitable for scar management therapy **Exclusion criteria** • Unable to give informed consent • Below 16 years of age • Prone to keloid scarring
Number of participants	150 participants (allows for 20% drop out). 20 participants will take part in semi-structured interviews.
Follow-up duration	Patients will be followed up in clinic at 6 months (26 weeks) post-baseline. Qualitative telephone interviews with up to 20 participants will take place after the 6 month follow-up.
Planned study period	36 months, including a recruitment period of 21 months.

Study aim	The aim is to assess the effectiveness of treating hypertrophic burns scars with PDL at an early stage of scar formation.
Study objectives	1. Determine whether there is a difference in scar quality between the two treatment groups. 2. Determine whether there is a difference in quality of life between the two treatment groups. 3. Estimate which is the cost-effective treatment option to inform decision-making. 4. Test factors that potentially moderate the clinical effectiveness of pulsed dye laser treatment on scar quality 5. Summarise the adverse event profile of pulsed dye laser in treatment of burn scars 6. Explore patient experiences and the psychological and social impact of their treatment [details regarding the analysis of this qualitative component are described in a separate plan].
Outcome measure data	**Primary outcome:** The primary outcome measure is patient-rated scar quality at 6 months as quantified by the Patient Observer Scar Assessment Scale (POSAS). **Secondary outcomes:** • Observer-rated POSAS at 6 months • Patient-rated POSAS at 6 and 12 weeks • Patient perception of scar change at 6 months • Objectively measured scar colour (erythema and melanin index) at 6 months • Patient-rated quality of life (QoL) (quantified by the Centre for Appearance Research burn scales (CARe) and the 12-item Short Form survey (SF-12)) at 6 months **Other data collected:** • Baseline characteristics • Healthcare resource use

	• Adverse events • Patient experience and psychosocial impact
Interventions	Patients will be allocated in a 1:1 ratio: 1. A course of 3 Pulsed Dye Laser (PDL) treatments, in addition to standard care (intervention). 2. Standard care alone (control).

2 Aims and objectives for the statistical and health economic analyses

To evaluate the effectiveness and cost-effectiveness of pulsed dye laser (PDL) treatment in addition to standard care (intervention arm) compared to standard care alone (control arm).

2.1 Primary outcomes

To determine whether there is a difference in patient-rated scar quality (using the Patient Observer Scar Assessment Scale (POSAS)) at 6 months between patients receiving PDL in addition to standard care (intervention arm) and those on standard care alone (control arm).

2.2 Secondary outcomes

To determine whether there is a difference between the two groups in scar quality based on:

- Observer-rated (blinded to treatment allocation) POSAS score at 6 months
- Patient-rated POSAS scores at 6 weeks and 12 weeks
- Objectively measured scar colour (erythema and melanin index) at 6 months
- Patient perception of scar change at 6 months

To determine whether there is a difference between the two groups in quality of life at 6 months based on:

- Centre for Appearance Research burn scales (CARe)
- 12-item Short Form survey (SF-12)

2.3 Additional analyses

- Test factors as potential moderators (age, scar location) of the effect of the PDL on patient-rated POSAS at 6 months
- To describe the adverse effect profile of PDL in treatment of burn scars.
- To conduct a within-trial economic evaluation of PDL to assess its cost-effectiveness.
- To qualitatively explore patient experience and psychosocial impact [details covered in a separate analysis plan]

3 Objectives and design of trial

This study is a two-arm, definitive superiority pragmatic parallel group multi-centre randomised controlled trial with qualitative component and health economic evaluation comparing pulsed dye laser treatment with standard care (intervention arm) to standard care (control arm) alone. Individual participants are randomised in a 1:1 ratio to receive either the intervention or control.

3.1 Objectives

The objectives are to:

1. Evaluate the effectiveness of pulsed dye laser treatment relative to standard care alone for improving scar quality
2. Evaluate the effectiveness of pulsed dye laser treatment relative to standard care alone for improving quality of life
3. Test factors that potentially moderate the clinical effectiveness of pulsed dye laser treatment on patient-rated scare quality
4. Summarise the adverse event profile of pulsed dye laser in treatment of burn scars
5. Evaluate the cost-effectiveness of pulsed dye laser treatment relative to standard care alone
6. Explore patient experience and psychosocial impact using a qualitative framework

This statistical and health economic analysis plan is focussed on the first five objectives. The analysis plan for the sixth objective will be fully described in a separate document.

3.2 Trial design

The study is a two-arm, definitive superiority pragmatic parallel group multi-centre (multi-site) randomised controlled trial with individual participants randomised to receive either pulsed dye laser treatment with standard care (intervention arm) or standard care alone (control arm). Participants are patients within 3 months of healing from a burn injury. Data collection occurs at baseline (pre-randomisation), 6 weeks, 12 weeks and 6 months follow-up (see Table 1).

Participants will be randomised in a 1:1 ratio to the intervention and control arms, stratified by study site.

Table 1: Summary of data collection points linked with study objectives

	Time point	**Post - allocation**			
Measure	**Enrolment**	**0 weeks**	**6 weeks**	**12 weeks**	**6 months**
Demographics	✓	x	x	x	x
1. Evaluate the effectiveness of pulsed dye laser treatment relative to standard care alone for improving scar quality					
Patient-reported POSAS	✓	x	✓	✓	✓
Observer-reported POSAS	✓	x	x	x	✓
Patient's perception of change in scar quality	x	x	x	x	✓
Scar colour measurement (E and M indices)	✓	x	x	x	✓
2. Evaluate the effectiveness of pulsed dye laser treatment relative to standard care alone for improving quality of life					
CARE	✓	x	x	x	✓
SF12	✓	x	x	x	✓
3. Test factors that potentially moderate the clinical effectiveness of pulsed dye laser treatment on patient-rated scare quality					
Baseline (study site, age and scar body region)	✓	x	x	x	x
4. Summarise the adverse event profile of pulsed dye laser in treatment in treatment of burn scars					
Adverse event register	✓	✓	✓	✓	✓
5. Evaluate the cost-effectiveness of pulsed dye laser treatment relative to standard care alone					
Healthcare resource use	✓	✓	✓	✓	✓
6. Explore patient experience and psychosocial impact using a qualitative framework					
Qualitative Interviews	x	x	x	x	✓

3.3 Sample size

No published data were found on the Minimal Clinically Important Difference (MCID) for the POSAS. A service evaluation study at Salisbury NHS Trust, on 15 patients treated with PDL, showed a change in the patient-rated POSAS from a mean (standard deviation (SD)) of 35.8 (10.6) at baseline to 25.7 (11.2) at 12 months. A study in The Netherlands on the effectiveness of silicone treatment on 46 scars from 23 patients showed mean (SD) pre-treatment scores of 31.0 (7.8) and post-treatment scores of 17.4 (11.5) [1]. The mean improvement at 12 months was 10.1 (Salisbury) and 13.6 (The Netherlands). Pooled over both studies and time-points, the SD was around 10. A one-point improvement on each of the 6 items would equate to an overall change of 6 points over the duration of the study of 6 months, which constitutes 11% of the range of the scale (minimum 6, maximum 60). The PPI group felt that a change of 6 points represented an important improvement in scar quality.

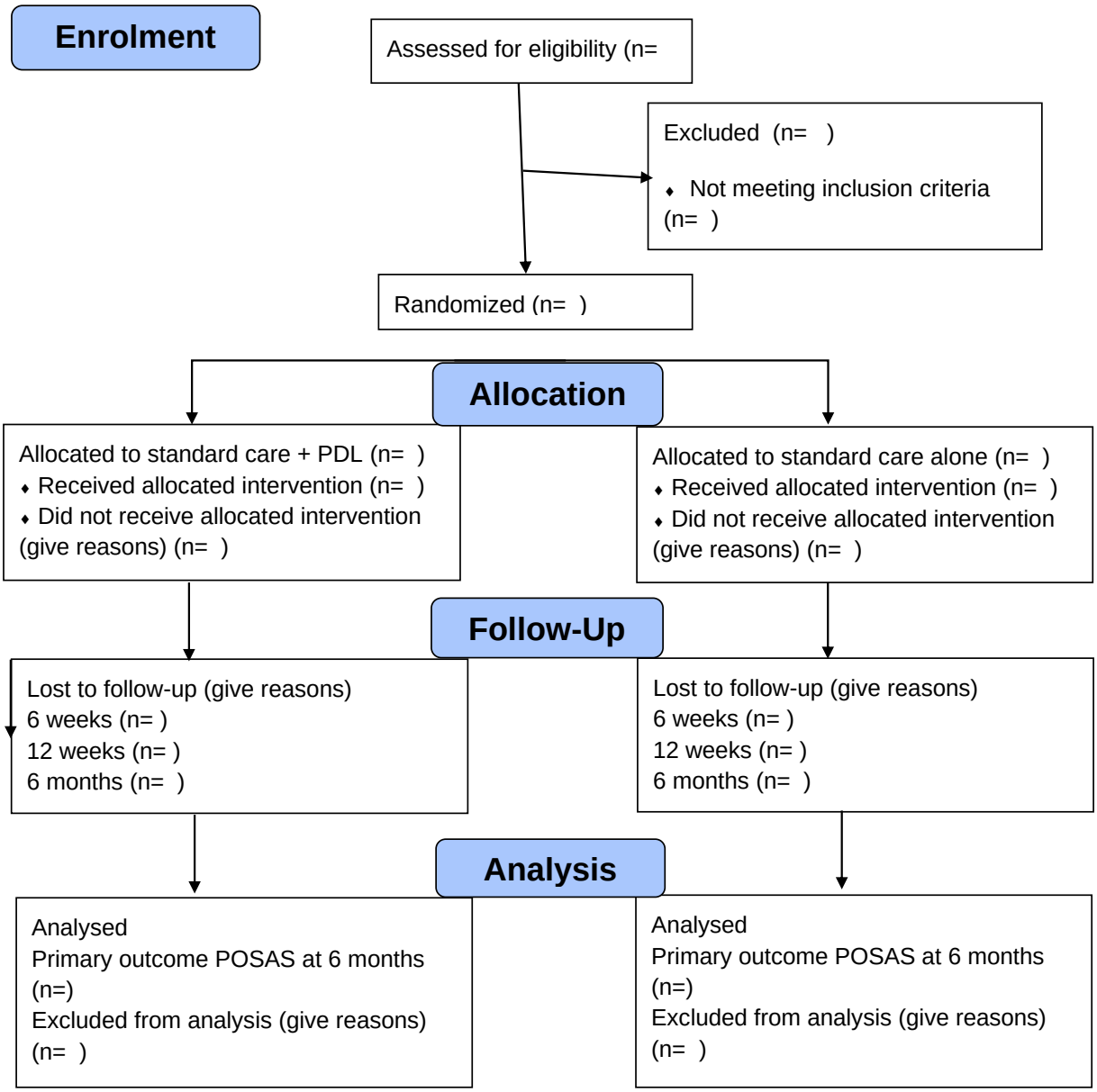
A 1:1 allocation ratio was chosen. This design would require 60 participants in the treatment arm and 60 in the control arm to give 90% power, assuming a 2-sided 5% significance level, standard deviation of 10 and effect size of 6. This gives a total of 120 participants and, allowing for 20% drop out, implies recruitment of 150 participants.

This recruitment of 150 patients over 21 months across 7 centres represents 1 per centre per month and is similar to that identified for adults in the PEGASUS trial [2]. The calculation does not take into account adjustment for baseline POSAS in the analysis. The PEGASUS study found a correlation of 0.545 between baseline and the 6-month follow-up POSAS scores [2].

4 Participant flow

4.1 CONSORT flow chart:

A CONSORT flow chart will be reported showing the flow of recruitment into the RCT (numbers available, approached, eligible, randomised, along with reasons if not approached or not eligible) and through the study (numbers with outcome data, reasons for withdrawing etc.).



5 Missing data

Outcome and resource use data will be sought for all randomised participants. Variables will be tabulated or plotted to identify implausible values and missing items. For missing items and implausible values, the original data will be checked at the source. A decision will be made, informed by clinical opinion, on whether to drop implausible values (i.e. consider values missing).

The main analysis of the clinical outcomes will use a complete case approach. Sensitivity analysis based on imputed data will also be undertaken.

6 Interim analysis

No interim analyses are planned.

7 Blinding

The trial statisticians will be blind to treatment allocation at least until after the statistical analysis plan has been signed and the final dataset has been locked. At analysis the statisticians will be unblinded as the nature of the analyses (e.g., estimation of the complier average causal effect) makes blinding impossible to maintain at that stage. The results of the statistical analysis will be presented to the rest of the trial team blinded to treatment arm. Once the interpretation of the results has been agreed within the trial team then the treatment arms will be un-blinded to the whole trial team.

8 Outcomes

8.1 Primary clinical outcome

The primary clinical outcome, measured at six months post-baseline, is the patient-reported version of the Patient and Observer Scar Scale (POSAS) score. Each of the patient- and observer-reported versions of the POSAS is comprised of 6 items each scored on a 10-point scale from 0 to 10. The total score is the sum of the items and has a possible scoring range from 6 to 60, with higher scores indicating better scar quality.

8.2 Secondary clinical outcomes

Secondary outcomes include the patient-reported version of the POSAS (at 6 weeks and 12 weeks post-baseline), the observer-reported version of the POSAS (at 6 months post-baseline), the Erythema and Melanin Indices of scar colour (at 6 months post-baseline), the patient's perception of change in scar quality between baseline and 6 months (measured at 6 months post-baseline), the Centre for Appearance Research (CARe) burn scales (at 6 months post-baseline) and the 12-item Short Form survey (SF-12) (at 6 months post-baseline).

The Erythema and Melanin Indices are recorded at 6 months post-baseline. Three Colorimeter readings are made at *each of* the treatment (burn) skin site and the reference (normal) skin site to quantify levels of *each of* erythema and melanin – i.e., there are 12 measurements altogether. Each measurement can take values between 0 and 100. The erythema score is calculated as the mean of the 3 replicates at the treatment skin site divided by the mean of the 3 replicates at the reference skin site. In a similar way, the melanin score is calculated as the mean of the 3 replicates at the treatment skin site divided by the mean of the 3 replicates at the reference skin site. Lower scores indicate a more favourable outcome.

Patient's perception of the improvement in scar quality between baseline and 6 months post-baseline will be quantified using a 7 point scale (1 – "Very much Worse"; 2 – "Much Worse"; 3 – "A little Worse"; 4 – "The same"; 5 – "A little Better"; 6 – "Much Better"; 7 – "Very Much Better").

The CARe burn scales (https://www.careburnscales.org.uk/the-scales) include a 44-item version for adults (18 years or over) and a 39-item version for young people (which can be used for those aged 16 and 17 in this study). Each item has a Likert scale response from 0 to 3, such that (after reverse-scoring some items) higher scores indicate better outcomes. There is no total score for the CARe burn scales, only subscale scores. Items scores are recoded and subscale scores are rescaled so that the subscale scores range from a possible 0 to 100. Higher scores are indicative of better outcomes.

For the Adult version of the CARe burn scales, scores can be derived for 12 different subscales:

- *Wound/scar discomfort* (4 items)
- *Physical Well-being* (2 items)
- *Social Situations*? (4 items)
- *Friend Support* (4 items)
- *Work Life* (2 items)
- *Family Support* (4 items)
- *Self-worth* (4 items)
- *Burn Wound/Scar Dissatisfaction* (3 items)
- *Intimacy/romantic relationships* (4 items)
- *Trauma Symptoms* (6 items)
- *Negative Mood* (4 items)
- *Positive growth* (3 items).

For the Young Person version of the CARe burn scales, scores can be derived for 6 different subscales:

- *Social Situations* (8 items)
- *Self-worth* (5 items)
- *Negative Mood*? (7 items)
- *Wound/Scar Dissatisfaction* (8 items)
- *Positive Growth* (5 items)
- *Romantic Appeal* (6 items).

The SF-12 (comprised of 12 items) will be used to quantify quality of life. Details of how the SF-12 will be analysed are provided in the economic evaluation section.

9 Analyses

9.1 Summary of baseline characteristics

Baseline and demographic variables will be summarised separately by trial arm and overall, using means and standard deviations (or medians and interquartile ranges) for continuous variables and numbers and percentages for categorical variables. These include: age (*16 to 17*"; *18 to 24*; *25 to 34; 35 to 44; 45 to 54; 55 to 64; 65 to 74; Over 75*), gender (1 – *Male;* 2 – *Female;* 3 - *Non-Binary;* 4 – *Other;* 5 - *Prefer Not to Say*), ethnicity (1 - *White - English, Welsh, Scottish, Northern Irish or British*; 2 - *White-Irish*; 3 - *White-Gypsy, Roma or Irish Traveller*; 4 - *White-Any other White background*; 5 - *Mixed or Multiple ethnic groups-White and Black Caribbean*; 6 - *Mixed or Multiple ethnic groups-White and Black African*; 7 - *Mixed or Multiple ethnic groups-White and Asian*; 8 - *Mixed or Multiple ethnic groups-Any other Mixed or Multiple ethnic background*; 9 - *Asian or Asian British-Indian*; 10 - *Asian or Asian British-Pakistani*; 11 - *Asian or Asian British-Bangladeshi*; 12 - *Asian or Asian British-Chinese*; 13 - *Asian or Asian British-Any other Asian background*; 14 - *Black, African, Caribbean or Black British – African*; 15 - *Black, African, Caribbean or Black British – Caribbean*; 16 - *Black, African, Caribbean or Black British - Any other Black, African or Caribbean Background*; 17 - *Other ethnic group – Arab*; 18 - *Other ethnic group - Any other ethnic group*; 19 - *Prefer Not to Say*), Fitzpatrick skin type scale (1 – *Pale white skin - Very sensitive - Always burns, never tans*; 2 – *White skin - Very sensitive - Usually burns; tans with difficulty*; 3 – *Light brown skin - Sensitive - Burns moderately, tans gradually*; 4 – *Moderate brown skin - Moderately sensitive - Rarely burns, tans well*; 5 – *Dark brown skin -Minimally sensitive - Very rarely or never burns*; 6 – *Deeply pigmented dark brown to black skin - Minimally sensitive - Never burns*), burn history (cause of the burn injury (1 – *Thermal flame*; 2 – *Thermal flash*; 3 – *Thermal contact*; 4 – *Thermal scald*; 5 – *Thermal radiation*; 6 – *Electrical*; 7 – *Chemical*; 8 – *Friction*; 9 – *Not Known*), total burn surface area (TBSA – percentage of the total area of the body that is burnt), time taken to heal ((a) calculated based on dates of burn and healing; (b) reported by clinician), the depth of burn (1 – *superficial*; 2 – *superficial dermal* (*superficial partial thickness*); 3 – *deep dermal* (*deep partial thickness*); 4 – *full thickness*; 5 – *mixed*; 6 – *unknown depth*), area/size of scar (in cm^2), anatomical location for the scar selected for inclusion to the study (one of three regions: 1 – *head/neck*; 2 – *torso*; 3 – *limbs*). The ethnicity categories may be combined if the counts are small.

Summary of these characteristics will enable us to assess the external validity of the trial and whether the trial arms were comparable at baseline (although no significance tests will be undertaken on these data).

9.2 Comparison of outcomes between trial arms

Participants will be analysed in the trial arm they were randomised to (i.e., using an “intention-to-treat” analysis) irrespective of the treatment they actually received. With the consent of the participants, we will attempt to collect complete data on everyone and use those data in the

analyses. The main findings will be based on complete case analyses. Sensitivity analyses will also be undertaken based on imputed datasets in which missing data are “filled in”.

This is a superiority trial and the hypothesis testing framework is specified so that a definitive conclusion can be obtained regarding the primary objective (primary outcome). The null hypothesis is that in the population of people with hypertrophic burn scaring, there is no difference in the mean patient-reported POSAS score at 6 months follow-up between those allocated to pulsed dye laser treatment with standard care (intervention arm) and those allocated standard care alone (control arm). Whilst hypothesis testing will also be used to address the secondary objectives, the results of these will only be interpreted as exploratory and may be used to inform future research or hypothesis generation.

All hypothesis testing will be undertaken at the (2-sided) 5% level of significance. The results from all comparisons between the trial arms will be reported with 95% confidence intervals and p values. No adjustments will be made for multiple comparisons. Secondary outcomes will be treated as exploratory.

For continuous outcomes (POSAS, erythema index, melanin index and the CARe burn scales), multiple linear regression will be used to compare mean scores between the intervention and control arms. The non-parametric bootstrap will be used in sensitivity analyses to ensure that the 95% confidence intervals for the mean difference are robust to non-Normality in the outcomes.

For the ordinal outcome, patient’s perception of the improvement in scar quality, we will assess whether the proportional odds assumption holds by using the Brant test. If the assumption holds we will use ordinal regression based on the proportion odds model to compare the measure between the trial arms. An odds ratio will be reported to quantify the intervention effect. If there is evidence that the proportional odds assumption does not hold, the outcome will be dichotomised (into “scar quality is better” versus “scar quality is the same or worse”) and logistic regression will be used to compare the binary outcome between the trial arms, reporting the odds ratio.

Crude (unadjusted) analyses of the intervention will be undertaken as well as analyses that are adjusted for potential prognostic factors (study site, burn scar location (head/neck versus torso versus limbs), age and the outcome score at baseline). The adjusted analyses will be reported as the main analyses. Recognising that small numbers of participants randomised within a site may adversely affect the model, we will try and ensure that each site has a minimum of 10 participants. Any site for which this doesn’t occur will be combined, for analysis purposes, with the site that is closest in distance.

As there are separate versions of the CARe burn scales for those aged 18+ (Adult version) and those aged 16/17 (Young Person version), analyses of this outcome will be undertaken separately for each age group. It is likely that there will be insufficient numbers to analyse the outcome for those aged under 18 years.

The frequency of adverse events (including blisters, scabbing, excessive oedema, excessive pigmentation change, excessive pain, worsening of the scar) and serious adverse events will be summarised overall and by trial arm status, using numbers and percentages.

9.3 Sub-group analyses

The effectiveness of the intervention may vary across different subgroups of patients, specifically scar location (1 – *head/neck*; 2 – *torso*; 3 – *limbs*) and age group (1 – 16 to 24; 2 – 25 to 64; 3 – 65+). For the primary outcome only (patient-reported Patient and Observer Scar Scale (POSAS) score at 6 months), tests of interaction will be undertaken to explore whether the intervention effect differs across subgroups defined by scar location and age. Statistical power for detecting these interaction effects will be reduced compared to analyses for detecting main effects. Furthermore, as the analyses are only exploratory rather than hypothesis driven any statistically significant p values will be interpreted with appropriate caution and will need to be corroborated in future research.

9.4 Additional analyses:

Some additional analyses (including sensitivity analyses) on the primary outcome will also be conducted:

(a) It is acknowledged that some participants (in either group) will receive treatment for more than one scar. In a supplementary sensitivity analysis these participants will be excluded from the analysis to see if the results change.

(b) While it is hoped that appointments will happen within 2 weeks of enrolment/scheduled follow-up, this may not be possible for some participants. A sensitivity analysis will be undertaken in which only participants that provide data within 2 weeks of scheduled follow-up are included in the analysis.

(c) The main analysis will analyse participants in the group to which they were randomised ("intention to treat" approach) regardless of whether they received the allocated treatment.

- As a supplementary analysis a per protocol analysis will also be undertaken where patients are only included in the analysis if they received the treatment they were allocated to. Patients that did not receive the allocated treatment for the full 6 months will be excluded. If patients crossed over from one treatment to the other during the 6 month period they will also be excluded from the per protocol analysis.

- Complier Average Causal Effect (CACE) analyses will be undertaken to quantify the effect of the intervention in the sub-population of people that comply with the intervention. Compliers will be defined as those who complete all 3 of the laser sessions. The CACE estimate will be obtained by undertaking two-stage least squares regression. The adjustment factors used in the main analysis (site, scar location, the baseline value of the patient-reported POSAS score and age) will be used as predictors in both stages of the CACE analysis.

(e) The robustness of the results to missing data will be assessed by undertaking analyses where multiple imputation is used to “fill in” missing values, making the assumption that the data are missing at random. Data will be imputed using the chained equations (fully conditional specification method). The multiple imputation model will include: all study outcomes across all study waves; site; scar location; age; number of laser sessions completed; trial arm status; total cost at the baseline and 6 month follow-up; utility scores at baseline and 6 month follow-up.

(f) The correlation between the erythema score and the patient-reported POSAS measure will be calculated at each wave to validate the use of the POSAS measure.

10 Economic approach/overview

10.1 Economic evaluation aims

The aim of the economic evaluation is to assess whether pulsed dye laser (PDL) treatment with standard care compared to standard care alone is a cost-effective intervention.

10.2 Economic evaluation objectives

The primary objective is to estimate the cost-effectiveness of PDL with standard care compared to usual care for people for individuals with burn scar injuries, using a within trial cost-utility analysis over a 6-month time horizon.

10.3 Overview of economic analysis

The within-trial economic evaluation will use individual patient data from the ELABs trial. The primary analysis will take the form of a cost-utility analysis (CUA). Using trial data, the results will be summarised using incremental cost-effectiveness ratios, which will be calculated by generating a ratio of the difference in mean costs and mean effects (QALYs).

10.4 Perspective

The primary economic analysis will be from an NHS/PSS perspective. An exploratory analysis will take a societal perspective, to include costs to patients.

10.5 Time horizon

The primary economic analysis will compare the costs and consequences between baseline data collection and 6-month follow-up.

11 Economic data collection and management

11.1 Identification of resources

The following items of resource use will be measured:

- Health and social care service resource use (e.g. GP visits, A&E attendances)
- Treatment costs (e.g. laser appointment)
- Personal expenditure on burn injury care

11.2 Measurement of resource use data

Resource use will be measured using participant questionnaires (primary and community care, participant costs) and data collected in case report forms (secondary care).

11.3 Valuation of resource use

All resource use will be valued in monetary terms using appropriate UK unit costs or participant valuations estimated at the time of analysis (2022-2023).

NHS reference costs will be employed to value hospital resource use (e.g A&E visits and outpatient attendances). Primary care and social care costs (e.g. GP visits) will be valued using PSSRU Unit Costs of Health and Social Care. Medication costs will be taken from the British National Formulary (BNF) [3]. Staff time will be costed using Agenda for Change pay scales [4]. If costs cannot be obtained from the aforementioned sources, the principal investigator will be consulted to develop appropriate estimates.

11.4 Identification and Measurement of outcomes

The primary economic outcome will be the QALY, derived from utility scores, obtained using the SF-6D algorithm [5]. The model used will be the scoring algorithm recommended at the time of analysis [6].

A cost-effectiveness analysis (CEA) may be undertaken using alternative outcome measures (e.g. POSAS score) if the TMG believes it to be appropriate (e.g. there is a clear benefit using the primary outcome that is not reflected in the economic outcome or data availability). This analysis would be considered exploratory.

11.5 Valuation of outcomes

The SF-6D algorithm (3) will be used to generate utilities at baseline and 6 months. QALYs will be generated using the area under the curve approach and adjusted for baseline utility [7].

12 Economic data analysis

12.1 Analysis population

Analysis will be conducted as intention-to-treat, as per the analysis of the primary effectiveness outcome.

12.2 Discount rates for costs and outcomes

As the analysis will have a 6-month time horizon, no discounting will be conducted.

12.3 Cost-effectiveness thresholds

The estimated mean QALYs and costs associated with the intervention and control will be combined with a feasible range of values for decision makers' willingness to pay (ʎ), to obtain distribution of net benefits at different levels of ʎ. The primary economic analysis will use a cost-effectiveness threshold of £20,000 per QALY.

12.4 Statistical decision rules

We will compare costs, QALYs and incremental net benefits with accompanying 95% confidence intervals when making inferences about differences between trial arms in costs and outcomes.

12.5 Analysis of resource use

The differences in resource use between the intervention and control arms will be presented but not compared statistically.

12.6 Analysis of costs

Differences in overall mean costs between the arms will be analysed using appropriate models based on the distribution of data. Covariates will include age, gender and stratification variables included in the primary outcome analysis (scar location and study site).

12.7 Analysis of outcomes

The primary economic evaluation outcome will be the QALY. Utility scores will be generated from responses to the SF-12, using the SF-6D algorithm [5] which uses UK tariff values. These will be used to form QALYs over 6 months, adjusting for baseline quality of life. An area under the curve approach will be used. Data will be analysed using appropriate models. Covariates will include age, gender and stratification variables included in the primary outcome analysis (scar location and study site).

12.8 Data cleaning

Data will be checked to identify any errors, outliers or misspellings. Free text responses to resource use questions will be checked against CRFs if necessary. Corrections will be documented in the STATA code.

12.9 Missing data

Missing data will be assumed to be missing at random. The main analysis will use a complete case approach. Sensitivity analyses based on imputed datasets will also be undertaken. The same set of imputations that are used for the analyses of clinical outcomes will be used for the cost effectiveness analyses (see Section 9.4).

12.10 Analysis of cost-effectiveness

In the primary analysis, cost and QALY data will be combined to calculate an incremental cost-effectiveness ratio (ICER) and net monetary benefit (NMB) statistic from the NHS and PSS perspective. The NHS and PSS perspective will comprise of all costs except those incurred by the participants (e.g. additional burns dressings purchased).

12.11 Sampling uncertainty

The non-parametric bootstrap approach will be used to determine the level of sampling uncertainly surrounding the mean ICER. 5,000 bootstrap replications of incremental costs and benefits will be produced, which will be plotted on cost-effectiveness planes and/or a cost-effectiveness acceptability curve (the latter if appropriate).

12.12 Subgroup analysis

Subgroup analysis may be carried out in order to inform policy-makers' decision-making with respect to the targeting of the intervention. Such subgroup analyses (for instance, looking at intervention effects in different groups) will be decided with the team in advance of data lock and may include:

- Age, gender or other social and personal circumstances which the TMG indicate are important
- Scar location
- Time to treatment from burn injury (binary outcome informed by the TMG)

Subgroup analysis will be considered exploratory.

12.13 Sensitivity analysis

Sensitivity analysis will explore uncertainties surrounding key parameters in the economic evaluation. This analysis will be determined prior to data lock.

13 References

1. M. B. A. van der Wal, P. P. van Zuijlen, P. van de Ven, and E. Middelkoop, “Topical silicone gel versus placebo in promoting the maturation of burn scars: a randomized controlled trial,” Plast Reconstr Surg, vol. 126, no. 2, pp. 524–531, Aug. 2010, doi: 10.1097/PRS.0b013e3181e09559

2. N. Moiemen et al., “Pressure garment to prevent abnormal scarring after burn injury in adults and children: the PEGASUS feasibility RCT and mixed-methods study,” Health Technol. Assess., vol. 22, no. 36, pp. 1–162, Jun. 2018, doi: 10.3310/hta22360

3. BMJ. British National Formulary.

4. NHS Employers. Agenda for Change Pay Scales 2022 [Available from: https://www.nhsemployers.org/articles/pay-scale-material-202223.

5. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics. 2002;21(2):271-92.

6. University of Sheffield. SF-6D [Available from: https://licensing.sheffield.ac.uk/product/SF-6D.

7. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics. 2005;14(5):487-96