STATISTICAL ANALYSIS PLAN

Clinical Trial

A multicenter, multinational, randomized, controlled, open label study, performed in children with thermal burns, to evaluate the efficacy and safety of NexoBrid as compared to standard of care (SoC) - (CIDS)

Sponsor:
MediWound, Ltd.
42 Hayarkon Street
North Industrial Area
Yavne, Israel 8122745

Trial protocol code: MW2012-01-01
EudraCT number: 2014-003066-24
ClinicalTrials.gov Identifier: NCT02278718
IND No.: 065448

Version Control Page

Version:	Version Date:	Prepared / Updated by:	Protocol Version:	Comments:
01	07-Jul-2020	Dr. Martin Scharpenberg Competence Center for Clinical Trials Bremen - Biometry	Version 9.01, 29-Jul-2019	
02	31-May-2021	Larissa Fertik, Bioforum Biostatistics Department	Version 10.01, 16-Nov-2020	Updates following FDA recommendations, protocol ammendment, further clarifications and specifications

1. Table of Contents

2. Introduction

2.1. Background

Eschar removal, “Debridement”, is the first stage of the comprehensive wound care process. This process of debridement constitutes a well-recognized and critical part of burn wound care. Timely and rapid debridement of eschar is essential to initiate the wound healing process and prevent further complications. The dead tissue, if not removed, often becomes heavily contaminated in 2 to 3 days and is the source of local and/or systemic infection or sepsis. The local inflammation and infection destroy healthy surrounding tissues and extends the original damage. To prevent these complications, it is imperative to evaluate the burn and remove all the offending eschar at the earliest possible opportunity. Only when the eschar is removed and the wound bed revealed, can the true damage be observed and assessed, and the appropriate wound closure modality prescribed [1] [2] [3] [4] [5] [6].

Removal of eschar may be accomplished by surgery or by non-surgical means. The choice of debridement method (surgical or non-surgical) depends on many variables such as the burn depth, anatomical site and size (expressed as % TBSA of the burn wound), the patient’s general condition, available donor site for autograft harvesting as well as the availability of surgical facilities (operating rooms) and staff.

The current SoC in the US relies primarily on surgical tangential excision to remove the offending eschar through use of sharp instruments such as scalpels and dermatomes [7] [8]. Non-surgical SoC are all based on maceration (in situ) of the eschar. They include collagenase ointment (Santyl), antimicrobial agents such as silver sulfadiazine, or various hydrogels.

Care of Pediatric Burn Patients

There is an unmet need for treatment modalities that can facilitate the care of children with burns. Patients below the age of 20 years, account for 31.9% of the total number of burn injuries in the USA [9]. Burns in pediatric patients exhibit the same pathophysiological responses as in adults but with typical pediatric characteristics. These characteristics include differences of relative sizes between areas, anatomy, physiology, emotional and psychological development, which influence the child’s response to injuries including burns.

NexoBrid is a non-surgical debridement product developed for treatment of second and third degree severe burns with improved safety and efficacy characteristics.

The development of NexoBrid as a product for enzymatic debridement of burns started several decades ago.

The safety and efficacy of NexoBrid (for burn eschar debridement) has been extensively investigated in 7 clinical studies [10], including two phase 3 pivotal trials, MW 2004-11-02 [11] and MW2010-03-02. Efficacy data generated during the NexoBrid clinical program suggest that the clinical benefits of NexoBrid address the unmet medical needs for both improved routine burn care as well as care in the event of a mass casualty incident (MCI), as demonstrated by the product's specific attributes:

1. An effective topical eschar removal agent (comparable to SoC)
2. Enables significantly earlier eschar removal compared to SoC, allowing direct visual assessment of wound depth and assisting in the decision on the proper subsequent wound closure strategy
3. Reduces the surgical burden (extent of excision and grafting); less traumatic than SoC
4. Reduces blood loss associated with eschar removal
5. Product use is not dependent on operating room facilities or surgeons (i.e., it can be applied at the patient's bedside) and therefore can alleviate bottlenecks in treatment in an MCI situation, as demonstrated in the MCI in Romania [12]
6. Enables avoiding of escharotomy surgery and its sequelae: Early application of NexoBrid may reduce the elevated interstitial pressure in circumferential burns and alleviate burn-induced compartment syndrome, thus reducing the need for eschar incision [13]

Indirect beneficial effects:

1. Reduction of incidence and extent of surgery (excision and grafting) by NexoBrid eschar removal treatment.
2. Autografting can be reduced by selective debridement with increased preservation of uninjured dermis and potential for spontaneous epithelialization. In choosing the approach of epithelialization over preserved dermis, autografting is reserved to full thickness wounds. In addition to the benefit of reduced surgery, decreased autografting reduces sacrifice of the subject's healthy uninjured skin and accompanying morbidity.

3. Blood loss is minimized by reduced excisional and autografting surgery.

2.2. Study Objective

The objective of this study is to evaluate the safety and effectiveness of NexoBrid in hospitalized pediatric patients with DPT and FT thermal burns of 1-30% TBSA and to compare NexoBrid to SoC. The specific study objectives are:

1. To demonstrate enzymatic eschar removal efficacy of NexoBrid by providing earlier, complete eschar removal,
2. To demonstrate enzymatic eschar removal efficacy of NexoBrid by reducing patients' surgical burden, eschar removal related blood loss and resulting in non-inferior final outcomes of cosmesis and function as compared to SoC,
3. To assess the safety of NexoBrid compared to SoC.

2.3. References to Study Protocol

This Statistical Analysis Plan refers to the study protocol version10.01 of 16 November 2020. The trial protocol code is MW2012-01-01.

2.4. Amendments

The study was initiated under the trial protocol version 4. Since then, there have been six amendments (to versions 5, 6, 7.02, 8, 9.01 and 10.01)[1].

2.5. Global Statistical Analysis

The CIDS MW2012-01-01 protocol is conducted as part of a Pediatric Investigational Plan in Europe and therefore changes in key binding elements in the Pediatric Investigational Plan are subject to the EMA's pediatric committee (PDCO) prior approval. The PDCO and the FDA had different requirements for the statistical analysis and therefore, as agreed with the FDA, the

[1] Only versions 7.02 and 9.01 were submitted under the IND

study conduct will be the same for the EU and US, however the SAPs will be different (Type B meeting minutes, May 2017). The SAPs will include the same analyses, but in a different order, according to each agency's requirements. This document is the non-EU version of the SAP.

3. Study Administrative Structure

3.1. Sponsor

Sponsor: MediWound, Ltd
42 Hayarkon Street
North Industrial Area
Yavne, Israel 8122745

Represented by: Prof. Lior Rosenberg
Chief Medical Officer
MediWound Ltd.

Keren David Zarbiv
VP Clinical Affairs
MediWound Ltd.

Nimrod Leuw
VP of Quality
MediWound Ltd.

3.2. Study Conduct (Data Management)

Study Conduct (CRO): Dr. Marco Schwarzer
GCP-Service International Ltd. & Co. KG
Anne-Conway-Str. 2
28359 Bremen

3.3. Statistics

Statisticians: Professor (emer.) Paul D. Feigin
Director of the Israel Data Science Initiative (IDSI)
Head of the Technion Statistics Laboratory
Faculty of Industrial Engineering and Management
Technion – Israel Institute of Technology
Haifa 3200003, ISRAEL

Larissa Fertik
Senior Biostatistician
Bioforum Biostatistics Department
3 Golda Meir St., Weizmann Science Park
Ness Ziona 7414001, Israel

4. Signatures

Keren David Zarbiv

Sponsor Representative
MediWound Ltd.

Keren David — Signature

01 June 2021 — Date

Professor (emer.) Paul D. Feigin

Director of the Israel Data Science Initiative (IDSI)
Head of the Technion Statistics Laboratory,
Faculty of Industrial Engineering and Management
Technion – Israel Institute of Technology

Signature

31-05-2021

Date

Larissa Fertik

Senior Biostatistician
Biostatistics Department,
Bioforum

Larissa Fertik Digitally signed by Larissa Fertik
Date: 2021.05.31 13:35:06 +03'00'

Signature

Date

5. List of Abbreviations

abbreviation	meaning
ABA	American Burn Association
ABL	Actual Blood loss
AE	Adverse Event
AFT	Accelerated Failure Time
ATC	Anatomical Therapeutical Chemical Classification
BICS	Burn-induced Compartment Syndrome
BMI	Body Mass Index
BSHS-B	Burn Specific Health Scale - Brief
CS	Clinically Significant
DDD	Defined Daily Dose
DPT	Deep Partial Thickness
DSMB	Data Safety Monitoring Board
EAP	Efficacy Assessment Period
EAR	Extremities at Risk
EBV	Estimated Blood Volume
ECG	Electrocardiogram
eCRF	electronic Case Report Form
EMA	European Medicines Agency
ER	Eschar Removal
FAS	Full Analysis Set
FDA	Food and Drug Agency
FPS-R	Faces Pain Scale - Revised

abbreviation	meaning
FT	Full Thickness
FU	Follow up
Hb	Hemoglobin
HD	Hospital Discharge
HEENT	Head, Eyes, Ears, Nose and Throat
HR	Hazard Ratio
ICF	Informed Consent Form
INR	International Normalized Ratio
ITT	Intention to Treat
KKSB	Competence Center for Clinical Trials Bremen
LEFS	Lower Extremity Function Scale
LOCF	Last Observation Carried Forward
LTFU24	24 months Long-Term follow up period
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
MVSS	Modified Vancouver Scar Scale
NCS	Not Clinically Significant
pCS	Post-treatment potentially clinically significant
PDF	Portable Document Format
POSAS	Patient Observer Scar Assessment Scale
PP	Per Protocol

abbreviation	meaning
PRBC	Packed Red Blood Cells
PTT	Prothrombin Time
QoL	Quality of Life
ROM	Range of Motion
RTF	Rich Text Format
SAE	Serious Adverse Event
SoC	Standard of Care
SOC	System Organ Class
SOP	Standard Operating Procedure
SPT	Superficial Partial Thickness
SSD	Silver Sulphadiazine
STFU12	12 months Short-Term Follow up period
SUSAR	Suspected Unexpected Serious Adverse Reaction
TBSA	Total Body Surface Area
TW	Target Wound
VAS	Visual Analogue Scale
WBC	White Blood Cell (Count)

6. General Definitions

6.1. General Definitions of Medical Terms

Full analysis set (FAS): The full analysis set includes all patients who are randomized into the trial. The analysis is focused on the planned treatment.

Complete eschar removal: is achieved at the end of the eschar removal treatment phase. The Eschar Removal assessment should be more or equal to 95% and will be clearly marked as the end of the eschar removal process and the initiation of specific treatment aimed to close the viable debrided bed by grafting or epithelialization [14].

Complete wound closure: defined as % of wound area epithelialized and/or closed by graft is >95% without drainage or dressing requirements confirmed at two consecutive study visits, 2 weeks apart.

Efficacy assessment period: defined as the time period up to the 12 weeks follow-up post wound closure confirmation.

6.2. Hardware and Software Used

The statistical analysis will be performed using SAS® version 9.4.

6.3. General format and content of tables, figures and patient data listings.

Tables, figures and listings of subjects will be created in SAS® version 9.4 and be converted in rich text format (RTF) or portable document format (PDF).

6.4. Quality Control

The statistical analysis must follow the statistical analysis plan.

The planned method of QC for all analyses is double programming.

6.5. General Traceability

All generated outputs will include headers and footnotes containing the following information:

- Page number (Page xx of xx)
- Name of generating program
- Output run time
- Data base extraction date
- Output file name
- Source data set(s)

7. Details of Trial Design

7.1. Trial Schedule

Table 1: Treatment flow chart – Study Schedule (all arms)

* Weekly FU assessments will be performed 7±2 days post eschar removal and, thereafter on a weekly basis until wound closure. Long term FU will be performed 6 and 12 weeks, 6,12,18 and 24 months after last wound closure confirmation.

#-If applicable

Parameter	Screening & Baseline	Pre-treatment - Within 1 h pre treatment – Both Arms	Treatment		Daily Assessments	Wound Management & Re-admission – surgical procedures	Weekly FU*	Hospital Discharge or premature discontinuation visit	FU (weeks 6, 12 months 6, 12, 18, 24)
			NexoBrid (S=Soaking R=Removal)	SoC (including additional procedures post NexoBrid)					
Signing informed consent	X								
Medical history	X								
Burn Description	X								
ConMeds & AEs	X	X	X	X	X	X	X	X	X
Physical examination	X							X	
Demographic data	X								
Vital signs	X				X				
Pain assessment	X	X	4h post R	Pre and 4h post	X				
Photograph of TWs	X	X	Post S	Pre & Post		Pre & post	X		X
Local Lab - HbA1C, βhCG	X								
Central Lab: Urianalysis		X	4h post R	Pre & 4h Post					

Parameter	Screening & Baseline	Pre-treatment - Within 1 h pre treatment – Both Arms	Treatment		Daily Assessments	Wound Management & Re-admission – surgical procedures	Weekly FU*	Hospital Discharge or premature discontinuation visit	FU (weeks 6, 12 months 6, 12, 18, 24)
			NexoBrid (S=Soaking R=Removal)	SoC (including additional procedures post NexoBrid)					
Central Lab: Biochemistry[2]		X	4h post R	Pre & 4h Post					
Central Lab: Hematology		X	4h post R	Pre & 4h Post		Pre & post			
Randomization	X								
Local Lab: Wound culture		X	Post S	Pre & Post					
Local Lab: Blood culture[3]		X	4h post R	4h Post					
PK[4]		X	X		X				
Immunogenicity[5]		X					X		X
Local Lab: PTT & INR		X	4h post R	4 h post – first treatment only					
Coverage				Start and end date, type of primary dressing		Start and end date, type of primary dressing			
Pressure Measurement (for Circumferential		#	#	#1 h pre & Post					

Parameter	Screening & Baseline	Pre-treatment - Within 1 h pre treatment – Both Arms	Treatment		Daily Assessments	Wound Management & Re-admission – surgical procedures	Weekly FU*	Hospital Discharge or premature discontinuation visit	FU (weeks 6, 12 months 6, 12, 18, 24)
			NexoBrid (S=Soaking R=Removal)	SoC (including additional procedures post NexoBrid)					
extremity wounds monitoring)									
Cosmesis, function and quality of life[6]									X
Record of scar modulating/reconstructive procedures									X
Range of Motion[7]									X
Blood transfusion	#		#	#	#	#			#
Histological biopsy	#			#	#	#	#	#	#

1 Physical examination will include height and weight. Height will be taken only once, at screening.

2 For patients weighing less than 18kg, serum chemistry will be collected at baseline only.

3 Blood culture is to be tested only if this is clinically indicated:

- in patients weighing less than 18kg,
- and/or in patients requiring two NexoBrid applications (either because of failure of the first treatment session or because >15%TBSA)

4 PK samples should be taken as detailed in Appendix 10 of the study protocol

5 Immunogenicity will be performed only for NexoBrid patients weighing at screening over 18kg at baseline, weeks 4, 8 and months 6 and 24 following wounds closure. In case that weekly visit 8 is not performed, this assessment should be taken on the relevant weekly/first follow up visit as close as possible to 8 weeks.

6 Cosmesis will be evaluated in donor site scars as well

7 Range of motion measurement of injured joints will be measured, if relevant. The measurement will be performed on the corresponding non injured joint as well, for comparison.

7.2. Criteria for Evaluation

7.2.1. Demographic Data and Other Baseline Characteristics

The following demographic data will be documented at screening/baseline:

- Date of birth
- Age Group (0 to 23 months, 24 months to 3 years, 4 to 11 years, 12 to 18 years)
- Gender
- Ethnicity
- Race
- Serum Pregnancy Test results (βHCG)

Furthermore, at screening/baseline the inclusion and exclusion criteria will be collected and the following parameters will be documented.

Burn history / description:

- Date and time of injury
- Etiology of injury (Fire/Flame, Scald, Contact)
- Place of injury (Outdoors, Indoors, Car, Other)
- Date and time of admission

General wound description:

- Wound number
- Anatomical location
- %TBSA of 2° SPT Burns (clinical assessment)
- %TBSA of 2° DPT Burns (clinical assessment)
- %TBSA of 3° FT Burns (clinical assessment)
- Information whether the respective wound will be designated as TW
- Information whether a wounded extremity is circumferential
- Information whether escharotomy had been performed prior to randomization
- Eschar description
- Information whether cleansing was performed

- Information whether all superficial keratin (blister) were removed from wound

Target wound identification (data collected per wound):

- Target wound number
- Anatomical location
- %TBSA of 2° SPT Burns (clinical assessment)
- %TBSA of 2° DPT Burns (clinical assessment)
- %TBSA of 3° FT Burns (clinical assessment)
- Information whether the wound was treated prior to enrollment

Information needed for randomization and treatment:

- Total wounds TBSA % per patient (clinical assessment)
- Overall total area of FT burns (clinical assessment)
- TWs FT total area (% TBSA, out of the total burned area of the patient, clinical assessment)
- Randomization result (treatment)
- Date of randomization

Physical examination:

- Height (cm or inch)
- Weight (kg or lb)

Furthermore, for each of the following body systems 4 categories are documented (normal, abnormal-NCS [not clinically significant], abnormal-CS [clinically significant], not done). In case of abnormal findings, they have to be documented. If abnormality is present for another body system, this has to be documented, too. Reasons for the category "not done" have to be documented.

- Head, eyes, ears, nose and throat (HEENT)
- Respiratory system
- Cardiovascular system
- Abdomen
- Musculoskeletal / Extremities
- Neurological
- Lymph Nodes

Vital signs:

- Temperature (incl. unit and method of measurement)
- Heart rate (beats/min)
- Respiratory rate (breaths/min)
- Systolic and diastolic blood pressure (mmHg)
- Assessment of vital signs (normal, abnormal-NCS, abnormal-CS)

Pain assessment:

- Pain measured via FPS-R
- Assessment of pain (normal, abnormal-NCS, abnormal-CS)

Local laboratory:

- Information whether patient is diabetic
- HbA1c (%)

Information on infections as well as the medical history will be recorded.

7.2.2. Efficacy Evaluation

7.2.2.1. Primary Endpoint: Time to Complete Eschar Removal

The following primary endpoint will be evaluated at a 'per patient' level i.e., on the full set of target wounds treated for each patient (please refer to protocol Section 2.1 for target wound definition) in accordance with the study protocol:

Earlier eschar removal (in days): Demonstrate superiority of NexoBrid over SoC for eschar removal (see definition in Section 6.1) as measured by a survival analysis of incidence of complete eschar removal for all TWs as a function of time in days from randomization. Eschar removal will be measured at the end of the debridement starting from randomization date, and marked as yes/no in the eCRF.

7.2.2.2. Secondary Endpoints

The following secondary endpoints will be evaluated in this study and compared between NexoBrid and SoC:

Incidence of Surgical Excision

Reduction of surgical need for excisional eschar removal is measured by an analysis of incidence (yes/no response) of surgical eschar removal (tangential/ minor/ avulsion/ Versajet and/or dermabrasion excision). The time frame for this incidence rate is until complete eschar removal has been achieved.

Blood Loss

Blood loss incurred during the eschar removal procedures will be evaluated. Actual blood loss will be measured by considering the estimated blood loss, changes in Hematocrit during the eschar removal procedures and the amount of units of blood transfused. Blood loss will be calculated using the following formula [15]:

$$Blood\ Loss\ (mL) = \frac{RBCV_{before} + (Transfusion\ RBCV) - RBCV_{after}}{Hct_{after}[\%] * 0.01}$$

RBCV (mL) = Red Blood Cell Volume (mL) = Body Weight (kg) * 80 (mL/kg) * (Hematocrit [%] * 0.01)

Transfusion RBCV (mL) = [Total Volume (mL) of Whole Blood Transfused * 0.3] + [Total Volume (mL) of Packed Red Blood Cells Transfused * 0.8]

The blood loss will be computed over the whole eschar removal period, using the formula above, i.e. the eschar removal process will be considered as one continuous procedure. Therefore, only the Hematocrit value before the first eschar removal procedure and after the last procedure will be used and all blood transfusions performed between these measurements will be included.

Additionally, the blood loss will be summed over all procedures carried out to remove eschar, with the hematocrit measured immediately before each procedure and 4 hours after its completion. The amount of blood transfused will be summed over all transfusions carried out during the debridement procedure and up to 24 hours from the end time of the debridement procedure.

Analyses with both blood loss values will be conducted. However, some missing values are expected. The analysis with less missing values will be regarded as more reliable and is therefore considered the main analysis. The other analysis is considered supportive.

Area of Autograft in DPT Wounds

Reduction in the need for Autograft in deep partial thickness wounds is measured by an analysis of percent area of deep partial thickness wounds autografted.

Incidence of Autograft in DPT Wounds

Reduction in the need for Autograft in deep partial thickness wounds is also measured by an analysis of incidence of Autograft procedures.

7.2.3. Safety Evaluation

7.2.3.1. Time to Reach Complete Wound Closure

Time to reach complete wound closure will be assessed in days, starting from randomization date.

7.2.3.2. Cosmesis and Function at 12 Months from Wound Closure

Cosmesis and Function assessment will be performed based on the Modified Vancouver Scar Scale (MVSS) at 6 and 12 weeks and at 6, 12, 18 and 24 months after the wound closure confirmation visit by a burn wound specialist who is masked to randomized wound treatment and to the other outcome assessments of the same wound. Table 2 gives an overview of the MVSS.

Table 2: Modified Vancouver Scar Scale (MVSS)

Pigmentation	***Pliability***	***Height***	***Vascularity***	***Pain***	***Pruritus***
(0) Normal	(0) Normal	(0) Flat	(0) Normal	(0) None	(0) None
(1) Hypopigmentation	(1) Supple-flexible with minimal resistance	(1) <2mm	(1) Pink	(1) Occasional	(1) Occasional
(2) Mixed	(2) Yielding- giving way to pressure	(2) 2-5 mm	(2) Red	(2) Requiring medication	(2) Requiring medication
(3) Hyperpigmentation	(3) Firm- inflexible, not easily moved, resistance to manual pressure	(3) >5mm	(3) Purple		
	(4) Banding-rope-like tissue that blanches with extension of the scar				
	(5) Contracutre-perment shortening of scar, producing deformity or distortion				

7.2.3.3. Cosmesis and Function at 24 Months from Wound Closure

See section 7.2.3.2.

7.2.3.4. Adverse Events

For each AE the following information will be determined:

- AE description
- Type of AE (General, Local – TW related, Local – not TW related)
- onset date and time
- outcome date and time
- outcome (resolved, resolved with sequelae, not resolved, death)
- intensity (mild, moderate, severe)
- action taken regarding study drug (yes, no)
- relation to study drug (not related, remotely related, possibly related, probably related, related)
- treatment required (none, medication, surgery, hospitalization, other)
- seriousness (serious, not serious)

Treatment-emergent Adverse Events (TEAEs) are defined as AEs that start during or after first NexoBrid or SoC treatment. TEAEs, related TEAEs, and Serious TEAEs will be summarized by Medical Dictionary for Regulatory Activities (MedDRA) system organ class, preferred term and severity. Events where the relation is considered “related”, “possibly related”, or “probably related” will be considered as related and events where the relation is considered ‘not related’ or ‘remotely related’ will be considered “not related”. Events with missing causality will be considered as "related" and will be included in the summary of related TEAEs. Serious AEs resulting in death and AEs resulting in withdrawal from the study will be listed separately.

For randomized patients, AEs that occur between signing the informed consent document and receipt of study treatment will be provided in a listing.

Except for a summary table on event level, TEAEs will be reported on a per-patient basis. This means that even if a patient reported the same event repeatedly (i.e., events mapped to the same preferred term) during the assigned study period, the event will be counted only once in tabulations. In the latter case the event will be assigned the worst severity and the strongest relationship to the study drug. The earliest date will be regarded as start date of the event and

the latest date will be regarded as stop date of the event. All AEs, including those occurring more than once will be listed.

The same rules for counting, severity and relationship apply for PTs mapped to the same system organ class.

Notes:

- The MedDRA version will be denoted in the footnote of tables and listings.
- Listings will be sorted within patients by start date of the adverse event according to treatment group.

7.2.3.5. Laboratory Assessments

The following tests will be performed:

Serum chemistry, hematology and urinalysis testing will be performed at a central laboratory during NexoBrid eschar removal phase. Any additional procedures require hematology tests, as described in the protocol, will be performed using Central lab. Pregnancy testing, HbA1c, wound cultures, blood cultures and biopsies (if indicated) will be performed locally on-site.

The following tests will be performed in special central laboratories: Pharmacokinetics test, Immunogenicity evaluation.

Serum Chemistry:

Urea, Creatinine, Glucose, Total Bilirubin, Cholesterol, Triglycerides, Uric Acid, Total Protein, Albumin, Globulin, Calcium (Ca++), Phosphorous, Sodium (Na+), Potassium (K+), Chloride (Cl-), Alkaline Phosphatase, SGOT (ASAT), SGPT (ALAT), LDH.

Hematology:

Hemoglobin, Hematocrit, RBC, MCV, MCH, MCHC, Platelets, Leukocyte count (WBC) with differential count.

HbA1C[2] will be assessed at a local lab at Screening for diabetic patients as captured during the medical history.

[2] HbA_1C will not be collected for patients weighing less than 18kg at screening

PTT and INR will be tested at a local lab pre treatment and post treatment.

Pregnancy Testing:

A serum or urine pregnancy (βHCG) test will be performed at a local lab for women of child bearing potential.

Urinalysis:

pH, Ketone, Specific gravity, Protein, Glucose, Bilirubin, Microscopic examination.

Qualitative Wound Culture:

Test burn wound for microbial flora (swab).

Blood Culture[3]:

Test blood for microbial flora

7.2.3.6. Additional Safety Endpoints

Additional safety endpoints to be evaluated in this study are:

- Vital signs
- Pain assessment (using FPS-R and as reported as AEs)
- Volume of blood transfusion given during hospitalization
- Immunogenicity evaluation for NexoBrid patients
- Analgesia and anesthesia medications
- Days of exposure to systemic antibiotic drugs
- Rates of hospital readmission
- Change in INR/PTT and incidence of change to > upper limit of normal (after treatment)
- Change in blood glucose and incidence of change to above upper limit of normal (after treatment)
- Long-term functionality evaluation of the extremities using the 'Lower Extremity Functional Scale'

[3] Blood culture is to be tested only if this is clinically indicated in patients weighing less than 18kg, and/or only if it is clinically indicated in patients requiring two NexoBrid applications (either because of failure of the first treatment session or because >15%TBSA)

- 'QuickDASH' questionnaires
- 'Range Of Motion' measurements
- Long-term Quality of Life using EQ5D and Burn Outcome Questionnaire (BOQ) for a subset of patients[4]

7.2.4. Exploratory Analyses

- Incidence of surgical Escharotomy procedures on circumferential extremities target wounds,
- Incidence of reduction in interstitial/compartment pressure in circumferential extremity wounds,
- Reduction in surgical needs as measured by an analysis of % wound area surgically excised for eschar removal.
- Reduction in surgical need as measured by analysis of incidence of surgically harvested donor sites scars,
- Reduction in surgical need as measured by analysis of % area of surgically harvested donor site scars.
- Blood loss related to eschar removal assessed by changes in Hemoglobin incurred during the eschar removal procedures
- PK evaluation in NexoBrid patients
- Cosmesis evaluation (MVSS and POSAS) will be used to assess the quality of the donor site scars,
- Incidence of wound closure
- Additional Cosmesis and Function evaluation will be performed using POSAS measured at 12 and 24 months from wound closure.
- Incidence and extent of Autograft in DPT wounds
- Duration of hospitalization

[4] BOQ will be completed only for patient from study sites in the USA. BOQ_{11-18} will be completed by patients aged 11-18. BOQ_{0-5} will be completed by parents of burned children under the age of 5 years and BOQ_{5-18} will be completed by parents of burned children between the ages of 5 and 18

8. Handling of Protocol Violations

Before database lock and statistical analysis, possible protocol violations will be discussed by data management, sponsor, PI and his/her proxy and the statisticians and classified as minor or major violations. The list with all protocol violations will be prepared by the data management.

A protocol violation will be classified as major if

- The treatment differs from the planned treatment (randomization arm), or
- There was drug misuse (this may include application time deviated in 25% or more, deviation of more than 20% in the TBSA area treated in one session at a time, if relevant, error in drug preparation (e.g. 2g mixed in 50g), storage temperature below -2°C - 8°C or >24h at room temperature and used without any stability confirmation by sponsor), or
- treatment started more than 5 days from injury, or
- one of the following inclusion criteria is violated:
    - Stage 1: Males and females between 4 years to 18 years of age,
    - Stage 2 (upon DSMB review): Males and females between 1 year to 18 years of age,
    - Stage 3 (upon DSMB review): Males and females between 0 years to 18 years,
    - Thermal burns caused by fire/flame, scalds or contact,
    - Patient total burns area ≥ 1% DPT and / or FT,
    - Patient total burns area should be ≤ 30% TBSA; SPT, DPT and/or FT in depth,
    - Informed consent can be obtained within 84 hours of the burn injury
    - At least one wound (a continuous burn area) that is ≥1% TBSA (DPT and/or FT) (this minimal wound size should not include face, perineal or genital)
    - Wound that is potentially intended for surgical eschar removal
    - Wound's blisters can be removed/unroofed, as judged by the investigator
- one of the following exclusion criteria is met:
    - Patients with electrical or chemical burns
    - The following pre-enrollment dressings: a. Flammacerium, b. Silver Nitrate ($AgNO_3$)

- Patients with pre-enrolment wounds which are covered by eschar heavily saturated with iodine or by SSD pseudoeschar (e.g. pseudoeschar as a result of >12 h SSD treatment),
- Patients with pre-enrollment escharotomy
- Enrollment in any investigational drug trial within 4 weeks prior to screening

In all other cases protocol violations will be classified as minor.

9. Populations Analyzed

The following sample size calculations are taken from the protocol.

9.1. Sample Size

The following sample size calculations are based on the consideration of the primary endpoint of this study. For the secondary endpoints, the power, which can be achieved with the planned sample size, is determined. Updated calculations were performed to support the final sample size of 145 patients enrolled into the study (provided in 13.3 13.3Update to Study Sample Size Due to COVID-19 Pandemic).

Estimated effects of NexoBrid from study MW2004-11-02 were used to design the study with adequate statistical power using the following steps.

a) Sample size calculations were based on study MW2004-11-02. This study included both adults and children. Similar effects were observed among the children population and the adult one;
b) The effect of NexoBrid versus standard of care was estimated together with the standard error of the estimate;
c) The size of the effect to be detected in this study, MW2012-01-01, was taken as the previous estimate decreased by half its standard error, in order to use a conservative value.

9.1.1. Primary Endpoint: Time to Complete Eschar Removal

Estimated log hazard ratio from study MW2004-11-02 (SoC versus NexoBrid) = -1.39, using a Cox proportional hazards model with treatment as only covariate.

Standard error = 0.28

Target log hazard ratio = -1.39 + 0.28/2 = -1.25; target hazard ratio λ= 0.29.

The formula for the number of events D (successful complete eschar removals) is:

$$D = \frac{4\left(z_\alpha + z_\beta\right)^2}{(\ln \lambda)^2}$$

(Piantadosi, *Clinical Trials: A Methodologic Perspective*, p.169)

z_α = 1.96;

z_β = 1.28 for 90% power.

If the proportion of events in NexoBrid group is denoted by p_1 = 0.59

and the proportion of events in SoC group is denoted by p_2 = 0.70,

which are the observed proportions in the study MW2004-11-02, the number of patients required $n = \frac{2D}{p_1+p_2} = 42$ (to the nearest even number), 21 in each group.

In accordance with the European regulatory requests and planned age distribution, we believe that a sample size of 80 patients per arm will detect with a 90% power a difference on the primary endpoint, which is well within what would be expected, based on the results from the previous study.

Based on the above calculations, we plan a study with total sample size of 160 patients; 80 (NexoBrid) + 80 (SoC)[5].

9.1.2. Secondary Endpoint: Incidence of Surgical Excision

Estimated proportions having surgical excision from study MW2004-11-02: NexoBrid 0.42; SoC 0.80

Standard errors: NexoBrid 0.057; SoC 0.044

Difference = 0.38; Standard error of difference = 0.073

Anticipated difference: 0.38-0.073/2 ≈ 0.34

We therefore take the anticipated proportion with surgical excision to be: NexoBrid 0.44 and SoC 0.78.

Using Fisher's exact test, the power achieved with a total sample size of 160 patients (80 in each group) using a two-sided significance level of 5% is 99.1%.

[5] Updated calculations were performed to support the final sample size of 145 patients enrolled into the study (provided in 13.3 13.3Update to Study Sample Size Due to COVID-19 Pandemic)

9.1.3. Secondary Endpoint: Blood Loss

Estimated mean blood loss from study MW2004-11-02: NexoBrid 421.93 ml; SoC 1024.27 ml

Standard errors: NexoBrid 129.52; SoC 240.98

Difference: -602.33; Standard error of difference: 273.58

Anticipated difference: -602.33 + 273.58 /2 = -465.54

Using a two-sided two-sample t-test the power with a total sample size of 160 patients and a two-sided significance level of 5% is 58.8%.

9.1.4. Secondary Endpoint: Area of Autograft in DPT Wounds

The proportion of subjects with at least one DPT wound was 62.6% in the study MW2004-11-02. Therefore, if 160 patients are enrolled to the CIDS study, we expect 100 patients to have at least one DPT wound and be thereby eligible for this analysis.

Estimated difference in mean area autograft in DPT wounds from study MW2004-11-02 (NexoBrid vs. SoC) = 13.30-23.63 = -10.33.

Standard error = 6.02

Target difference Δ = -10.33 + 6.02/2 ≈ -7.32.

Using a two-sided two-sample t-test the power with a total sample size of 160 patients using a two-sided significance level of 5% is 23.1%.

9.1.5. Secondary Endpoint: Incidence of Autograft in DPT Wounds

Estimated log OR using GEE approach from study MW2004-11-02 (NexoBrid vs SoC): -0.676

Standard Error: 0.40

Anticipated log OR: -0.676+0.40/2=-0.476

Using the normal approximation to the Wald test statistics (see for example Shih (1997)) we obtain a power of 53.3% with a total sample size of 160 patients using a two-sided significance level of 5%.

9.2. Safety Set

The safety set includes all patients who received a treatment and analysis is focused on the treatment actually performed.

9.3. Full analysis Set (Intent-to-Treat population)

The full analysis set (FAS) includes all patients who were randomized into the trial. The analysis is focused on the planned treatment. The analysis of the ITT population (FAS) provides the primary and secondary efficacy and the comparative safety analyses.

9.4. Per Protocol Set

The per protocol (PP) set includes all subjects who do not have major protocol violations (cf. section 8). Analysis of the PP population is for supportive purposes. The PP analysis will include patients in the treatment group according to treatment received.

9.5. Enrolled population

The enrolled population consists of all subjects who passed through the screening processes and signed informed consent.

10. Data Handling

10.1. Handling of Missing Data, Outliers and Implausible Data

10.1.1. Missing Date/Time Variables

If time data is missing (e.g. for ICF or injury), i.e. a date variable is given but without the time, the time is set to 12 pm (middle of the day).

In case 12 weeks follow-up date is missing (needed for the definition of the efficacy assessment period), it will be imputed by adding 42 days to the 6 weeks follow-up date. In case 6 weeks follow-up date is missing, it will be imputed by adding 12 weeks (84 days) to the wound closure confirmation date. In case wound closure confirmation date is missing, it will be imputed by subtracting 90 days from the 6 months follow-up date. In case 6 months follow-up date is missing, it will be imputed by subtracting 270 days from the 12 months follow-up date. In case 12 months follow-up date is missing, it will be imputed by subtracting 450 days from the 18 months follow-up date. If the wound closure confirmation date and all further follow-up dates are missing, it will be imputed by the study termination date.

In case the 12 months follow-up date is missing (needed for the definition of the stages of the analyses, see Section 11), it will be imputed by adding 12 months to the wound closure confirmation date. If the wound closure confirmation date is missing, the 12 months follow-up date will be imputed by adding half the difference (in days, rounded to the next highest whole number) between the 6 months follow-up date and the 18 months follow-up date to the 6 months follow-up date. If the 6 months follow-up date is not missing (and the 18 months follow-up date is missing), the 12 months follow-up date will be imputed by adding 180 days to the 6 months follow-up date. If the 18 months follow-up date is not missing (and the 6 months follow-up date is missing), the 12 months follow-up date will be imputed by subtracting 180 days from the 18 months follow-up date. If the wound closure confirmation date and all further follow-up dates are missing, the 12 months follow-up date will be imputed by the study termination date.

In case the 24 months follow-up date is missing, it will be imputed by the study termination date.

In case the end date of an adverse event or a concomitant medication is missing, it will be imputed by the maximum of the end date of the analysis stage, respectively the study termination date.

A missing start date of an ongoing concomitant medication or adverse event will be imputed by the date of informed consent.

10.1.2. Primary Endpoint: Time to Complete Eschar Removal

The main analysis of the primary endpoint time to complete eschar removal will compute missing values as censored at the date of the last non-missing eschar removal assessment (typically last debridement procedure). If no eschar removal assessment after baseline is available, the subject will be censored with time 0.

In addition, three sensitivity analyses will be performed:

1. For the first sensitivity analysis, subjects with no available eschar removal assessment will be censored with Median Time, estimated using the Kaplan-Meier method based on subjects from both treatment groups having a non-missing value on this endpoint (including subjects censored at the date of the last non-missing eschar removal assessment).
2. For the second sensitivity analysis, subjects with no available eschar removal assessment will be censored as follows (a worst case scenario):
    - For NexoBrid – with the maximal time obtained for subjects randomized to NexoBrid group (including subjects censored at the date of the last non-missing eschar removal assessment)
    - For SoC – with the Median Time, estimated using the Kaplan-Meier method based on subjects from SoC group having a non-missing value on this endpoint (including subjects censored at the date of the last non-missing eschar removal assessment)
3. For the third sensitivity analysis, subjects with no available eschar removal assessment will be censored as follows (a best case scenario):
    - For NexoBrid – with the Median Time, estimated using the Kaplan-Meier method based on subjects from NexoBrid group having a non-missing value on

this endpoint (including subjects censored at the date of the last non-missing eschar removal assessment)

- For SoC – with the maximal time obtained for subjects randomized to SoC group (including subjects censored at the date of the last non-missing eschar removal assessment)

10.1.3. Secondary Endpoints

10.1.3.1. Incidence of Surgical Excision

Little or no missing data are expected for the secondary endpoint incidence of surgical excision. The main analysis will include all randomized patients, with those having a missing value assumed to have "failed", i.e. to have received surgery. Two sensitivity analyses will be performed. First, in each treatment group, only patients with documented surgical excision will be counted and this number divided by the total number of randomized subjects with non-missing endpoint. A second analysis will include all randomized patients and count all patients with missing data (for this endpoint) as "success" (i.e. no surgical excision performed).

Note: The main analysis specified in the protocol was the complete case analysis. However, in previous discussions concerning the study MW2010-03-02, FDA suggested to use all subjects for the main analysis of this endpoint. Therefore, in the previous study, the main analysis included all subjects with those having a missing value assumed to have "failed", i.e. to have received surgery. This approach was also adopted here.

10.1.3.2. Blood Loss

As defined in Section 10.3.3.2.1, the measure of blood loss per patient will be computed using two different methods:

- Method 1: treating the whole eschar removal process as one continuous procedure
- Method 2: summing the blood loss computed per procedure over all procedures carried out to remove eschar

Missing blood loss values will be imputed separately for each method. The procedures can be described as follows:

1. Test for normality:

- For Method 1: The measure of blood loss defined in Section 10.3.3.2.1 will be computed for each patient.
- For Method 2: The measure of blood loss defined in Section 7.2.2.210.3.3.2.1 will be computed for each procedure and aggregated to patient level.

An ANOVA model (on a patient level) incorporating the stratification factors treatment center, %TBSA, proportion of the FT area and age will be used to detect differences between both treatment groups. The normality of the data will be tested on residuals of this model using the Shapiro-Wilk test. If the normal distribution hypothesis is not rejected at the 5% significance level, then the ANOVA model will constitute the main analysis model. In case the normal distribution hypothesis is rejected at the 5% significance level, then a stratified Wilcoxon test, using the same stratification variables as outlined for the ANOVA, will be used.

2. Imputation:
   If the normality condition is satisfied, the following will be performed:
   - For Method 1: Fit a linear regression model with target variable blood loss on a patient level (as computed in Section 10.3.3.2.1) and independent variables age, wound area, depth of wound (categorized as all DPT / mixed / all FT) and course of debridement procedure (as defined in Section 10.3.3.2.5) to the subset containing patients with complete data in those variables for both treatment groups separately.
   - For Method 2: Fit a linear regression model with target variable blood loss on a procedure level (as computed in Section 10.3.3.2) and independent variables age, size of wounds treated in this procedure (wound area), proportion of area SPT of wounds treated in this procedure (as defined in Section 10.3.3.2.2), proportion of area DPT of wounds treated in this procedure (as defined in Section 10.3.3.2.3), proportion of area FT of wounds treated in this procedure (as defined in Section 10.3.3.2.4) and type of procedure (surgical/non-surgical) to the subset containing patients with complete data in those variables for both treatment groups separately. This model will treat all the procedures as independent.

Impute the missing entries of the incomplete data sets m = 25 times as follows. Draw a set of regression coefficients from their sampling distribution obtained from fitting the respective linear models outlined above. Sample a variance as the quotient between the residual variance from the linear model fitted to the data and a random variable which is chi-square distributed with corresponding degrees of freedom (a different distribution for the regression coefficients and variance will be obtained in the different treatment groups). With these sampled coefficients calculate the predicted blood loss values for the missing blood loss values using the covariate information. To these predicted blood loss values add the realizations of a normally distributed random variable with zero mean and standard deviation equal to the sampled standard deviation. The imputation outlined above will be performed using PROC MI in SAS.

If the normality condition is not satisfied, then the multiple imputation method known as predictive mean matching will be used. Random draws from the five nearest neighbors for each missing value, and m=25 multiply imputed datasets will be used. The predictive mean matching method is implemented in the SAS procedure PROC MI.

- For Method 1: The random seed will be 12467.
- For Method 2: The random seed will be 1732. On the imputed datasets the blood loss on a patient level will be calculated.

3. Analysis: Analyze each of the m complete datasets with an ANOVA or the stratified Wilcoxon test according to the test result in step 1 using the re-randomization test outlined in Section 11.1.2. An estimated treatment difference, δ (NexoBrid versus SoC), from the analysis model and a two-sided p-value, p, from the permutation test (as described in Section 11.1.2) will be obtained for each imputed dataset. In case the analysis is performed with the stratified Wilcoxon test, Hodges-Lehmann (HL) estimate for treatment difference will be used (options ALIGN=STRATA WILCOXON HL in SAS NPAR1WAY procedure).
4. Pooling: Integrate the m analysis results into a final result, as follows. A “working standard error”, s, of the estimated treatment difference for each imputed dataset will be obtained by $s = \delta/\Phi^{-1}(1-p/2)$ if $\delta>0$, or $s = \delta/\Phi^{-1}(p/2)$ if $\delta<0$, where Φ^{-1} is the inverse of the standard normal cumulative distribution function. Then Rubin’s rules will be used to combine the set of values of δ and s (using PROC MIANALYZE in SAS) - to give an overall estimate of δ and its standard error.

5. The p-values ($p_{1, ..., } p_m$) will be combined using the Licht-Rubin formula (as described in Section 4.1 of [16]), as follows:
    1) Compute $z_i = \Phi^{-1}(1 - p_i)$ for each i
    2) Compute the mean: $\bar{z} = \frac{1}{m}\sum_{i=1}^{m} z_i$
    3) Compute the between variance: $\mathrm{B} = \frac{1}{m-1}\sum_{i=1}(z_i - \bar{z})^2$
    4) Compute the total variance: $\mathrm{T} = 1 + (1 + m^{-1})\mathrm{B}$
    5) Compute the relative variance increase due to non-response: $\mathrm{r} = (1 + m^{-1})\mathrm{B}$
    6) Compute the degrees of freedom: $\nu = (m - 1)(1 + r^{-1})^2$
    7) The final p-value is then: $\Pr\left[\mathrm{t}_\nu > \bar{\mathrm{z}}/\sqrt{\mathrm{T}}\right]$ where t_ν is the t-distribution with ν degrees of freedom.

Analyses with both blood loss values will be conducted. However, the analysis with less missing values will be regarded as more reliable and is therefore considered the main analysis. The other analysis is considered supportive.

10.1.3.3. Area of Autograft in DPT Wounds

Very few cases of missing data are expected to be observed in the secondary endpoint of percentage of wound area autografted in DPT wounds. Any missing data for percent area autografted will be imputed using the multiple imputation method. The procedure can be described as follows:

1. Imputation: The imputation model will be built on a subject level by multiple linear regression of known percent area autografted on center, total area of TWs, %TBSA, number of TWs (1,2,3-4,≥5), age and sex, and will be carried out separately within each treatment group. Once the model is estimated, the imputed values will be obtained from the value predicted by the model plus a residual term that will be drawn at random from the estimated residuals. The random seed will be 141421. Twenty five (m=25) imputed datasets will be created.
2. Analysis: Each of the m complete datasets will be analyzed using a linear regression model. The explanatory variables in the model will include treatment, center, age, %TBSA, proportion of the FT area of a patient and number of TWs (1,2,3-4,≥5). An

estimated treatment difference, δ (NexoBrid versus SoC), from the analysis model and a two-sided p-value, p, from the permutation test (as described in Section 11.1.2) will be obtained for each imputed dataset.

3. Pooling: The m analysis results will be integrated into a final result – an overall estimate of δ, its standard error, and the final p-value, as described in Section 10.1.3.2 (items 4 and 5).

10.1.3.4. Incidence of Autograft in DPT Wounds

Little or no missing data are expected for the secondary endpoint incidence of Autograft performed in deep partial thickness wounds. The main analysis (on a target wound level) will include only DPT target wounds from all randomized patients, with those having a missing value assumed to have "failed", i.e. to have received autograft. Two sensitivity analyses will be performed. First, in each treatment group, only DPT target wounds with documented Autograft will be counted and this number divided by the total number of DPT target wounds from randomized subjects with non-missing endpoint (i.e. wounds with missing values will be excluded, a complete case analysis). A second analysis will include DPT target wounds from all randomized patients and count all wounds with missing data (for this endpoint) as "success" (i.e. no autograft performed).

Note: The main analysis specified in the protocol was the complete case analysis. However, for the same reasons as detailed for the secondary endpoint incidence of surgical excision (see 10.1.3.1), the main analysis will be based on DPT target wounds from all randomized subjects.

10.1.4. Safety Endpoints

10.1.4.1. Time to Reach Complete Wound Closure

The analyses of the safety endpoint time to complete wound closure on a target wound level will compute missing values as censored at the date of the last non-missing wound closure assessment. If no wound closure assessment after baseline is available, the target wound will be censored with time 0.

For additional analyses on a subject level, if a subject has a wound which did not reach complete closure, the time will be censored at last date of wound closure assessment. If a

subject has a wound with no wound assessment available after baseline, the subject will be censored with time 0.

Note: the additional analysis on a subject level had been added per FDA recommendation.

10.1.4.2. Cosmesis and Function at 12 Months from Wound Closure

A higher ratio of missing values is expected in the case of later MVSS values, especially for 24 months. It is assumed that the 6 weeks MVSS value exists. If this is not the case the missing MVSS value will be replaced by the mean MVSS value of all patients who belong to the same treatment group and the same stratum (see Section 11.1.11) at that time point. If no non-missing value in that stratum is present, the missing values will be imputed by the 6 weeks mean MVSS value of all patients who belong to the same treatment group. For missing MVSS values at later time points (12 and 24 months) the following imputation method will be used:

For each time point (6 and 12 weeks, and 6 months) and each treatment arm separately, a linear regression model with target variable MVSS at 12 months and explanatory variables MVSS at the respective time point and all stratification variables (see Section 11.1.11) is fitted to the data of patients having both MVSS values. Missing MVSS values at 12 months are then imputed with the prediction of the linear regression model, which has the last non-missing MVSS value before the 12 months measurement for the individual as explanatory variable (i.e. for patients who have the 6 months MVSS measurement but not the 12 months measurement, the 12 months measurement is imputed from the 6 months measurements; for those who have the 12 weeks measurement but not the 6 months or 12 months measurement the 12 months measurement is imputed from the 12 weeks measurement; and for those who have the MVSS measurement at 6 weeks but not at 12 weeks, 6 months or 12 months the 12 months measurement will be imputed from the 6 weeks measurement). To this prediction, the realization of a normally distributed random variable with zero mean and standard deviation equal to the residual standard deviation from the linear model is added. The random seed for this imputation is 45318.

Similarly, for each time point (6 and 12 weeks, 6, 12 and 18 months) and each treatment arm separately, a linear regression model with target variable MVSS at 24 months and explanatory variables MVSS at the respective time point and all stratification variables (see Section 11.1.11) is fitted to the data of patients having both MVSS values. Missing MVSS values at 24 months are then imputed with the prediction of the linear regression model, which

has the last non-missing MVSS value before the 24 months' measurement for the individual as explanatory variable. To this prediction, the realization of a normally distributed random variable with zero mean and standard deviation equal to the residual standard deviation from the linear model is added. The random seed for this imputation is 45448.

Using this imputation method deviates from the handling of missing data specified in the protocol. There, simple LOCF was intended. However, LOCF is not considered a suitable imputation method for the MVSS values as early MVSS values are not considered predictive for late MVSS measurements in the same individual. Furthermore, LOCF is heavily criticized in the statistics community since it may lead to biased results even when the missing occurs completely at random (see [17], [18]).

Additionally, best case – worst case imputations will be used as a supporting sensitivity analysis. One using the worst case MVSS value of 9 (observed maximum MVSS value in the SoC group in the previous study MW2004-11-02) for all missing MVSS data and one using the best case value of 0 (observed minimum MVSS value in the SoC group in the previous study MW2004-11-02).

10.1.4.3. Cosmesis and Function at 24 Months from Wound Closure

See 10.1.4.2.

10.1.4.4. General Parameters of Safety

Lower Extremity Functional Scale

The calculation of the total score will be made if less than or equal to 10% of the items were left blank. If more than 10% of items are missing the score cannot be calculated. In the event of missing items, the total score will be adjusted by dividing the total by the proportion of completed items.

Disabilities of the Arm, Shoulder and Hand (Quick-DASH)

The calculation of the total score will be made if less than or equal to 10% of the items were left blank. If more than 10% of items are missing the score cannot be calculated.

Range of Motion (ROM) (at 12 and 24 months)

In case of missing data of ROM, LOCF will be used for imputation. If a subject has no ROM measurement at all, it will be excluded from the general linear model analysis of the ROM.

EQ-5D (Quality of Life)

Missing EQ-5D values will not be imputed, but form a separate category in the analysis of this endpoint.

10.1.5. Exploratory Endpoints

10.1.5.1. POSAS

The Patient Observer Scar Assessment Scale (POSAS) consists of two parts: a patient scale and an observer scale. Both scales contain six items that are scored numerically with 1-10. If for a wound an item of the observer scale is missing the respective value will be imputed by the mean value of the other items from the observer scale of that wound. Analogously, if for a wound an item of the patient scale is missing the respective value will be imputed by the mean value of the other items from the patient scale of that wound.

If more than one value is missing within a scale for one wound, the respective scale and the overall POSAS value for that wound (see Section 10.3.5.7) will not be calculated and regarded as missing value.

10.2. Handling of Withdrawals and Dropouts

Withdrawn subjects will be evaluated in the FAS and safety set. Missing values are imputed as mentioned in Section 10.1.

Subjects with premature study discontinuation before end of the efficacy assessment period (i.e. before the 12 weeks follow-up post wound closure confirmation) lead to problems for analyses which include data from re-admissions. If a subject discontinued study participation during the efficacy assessment period, it is not known whether the subject had a(nother) re-admission and data are missing. Consequently, the corresponding variables will be set to missing for all those subjects except for “Lost-to-follow-ups” who completed the eschar

removal treatment. The exception is done under assumption that subjects who are lost to follow up would have come back for re-admission if this had been necessary (Without this exception a lot of information would have been lost in MW2010 since the percentage of subjects with re-admission is higher for “Lost-to-follow-ups” then for the rest of the population).

10.3. Data Transformation

For all continuous variables, mean, standard deviation, minimum, maximum, median, 25% quantile and 75% quantile will be calculated per treatment group and overall.

For all categorical variables counts and frequencies of categories will be calculated per treatment group and overall.

10.3.1. Demographic Data and Other Baseline Characteristics

10.3.1.1. Baseline Values

Baseline value of a variable is defined as the last available measurement of this variable before treatment (screening/pre-treatment data).

10.3.1.2. Age

Age at entry will be computed from the date of birth to the informed consent (IC) date by the following formula:

$$Age\ (in\ years)\ =\ (IC\ date\ -\ the\ date\ of\ birth)/365.25.$$

10.3.1.3. Day to Month Conversion

$$1\ month\ =\ 30.4375\ days$$

10.3.1.4. Height

A patient's height is either given in cm or inch. Height in inch will be transformed to height in cm and vice versa using the formulae

$$height(cm) = 2.54 \times height(inch)$$

$$height(inch) = 0.3937 \times height(cm)$$

Hence, height data for each patient will be available in inch as well as in cm.

10.3.1.5. Weight

A patient's weight is either given in kg or lb. Weight in lb will be transformed to weight in kg and vice versa using the formulae

$$weight(kg) = 0.4536 \times weight(lb)$$

$$weight(lb) = 2.2046 \times weight(kg)$$

Hence, weight data for each patient will be available in lb as well as in kg.

10.3.1.6. Body Mass Index (BMI)

A patient's BMI will be calculated using the following formula:

$$BMI(kg/m^2) = \frac{weight(kg)}{height(m)^2}$$

10.3.1.7. Time Since Injury

The time since injury will be computed as the time in hours between injury and randomization.

10.3.2. Primary Endpoint: Time to Complete Eschar Removal

Information about date will be transformed to time in days from randomization. Time until complete eschar removal will be calculated as the time (in days) until complete eschar removal has been achieved for a TW. From these times, time to reach complete eschar removal at a patient level, i.e. for all TW's of an individual patient will be calculated. For wounds/patients that do not reach complete eschar removal, their time will be censored at the last non-missing eschar removal assessment (typically the last debridement procedure).

Furthermore, for a sensitivity analysis, time to complete eschar removal will be computed as above but in days from ICF.

For each patient the number of TWs (1, 2, 3-4, ≥5) will be computed and used as an explanatory variable in the analysis of this endpoint.

For each patient the proportion of FT area burned will be computed.

10.3.3. Secondary Endpoints

10.3.3.1. Incidence of Surgical Excision

For each patient a binary (yes/no) variable will be computed, indicating whether this patient needed surgical excision for eschar removal. The time frame for this analysis is the time until complete eschar removal has been achieved.

Furthermore, for each target wound a binary (yes/no) variable will be computed, indicating whether this wound needed surgical excision for eschar removal.

10.3.3.2. Blood Loss

10.3.3.2.1. Blood Loss Calculation

Blood loss will be calculated using the following formula [15], [48]:

$$Blood\ Loss\ (mL) = \frac{RBCV_{before} + (Transfusion\ RBCV) - RBCV_{after}}{Hct_{after}[\%] * 0.01}$$

RBCV (mL) = Red Blood Cell Volume (mL) = Body Weight (kg) * 80 (mL/kg) * (Hematocrit [%] * 0.01)

Transfusion RBCV (mL) = [Total Volume (mL) of Whole Blood Transfused * 0.3] + [Total Volume (mL) of Packed Red Blood Cells Transfused * 0.8]

The blood loss will be computed over the whole eschar removal period, using the formula above, i.e. the eschar removal process will be considered as one continuous procedure. Therefore, only the Hematocrit value before the first eschar removal procedure and after the last procedure will be used and all blood transfusions performed between these measurements will be included.

Additionally, the blood loss will be summed over all procedures carried out to remove eschar, with the hematocrit measured immediately before each procedure and 4 hours after its completion. The amount of blood transfused will be summed over all transfusions carried out during the debridement procedure and up to 24 hours from the end time of the debridement procedure.

Analyses with both blood loss values will be conducted. However, some missing values are expected. The analysis with less missing values will be regarded as more reliable and is therefore considered the main analysis. The other analysis is considered supportive.

10.3.3.2.2. Proportion of SPT Area of Wounds Treated per Procedure

For the multiple imputation outlined in Section 10.1.3.2 information on the proportion of the SPT area of wounds treated per procedure will be computed using the following formula:

$$Prop(SPT) = \frac{\sum_i \%SPT(W_i) * \%TBSA(W_i)}{\sum_i \%TBSA(W_i)},$$

where both sums range over all target wounds treated in a debridement procedure, $\%SPT(W_i)$ is the percent area of SPT of target wound i before start of first treatment and $\%TBSA(W_i)$ is the area of the target wound i before start of first treatment. $Prop(SPT)$ will be computed for each procedure in each patient separately.

10.3.3.2.3. Proportion of DPT Area of Wounds Treated per Procedure

For the multiple imputation outlined in Section 10.1.3.2 information on the proportion of the DPT area of wounds treated per procedure will be computed using the following formula:

$$Prop(DPT) = \frac{\sum_i \%DPT(W_i) * \%TBSA(W_i)}{\sum_i \%TBSA(W_i)},$$

where both sums range over all target wounds treated in a debridement procedure, $\%DPT(W_i)$ is the percent area of DPT of target wound i before start of first treatment and $\%TBSA(W_i)$ is the area of the target wound i before start of first treatment. $Prop(DPT)$ will be computed for each procedure in each patient separately.

10.3.3.2.4. Proportion of FT Area of Wounds Treated per Procedure

For the multiple imputation outlined in Section 10.1.3.2 information on the proportion of the FT area of wounds treated per procedure will be computed using the following formula:

$$Prop(FT) = \frac{\sum_i \%FT(W_i) * \%TBSA(W_i)}{\sum_i \%TBSA(W_i)},$$

where both sums range over all target wounds treated in a debridement procedure, $\%FT(W_i)$ is the percent area of FT of target wound i before start of first treatment and $\%TBSA(W_i)$ is the area of the target wound i before start of first treatment. $Prop(FT)$ will be computed for each procedure in each patient separately.

10.3.3.2.5. Course of Debridement Procedure

For the multiple imputation outlined in Section 10.1.3.2 for each patient a categorical variable describing the course of the debridement process will be calculated. The different levels of the variable are defined as:

- "NexoBrid" or "NexoBrid - Nonsurgical Procedure" or "NexoBrid - Nonsurgical Procedure - Surgical Procedure" or "NexoBrid - Surgical Procedure"
- "Nonsurgical Procedure"
- "Nonsurgical Procedure - Surgical Procedure"
- "Surgical Procedure - Nonsurgical Procedure - Surgical Procedure"
- “Surgical Procedure”

10.3.3.3. Area of Autograft in DPT Wounds

For each DPT target wound, percent area wound autografted will be calculated as the sum of areas autografted in all autografting procedures of the specific wound, as collected on CRF wound management page (field "% area wound grafted clin. ass."). Treated DPT target wounds that had not been autografted will receive the value of 0.

For the purpose of the Main analysis, percent area wound autografted will be calculated on a subject level.

For each subject with at least one DPT target wound, the percent area wound autografted in DPT wounds will be calculated as follows:

$$TBSA\ Autograft\ in\ DPT = \frac{\sum_i TBSA_i * \%Autograft_i}{\sum_i TBSA_i}$$

where $TBSA_i$ is the % TBSA of the ith DPT target wound (the original TW assessment defined in the treatment visit), $\%Autograft_i$ is the corresponding percent area autografted of that wound and the sum is taken over all DPT target wounds of a patient.

This is a weighted average of the % area autografted over all DPT target wounds for that patient, and is equivalent to the formula:

$$TBSA\ Autograft\ in\ DPT = \frac{\sum_i areaAutografted_i}{\sum_i TBSA_i} * 100\%$$

where $areaAutografted_i$ is the area autografted in the ith DPT target wound.

Patients with only FT wounds and having no DPT wounds will be excluded from this analysis.

10.3.3.4. Incidence of Autograft in DPT Wounds

For each DPT TW a binary (yes/no) variable will be defined, indicating whether this wound needed autografting.

10.3.4. Safety Endpoints

10.3.4.1. Time to Reach Complete Wound Closure

Information about date will be transformed to time in days from randomization. Missing values will be addressed as described in Section 10.1.4.1.

For the purpose of the main analysis, time until complete wound closure (for definition see Section 6.1), will be defined as the time (in days) until complete wound closure has been achieved at a wound level. The definition of complete wound closure demands that the wound closure is to be confirmed at two consecutive study visits, 2 weeks, i.e. 14 (± 2 days), apart. A wound is only to be regarded as closed in this analysis if the wound was closed at a wound closure assessment and this was confirmed at an additional assessment within 12-16 days from the initial closure assessment. In case wound closure is not confirmed in the pre-defined time frame the wound is considered not completely closed at the first closure assessment.

Additionally, for the SoC group, a variable for time to wound closure will be calculated as above, but in the non-inferiority comparison, a 7 days non-inferiority margin will be added.

For additional analyses on a subject level, time to complete wound closure of all target wounds (last healing wound) will be computed, i.e. the time from randomization (in days) to subject's last wound closure.

For supportive analyses, the time to reach 100% wound closure will be computed on a wound level and on a subject level, where the complete wound closure is defined as 100% of wound area epithelialized and/or closed by graft without drainage or dressing requirements confirmed at two consecutive study visits, 2 weeks apart.

10.3.4.2. Cosmesis and Function at 12 Months from Wound Closure

MVSS sub scores at 12 months for pigmentation, pliability, height, vascularity, pain and pruritus will be added up to obtain the overall MVSS score (range 0 to 18) at 12 months per target

wound. The patient MVSS score will be the average of the MVSS over all TWs. To note, in case the individual item score is missing, the TW score and the patient score will be defined as missing.

10.3.4.3. Cosmesis and Function at 24 Months from Wound Closure

MVSS sub scores at 24 months for pigmentation, pliability, height, vascularity, pain and pruritus will be added up to obtain the overall MVSS score (range 0 to 18) at 24 months per target wound. The patient MVSS score will be the average of the MVSS over all TWs. To note, in case the individual item score is missing, the TW score and the patient score will be defined as missing.

10.3.4.4. General Parameters of Safety

Volume of Blood Transfusions Given

The volume of blood transfusions given during hospitalization (including readmissions) will be according to the volume (in milliliters) of whole blood and packed red blood cells transfused, presented separately, as collected on eCRF. In addition, Transfusion RBC Volume (in milliliters) will be calculated using the following formula [15], [48]:

$$Transfusion\ RBC\ Volume\ (mL) = [Total\ Volume\ (mL)\ of\ Whole\ Blood\ Transfused \times 0.3] + [Total\ Volume\ (mL)\ of\ Packed\ Red\ Blood\ Cells\ Transfused \times 0.8]$$

For subjects who discontinued study participation during the efficacy assessment period, it is not known whether a(nother) transfusion occurred and relevant data are missing. Consequently, the variable is set to missing for all those subjects except for “Lost-to-follow-ups” who completed the eschar removal treatment. The exception is done under the assumption that subjects who are lost to follow up would have come back for readmission if this had been necessary (see also safety endpoint “rates of hospital readmission”).

Analgesia and Anesthesia

Relevant analgesia and anesthesia medications will be presented in the concomitant medications tables.

Antibiotic Use

The number of days of exposure to systemic antibiotic drugs will be calculated for each patient. This will be done overall and also for antibiotic use due to AE involvement (as indicated in the eCRF) and for prophylactic antibiotic use (as indicated in the eCRF) separately.

Only records with ATC level 2 equal to “ANTIBACTERIALS FOR SYSTEMIC USE" are used. In case of a missing start date for a medication marked as pre-existing, the start date is imputed by the date of informed consent. In case of a missing stop date for a medication ongoing at study termination, the stop date is imputed by the study termination date. No other imputation is planned, i.e. if the start or stop date for antibiotics is missing, the days of antibiotic use will be counted as missing for the respective subject.

The calculation of the number of days only considers days between randomization and wound closure confirmation. Each day is only counted once; the actual frequency is not considered. Values for subjects with premature study discontinuation prior to the end of the efficacy assessment period are set to missing.

Rates of Hospital Readmission

For each patient a binary (yes/no) variable will be computed indicating whether this patient had a hospital readmission or not. Additionally, for each hospital readmission, a binary (yes/no) variable will be collected indicating whether this readmission was planned or not. A readmission will be considered as being planned when the reason for readmission was indicated as “planned wound management” or “scar modulation” in the eCRF. Readmissions for which the reason was given as “AE” in the eCRF are considered as being unplanned. Hospital readmission for which the reason is given as “other” will be reviewed by the sponsor in order to decide whether they are to be classified as planned or unplanned.

Furthermore, a variable giving the number of planned hospital readmissions and a variable giving the number of unplanned readmissions will be computed for each patient.

Relative and absolute frequencies of hospital readmission (overall, planned, and unplanned) will be calculated per treatment arm.

For subjects who discontinued study participation during the efficacy assessment period, it is not known whether a(nother) readmission occurred and relevant data are missing.

Consequently, the variable is set to missing for all those subjects except for “Lost-to-follow-ups” who completed the eschar removal treatment. The exception is done under the assumption that subjects who are lost to follow up would have come back for readmission if this had been necessary (Without this exception a lot of information would have been lost in MW2010 since the percentage of subjects with re-admission is higher for “Lost-to-follow-ups” then for the rest of the population).

Change in PTT/INR

Change in PTT/INR will be computed as PTT/INR on 4h ± 15 min post first application (NexoBrid arm) respectively to 4h post first SoC treatment (SoC arm) minus PTT/INR pre the respective procedure for each patient.

For the classification of normal and abnormal PTT/INR values (low or high), the reference limits of the local labs will be used. If INR reference ranges will not be available for specific sites, common normal range of 0.8-1.2 will be used for these sites [19].

In addition, for each subject, for each parameter (INR/PTT) a variable will be computed, indicating whether the 4 hour post first treatment measurement was post treatment potentially clinically significant high (pCS high) or post treatment potentially clinically significant low (pCS low).

For the definition of pCS high and pCS low, see Section 13.2.2.

Lower Extremity Function Scale (LEFS)

The LEFS will be assessed at the following time points: at 6, 12 weeks and months 6, 12, 18 and 24 after wound closure. It consists of 20 items which can each be answered on a scale of 0 to 4. In the database the answers will be coded as 1 to 5. Therefore, the total LEFS for a patient will be computed as

$$LEFS = \sum_{i=1}^{20} a_i - 20,$$

where a_i is the answer to item i (coded 1 to 5). As indicated in Section 10.1.4.4, the calculation of the total score will be made if less than or equal to 10% of the items were left blank. If more than 10% of items are missing the score cannot be calculated. In the event of missing

items, the total score will be adjusted by dividing the total by the proportion of completed items, as follows:

$$LEFS = \frac{\sum_{i=1}^{k} a_i}{k/20} - 20$$

where k is the number of completed items and a_i are the non-missing values.

Disabilities of the Arm, Shoulder and Hand (QuickDASH)

The QuickDASH will be assessed at the following time points: at 6, 12 weeks and months 6, 12, 18 and 24 after wound closure. It consists of 11 items which are scored 1-5. As indicated in Section 10.1.4.4, the calculation of the total score will be made if less than or equal to 10% of the items were left blank. If more than 10% of items are missing the score cannot be calculated. Hence, at least 10 of the 11 items must be completed for a score to be calculated. The assigned values for all completed responses are simply summed and averaged, producing a score out of five. This value is then transformed to a score out of 100 by subtracting one and multiplying by 25. Hence, the QuickDASH score for a single patient is calculated by:

$$QuickDASH = \left(\frac{sum\ of\ n\ responses}{n} - 1\right) * 25.$$

Range of Motion (ROM)

For each visit (12 months FU and 24 months FU), a binary (yes/no) variable will be calculated, indicating whether the range of motion assessment for that patient was abnormal or not. The range of motion assessment will be considered abnormal, if the measurement of at least one joint under consideration does not lie within the normal ranges of motion given in appendix 13 of the study protocol at the respective time point. For handling of missing data see Section 10.1.4.4.

EQ-5D (Quality of Life)

EQ-5D-Y consists of 5 items with 3 possible answers each. It will be evaluated per the EQ-5D-Y user guide [20]. Note that for children aged 2 – 11 years, the parent/caregiver will complete a proxy version of the EQ-5D for the patient, children aged 12-18 will complete the

questionnaire themselves. EQ-5D will not be assessed for children younger than 2 years (based on age of first FU visit).

For each item, relative and absolute frequencies of the answers will be computed in total and per age group, defined in the stratification variables (see Section 11.1.11). The table will be presented by treatment arm and overall. The analysis will be done per age at randomization. Furthermore, summary statistics of the EQ VAS will be computed in total and per age group. The table will be presented by treatment arm and overall.

The EQ-5D states will be converted to a single summary index using the USA value set from [21]. For subjects with incomplete questionnaires, the summary index will be a missing value.

Immunogenicity Evaluation

The immunogenicity evaluation will be done in a separate analysis and is described in a separate document.

Pyrexia and Hypothermia

Pyrexia (and related terms) and hypothermia reported as AEs will be presented in the AEs tables.

Systemic Infection

Systemic infections reported as AEs will be presented in the AEs tables.

Adverse Events

Counts and frequencies of subjects reporting treatment emergent adverse events (TEAEs), related TEAEs, and serious TEAEs per treatment arm will be computed per system organ class (SOC) and preferred term (PT). This will be done separately by:

- Severity (mild / moderate / severe AEs)
- Relatedness (not related or remotely related / possibly related or probably related or related AEs)

- Time of onset (during the treatment session / within 24 hours after treatment / 24 – 72 hours after treatment / more than 72 hours up to the end of the first week after treatment / during week 2 to week 4 after treatment / more than 4 weeks after treatment)

Note: During the treatment session is defined as the time frame from start date/time of first eschar removal procedure to end date/time of last eschar removal procedure.

All serious adverse events will be listed for the enrolled population.

For more details see Section 11.1.5.4.

For each of the two stages of analysis (see Section 11 for details), the safety data gathered in that stage will be analyzed separately. That is, in the first stage analysis (acute + 12M follow-up), all AEs reported until the 12 months follow-up will be analyzed separately for the following sub-periods: 1) until the 12 weeks follow-up, 2) between the 12 weeks follow-up and the 12 months follow-up (including AEs from the first sub-period that are still ongoing), 3) combined until the 12 months follow-up. In the second stage analysis (24 months follow-up), all AEs reported between the 12 months follow-up and the 24 months follow-up (including AEs from the first stage that are still ongoing) will be analyzed. Furthermore, a combined analysis, including all reported AEs, will be provided at the end of the trial.

Central Laboratory Values

For all analyses involving data from the central laboratory only samples taken within the allowed time range are considered

- Pre procedure: Within 12 hours before start of the procedure until up to 15 min after start.
- Post procedure: Within 24 hours after end of the procedure

Samples not taken in the allowed time range are excluded from the analysis.

Imputation of data from the central laboratory is only done for planned missing values. Version 10.01 of the study protocol specifies some planned missing laboratory samples. On the one hand, a laboratory sample before the additional application of NexoBrid is not required. On the other hand, a laboratory sample before a nonsurgical procedure is only prescribed for the first eschar removal procedure and a laboratory sample after a nonsurgical procedure is only

required for the last eschar removal procedure. To enable analysis of pre-post differences on procedure level, the planned missing values are imputed as follows:

1. Planned missing measurements before an additional application / a non-surgical procedure
   Values of a sample taken after the previous eschar removal procedure are used if the laboratory sample was taken within 12 hours before start of the eschar removal procedure.
2. Planned missing measurements after a non-surgical procedure
   Values of a sample taken before the subsequent eschar removal procedure are used if the laboratory sample was taken within 24 hours after end of the nonsurgical procedure.

For each subject, each laboratory parameter and each time point of measurement a variable will be defined, indicating whether this measurement was post treatment potentially clinically significant high (pCS high) or post treatment potentially clinically significant low (pCS low).

From this, for each subject and each laboratory parameter, an indication whether this subject experienced any pCS high measurement post baseline in this parameter and/or whether this subject experienced any pCS low measurement post baseline in this parameter will be produced.

For the definition of pCS high and pCS low, see Section 13.2.2.

Vital Signs

Body temperatures given in °F will be transformed to °C using the following formula:

$$Temperature[°C] = \frac{5}{9}(Temperature[°F] - 32)$$

Differences in the vital signs assessments (body temperature, heart rate, respiratory rate, systolic blood pressure, diastolic blood pressure) between daily visits will be calculated by

$$Diff_d = VS_d - VS_{d-1}$$

where VS_d is the vital sign assessment at day *d* (daily follow-up) and d is varied from 2 to 7.

Note: Vital sign assessments will only be recorded while a subject is hospitalized. It may however be the case that certain subjects will be discharged during the first week, making some of the measurements not available. The respective subjects will be excluded from all vital sign analysis referring to time points after hospital discharge.

Additionally, for each subject, each vital sign and each time point of measurement a variable will be defined, indicating whether this measurement was post treatment potentially clinically significant high (pCS high) or post treatment potentially clinically significant low (pCS low) (based on definitions in Section 13.2.1). Baseline categories (pCS low, pCS high and not pCS) will be created using the reference values from Section 13.2.1 without the increase/decrease criteria.

In addition, for each subject and each vital sign an indication whether the subject experienced any pCS high and/or any pCS low result post baseline (within the first 7 days) in this vital sign will be produced.

For subjects discharged before day 7, this indication will be calculated only taking into account the time during hospitalization.

Pain Assessment

Pain assessments will only be recorded while a subject is hospitalized. It may however be the case that certain subjects will be discharged during the first week, making some of the measurements not available. The respective subjects will be excluded from all pain analysis referring to time points after hospital discharge.

For each eschar removal procedure and each daily follow-up pain assessment the change in pain from baseline will be calculated as

$$Diff_d = Pain_d - Pain_{Base}$$

where $Pain_d$ is the pain assessment at the respective time point. For the calculation of the difference in pain for the eschar removal procedures the pain assessment post that procedure will be used. The pain value captured pre first procedure will be considered as baseline. If this is missing, the pain assessed at screening will be used.

In addition, for each patient, an indication will be produced whether:

- The patient experienced any post treatment abnormal (NCS) pain assessment according to investigator's assessment
- The patient experienced any post treatment abnormal (CS) pain assessment according to investigator's assessment

For subjects discharged before day 7, this indication will be calculated only taking into account the time during hospitalization.

10.3.5. Exploratory Analyses

10.3.5.1. Incidence of surgical Escharotomy procedures on circumferential extremities target wounds

For each circumferential extremity target wound a binary (yes/no) variable will be computed, indicating whether this wound was escharotomized. Furthermore, a binary (yes/no) variable for each patient with at least one circumferential extremities target wound will be computed indicating whether this patient had at least one circumferential extremities target wound which was escharotomized.

10.3.5.2. Incidence of reduction in interstitial/compartment pressure in circumferential extremities target wounds

Interstitial/compartment pressure in circumferential extremities target wounds is measured within 1 hour (+15 min) prior to start of each eschar removal procedure and post-eschar removal. For each circumferential target wound and each eschar removal procedure a binary (yes/no) variable will be computed indicating whether reduction in interstitial/compartment pressure in that wound was present during the procedure.

10.3.5.3. % wound area surgically excised for eschar removal

% of wound area surgically excised for eschar removal, A_{total}, will be calculated per patient using the following formula:

$$A_{total} = \frac{\sum_i (a_i * \%TBSA_i)}{\sum_i \%TBSA_i}$$

where $\%TBSA_i$ is the area of TW i (in % TBSA) and a_i denotes the % of wound area surgically excised for eschar removal of TW i. The sums refer to all TWs of a patient.

For TWs that were treated for eschar removal with NexoBrid and/or non-surgical procedures only, the a_i will receive the value of 0.

For TWs that were treated for eschar removal with at least one surgical procedure and no corresponding % area excised is available, the value of a_i will be set to missing. Consequently, for subjects having such TWs, A_{total} will be set to missing.

10.3.5.4. Incidence of surgically harvested donor sites wounds

For each patient, the number of donor site wounds will be calculated.

Not all donor sites are numbered during wound management and re-admissions. Therefore, Donor sites will be identified in SDTM domain PR and duplicates for donor sites already numbered will be deleted. For donor sites not already numbered but only described by the anatomical location, only one entry per procedure date and anatomical location will be kept. The remaining entries are all considered as separate donor sites and numbered under consideration of already numbered donor sites. That means, even if the anatomical location is equal, entries with no donor site number are considered as separate donor sites if the procedure date differs.

10.3.5.5. % area of surgically harvested donor site wounds

For each autograft procedure the % TBSA of the donor site is captured. These values will be summed up over all autograft procedures for each patient giving the % area of surgically harvested donor site wounds.

10.3.5.6. Blood loss using changes in hemoglobin following eschar removal procedures

For each debridement procedure, the change in hemoglobin will be calculated as the difference of the hemoglobin value before the procedure and the hemoglobin value 4 hours post the procedure. This difference in hemoglobin values will be summed for each patient.

Furthermore, for each patient, the difference of the hemoglobin value prior to the first debridement procedure and the hemoglobin value 4 hours post the last debridement procedure will be calculated.

10.3.5.7. POSAS

The Patient Observer Scar Assessment Scale (POSAS) consists of two parts: a patient scale and an observer scale. Both scales contain six items that are scored numerically with 1-10. For each wound of a patient the total score of the patient scale will be computed as:

$$POSAS_P = \sum_{i=1}^{6} p_i,$$

where p_i is the score the patient assigned to item i. The total score of the observer scale will be computed analogously as:

$$POSAS_O = \sum_{i=1}^{6} o_i,$$

where o_i is the score the observer assigned to item i.

A total score for the POSAS is then obtained by summing the patient scale score and the observer scale score:

$$POSAS_{tot} = POSAS_P + POSAS_O.$$

For each patient the mean over all TWs will be computed separately for $POSAS_{tot}$, $POSAS_P$ and $POSAS_O$.

The POSAS will be computed for TWs at the long-term follow up visits.

10.3.5.8. Duration of hospitalization

Duration of acute hospitalization will be computed as the difference in days between hospital admission and hospital discharge (initial hospitalization only). Subjects without a discharge date will be censored at the date of death (if applicable) or date of last known assessment performed during initial hospitalization period.

10.3.6. Laboratory Values

The following transformations will be applied to the central laboratory data obtained pre and post the eschar removal procedures:

Laboratory findings that are given in various units will be transformed to standard units. Laboratory values will be categorized as “abnormal (low)”, “normal” and “abnormal (high)” based on the reference limits of the labs.

Pre-post differences of laboratory values for each eschar removal procedure will be calculated (difference between post-procedure value and pre-procedure value).

Absolute and relative frequencies of “abnormal (low)”, “normal” and “abnormal (high)” findings will be calculated per laboratory value, eschar removal procedure and treatment arm.

10.4. Multiple Visits Per Time Point

Study visit time-points are scheduled at screening, pre eschar removal, eschar removal (treatment), post-treatment, hospital discharge, follow-up (weekly until wound closure, wound closure confirmation), long-term follow-up at 6 and 12 weeks and at 6, 12, 18 months and 24 months post wound closure confirmation and re-admission to hospital/day care (if done). One visit is expected for each time-point.

For all endpoints, if a subject has more than one visit with data at a time-point, the later non-missing evaluation will be used for analyses.

10.5. Implications of COVID-19 Pandemic to the Study Data

This section describes how deviations related to COVID-19 pandemic will be handled.

10.5.1. Handling of Missing Values

Data missing due to COVID-19 will be handled in the same way as missing data for any other reason, i.e. the procedures for handling of missing data described in the SAP in Section 10.1 will also be applied to missing values due to COVID-19.

10.5.2. Handling of Remotely Captured Data

Due to COVID-19 restrictions, the sponsor has submitted a protocol memo (see SAP Section 13.4), dated 09 Apr 2020, to sites allowing them to perform long-term follow-up visits

remotely. All visits could have been performed remotely, but at the time of COVID-19 outbreak almost all patients had already been post 12 weeks FU visit, so it is expected a large portion of remote assessments starting from the 6 months FU visit and onwards.

Apart from MVSS and POSAS assessments, all endpoints, allowed to be evaluated remotely during long-term follow-up visits, are self-reported. Therefore, no influence of remote study visits is expected. For MVSS and POSAS assessments, the possible influence of remotely captured data will be addressed as described in the following sub-sections.

10.5.2.1. Cosmesis and Function at 12 Months from Wound Closure

For the purpose of the main analysis, all values (both on-site and remotely collected) will be used, with missing values imputed as described in SAP Section 10.1.4.2.

For the purpose of the sensitivity analysis, to explore the possible influence of remotely captured data, the following procedure will be applied:

1. For a time point t=12 months take a subset of subjects in each treatment group who have *on-site* evaluations X_{12m} or *remote* evaluations Y_{12m} (at least 10 subjects with *on-site* evaluations at 12 months, and at least 10 subjects with *remote* evaluations at 12 months in each treatment group), and also have *on-site* evaluations for all previous time points (6 weeks, 12 weeks and 6 months). In case such a subset cannot be obtained due to absence of sufficient *on-site* values at time point 6 months, a subset with *on-site* evaluations for time points 6 weeks and 12 weeks will be considered. Denote the set of relevant previous time points as $(t-1,\dots,1)$.
2. Separately for each treatment group, regress Y_{12m} on $x_{t-1}=(1,X_{t-1},\dots,X_1)$ and all stratification variables (*S*) (see Section 11.1.11), where Y_{12m} are the *remote* evaluations at 12 months (from a subset of subjects obtained at the first step), x_{t-1} are the *on-site* evaluations at previous times (for subjects with Y_{12m}) and "1" is to indicate regression with intercept. For each regression:
    a. Obtain the estimate of residual variance, denote by $\widehat{\sigma_R^2}$.
    b. Obtain a prediction formula $\widehat{Y_{12m}}=H(x_{t-1},S)$ to predict Y_{12m} given x_{t-1} and S.
    c. Apply this formula to the (x_{t-1},S) of the subjects who had on-site evaluations X_{12m} (from a subset of subjects obtained at the first step). For each such subject there will be both X_{12m} and $\widehat{Y_{12m}}$.

d. Estimate the mean Bias $\widehat{\mu_B}$ using the mean of the differences $D_{12m} = X_{12m} - \widehat{Y_{12m}}$ for those subjects.

e. Estimate the variance of the Bias $\widehat{\sigma_B^2}$ by the estimate of the variance ($\widehat{\sigma_D^2}$) of the differences $D_{12m} = X_{12m} - \widehat{Y_{12m}}$ plus $\widehat{\sigma_R^2}$. That is $\widehat{\sigma_B^2} = \widehat{\sigma_D^2} + \widehat{\sigma_R^2}$.

3. For the sensitivity analysis of MVSS at 12 months:

a. Take the *remote* evaluations Y_{12m} and add to each a random normal variate with mean $\widehat{\mu_B}$ and variance $\widehat{\sigma_B^2}$ (for the appropriate treatment group). This will provide a "correction" to *remote* evaluations, as if they were observed during an on-site visit.

b. Repeat step a for each previous time point separately to get the "corrected" *remote* evaluations. All these "corrections" (biases) would be sampled from the same fixed distribution with mean $\widehat{\mu_B}$ and variance $\widehat{\sigma_B^2}$, estimated at step 2 above.

c. Impute missing MVSS values at 12 months, applying the imputation procedure from SAP Section 10.1.4.2 on observed *on-site* evaluations and "corrected" *remote* evaluations.

10.5.2.2. Cosmesis and Function at 24 Months from Wound Closure

For the purpose of the main analysis, all values (both on-site and remotely collected) will be used, with missing values imputed as described in SAP Section 10.1.4.2.

For the purpose of the sensitivity analysis, to explore the possible influence of remotely captured data, the following procedure will be applied:

1. Repeat step 3a from the above section (10.5.2.1) for additional follow-up time points 18 months and 24 months, to get the "corrected" *remote* evaluations.
2. Impute missing MVSS values at 24 months, applying the imputation procedure from SAP Section 10.1.4.2 on observed *on-site* evaluations and "corrected" *remote* evaluations from all previous time points (for time points up to 12 months previously "corrected" *remote* evaluations can be used).

10.5.2.3. Exploratory Analyses of MVSS and POSAS

For the purpose of exploratory analyses of MVSS (see SAP Sections 11.1.6.8 and 11.2.5.2) and POSAS (see SAP Sections 11.1.6.7 and 11.2.5.1) all observed values (both on-site and remotely collected) will be used.

11. Statistical Analysis of Target Variables

The analysis is planned to be carried out in two stages, as described below.

1. The first analysis will be performed at the end of the 12 months follow up period. This analysis will be the only inductive analysis of the trial and will include statistical tests for the primary and secondary endpoints as described above, as well as 12 months safety endpoints. At this time-point, data will be available for all patients on the primary and secondary end-points, as well as the 12 months safety endpoints. Missing values for early drop outs etc. will be handled as described in Sections 10.1 and 10.2. The complete data set documented so far will be locked and analyzed as described above. The 24 months data, although captured in the eCRF for a few of the subjects at the stage of end of the 12 months follow up period, will not be included, revealed or analyzed during these analyses.

2. The second and final analysis covers the data of the full 24 months safety follow up. It will be conducted after the last patient has reached the 24 months assessment, 12 months after first stage of analysis. At this analysis, all accumulated safety endpoints at the 24 months follow up will be evaluated.

Long term follow up assessments of Cosmesis (MVSS and POSAS), quality of life (EQ5D and BOQ) and functionality (LEFS, QuickDASH and ROM) will be performed by a “blinded assessor”. The assessor signs a “blinded assessor memo” which describes the procedures required to maintain the investigator blinded from the treatment arm. The blinded assessor will remain blinded throughout the whole duration of the trial.

The data sets for analysis will be 100% source document verified and signed by the site investigator. Therefore, data changes are prevented during the further conduct of the study. Data management quality control actions will ensure that data used in the analysis is not changed afterwards (see the Data Management Plan for further details).

For a better overview of this SAP, the statistical analysis of target variables described below will be displayed separately for these 2 stages of analyses.

All variables not mentioned explicitly in the following will be analyzed descriptively (as part of the case-wise listing).

Unless otherwise specified, all statistical tests will be conducted against a two-sided alternative hypothesis, employing a significance level of 0.05.

11.1. First Stage Analysis –12 Months Follow Up

11.1.1. Demographics Data and Other Baseline Characteristics

For all considerations regarding the demographics and other baseline characteristics the full analysis set will be used.

The following demographic and baseline information will be tabulated:

- Age Group
- Age
- Gender
- Ethnicity
- Race
- Height (cm)
- Weight (kg)
- BMI
- Physical examinations (categorized as normal, abnormal-NCS, abnormal-CS, not done):
    - Head, eyes, ears, nose and throat (HEENT)
    - Respiratory system
    - Cardiovascular system
    - Abdomen
    - Musculoskeletal / Extremities
    - Neurological
    - Lymph Nodes
- Vital signs:
    - Temperature
    - Heart rate
    - Respiratory rate

    - Systolic and diastolic blood pressure (mmHg)
    - Assessment of vital signs (normal, abnormal-NCS, abnormal-CS)
- Pain assessment:
    - FPS-R
    - Assessment of pain (normal, abnormal-NCS, abnormal-CS)
- Local laboratory:
    - Diabetes status (yes/no)
    - HbA1c (%)
- Burn history (on a patient level):
    - Time since injury to randomization (hours)
    - Etiology of injury (Fire/Flame, Scald, Contact)
    - Place of injury (Outdoors, Indoors, Car, Other)
- General wound description (on a wound level):
    - Anatomical location (face, head, neck, etc.)
    - %TBSA of 2° SPT Burns (clinical assessment)
    - %TBSA of 2° DPT Burns (clinical assessment)
    - %TBSA of 3° FT Burns (clinical assessment)
    - Information whether escharotomy had been performed prior to randomization
    - Eschar description (white, moist, dry, etc.)
    - Information whether cleansing was performed
    - Information whether all superficial keratin (blister) were removed from wound
- Target wound description (assessments at screening):
    - General wound description (see above; on a wound level; for target wounds only)
    - Information whether the wound was treated prior to enrollment (TWs)
    - Treatment of TWs prior to enrollment (SSD, Iodine, etc.)
- Target wound description (on a patient level, assessments at screening):
    - Total wounds %TBSA
    - TWs DPT total area (%TBSA) (clinical assessment)
    - TWs FT total area (%TBSA) (clinical assessment)

    - Overall TWs depth per patient (all full thickness, mixed, all deep partial thickness) (clinical assessment)
    - Overall TW area (SPT+DPT+FT) (clinical assessment)
- Criteria for inclusion (met/not met)
- Counts of patients/wound per strata (%TBSA cut point 15%, % FT burns cut point 20 %, center)

All the above baseline information will be presented using descriptive statistics by treatment group and in total.

If any of the baseline factors gender, percent SPT area (the area of SPT as a percentage of the total area of all TWs, classified as <25% or ≥25%), time from injury to randomization or number of TWs (1, 2, 3-4, ≥5) are found to be significantly different between the treatment groups at the 15% level (Chi-squared tests (or in case of small estimated cell counts Fisher's exact test) for categorical variables and one-way analysis of variance for continuous variables) then the factor will be included as an extra adjusting covariate in the supportive analysis models for the primary and secondary efficacy endpoints.

Furthermore, a table for the disposition of patients will be produced.

11.1.2. Re-Randomization Test

Treatment allocation will be performed in this trial by minimization, therefore the p-values reported in this trial for the main analyses of the primary and secondary endpoints, performed on the FAS, will be based on the re-randomization test. To note, all other analyses (sensitivity, supportive, additional), including analyses on PP population, will be performed without the re-randomization test.

The re-randomization test is advocated by Proschan et al (Biometrics 2011; 67:1135-41) for trials that employ minimization, and belongs to the class of tests that are based on permuting the treatment allocations of the patients.

The test works as follows. Suppose the analysis for an endpoint is performed as described in Sections 11.1.3 or 11.1.4 below and a p-value for the estimated treatment difference is p*. We now return to the list of patients in the same order as they entered the trial and re-apply the minimization procedure to their treatment allocation, as described in Appendix 13.1. Because

the procedure involves random allocations at various junctures during the trial, the next time we apply the procedure it will in general yield a different set of treatment allocations for the patients. Having obtained these new treatment allocations, we then re-analyze the data as if these had been the true treatment allocations, and obtain a new p-value p_1 and corresponding test statistic s_1 for the test of the treatment difference. We repeat this step a large number of times (in this trial we will use 8000 repeats) and obtain a set of 8000 pairs of p-values and test statistics (p_1, s_1), (p_2, s_2), …, (p_{8000}, s_{8000}). If exactly n of these 8000 p-values are less than or equal to the value p* (obtained on the real data) with the test statistic in the direction of advantage to NexoBrid, then the re-randomization test p-value is assigned the value 2*n/8000.

The random seed for the re-randomization will be 48920.

Following teleconference with the FDA on April 2017, it was decided that the form of the minimization used for treatment allocation in the trial will be modified from strict minimization (without a random element) to minimization with a random element. The re-randomization test can naturally accommodate this modification. When re-allocating treatments for this re-randomization analysis, those patients allocated treatment in the trial using strict minimization will be re-allocated treatment using strict minimization, while those patients who entered later and were allocated treatment using minimization with a random element will be re-allocated treatment using minimization with a random element.

The minimum and maximum numbers of patients allocated to each treatment group over the 8000 re-randomized samples will be reported.

11.1.3. Primary Efficacy Analysis: Time to Complete Eschar Removal

The primary efficacy variable will be analyzed using the FAS. Missing data for the main primary analysis and for the three sensitivity analyses will be handled by the methods described in Section 10.1.2. In addition, the primary analysis will be repeated for the PP collective as sensitivity analysis.

Time until complete eschar removal will be defined as the time until complete eschar removal has been achieved for all TWs. This will be measured as time from the randomization date (days). Kaplan-Meier curves will be presented graphically to display the distribution of time to complete eschar removal under the two treatments. Median time to complete eschar removal

will be estimated for each treatment group with a 95% confidence interval. Furthermore, descriptive statistics for time to complete eschar removal will be presented. The treatment groups will be compared using a Cox regression model. The comparison will be adjusted for clinical center, age, %TBSA, proportion of the FT area of a patient and number of TWs (1,2,3-4, ≥5) by including each of them in the Cox regression model together with the treatment variable (NexoBrid vs. SoC). The treatment groups will be compared by testing the null hypothesis of no difference, comparing the ratio of the estimated treatment coefficient to its standard error to a standard normal distribution.

The Cox Regression analysis outlined above will be performed if the proportional hazards assumption appears to hold. This will be checked by including in the regression model a variable representing the interaction between time since randomization and treatment group. This coefficient will be tested (on actual sample allocation, without the re-randomization test) to be nonzero. Additionally, a log likelihood test comparing the model fit of a model with all time by covariate interactions added with a model without these interactions will be computed. The proportionality assumption will be rejected if one of these tests yields a significant result at the 10% level. In this case, we will use a generalized Wilcoxon-Gehan test. The Wilcoxon-Gehan test will be performed using the Wilcoxon option (also provided by default) in the SAS procedure PROC LIFETEST using the following SAS Code:

```
PROC LIFETEST DATA=burn_data;
TIME time*status(0);
STRATA agegroup ftgroup numtw tbsagroup centergroup / GROUP =
treatment_group;
RUN;
```

Here burn_data is the data set containing all necessary information. The variable time is the time to complete eschar removal, status is a censoring variable (0 for censor and 1 for complete ER event) and agegroup, ftgroup, numtw, tbsagroup, and centergroup are the variables, which the model is adjusted for. Agegroup, ftgroup, tbsagroup and centergroup are the stratification levels defined in Section 11.1.11 and numtw is the number of target wounds (1, 2, 3-4, ≥5).

Supportive analyses will include adjustment for other baseline variables that are imbalanced, based on the analyses described in Section 11.1.1.

The comparison between treatments will also be conducted on the time to complete eschar removal measured from the date of ICF as an additional analysis.

As an additional analysis, time to complete eschar removal will be analyzed on a target wound level using a method of survival analysis with clustered data that is based on appropriate assumptions. By “clustered data”, we refer to the multiple target wounds that can occur in a patient. The proportional hazards assumption will be checked in the same way as in the patient level analysis. I.e. a Cox regression model will be fitted with covariates treatment, center, age, %TBSA, proportion of the FT area of a patient, number of TWs (1,2,3-4, ≥5) and a variable representing the interaction between time since randomization and treatment group. The latter coefficient will be tested to be nonzero. Additionally, a log likelihood test comparing the model fit of a model with all time by covariate interactions added with a model without these interactions will be computed. The proportionality assumption will be rejected if one of these tests yields a significant result at the 10% level. If the proportional hazards assumption is appropriate we will use a marginal Cox regression analysis with a robust sandwich estimator. If not, then we will use a parametric frailty model. Either method can be implemented in SAS (see Gharibvand and Liu, 2009 for details). The marginal Cox regression model is implemented in the PHREG procedure and the robust variance option is implemented by specifying the COVS(AGGREGATE) option. This induces the use of the robust variance method of Binder et. al. [22]. The parametric frailty model will be implemented using the following SAS-code:

```
PROC PHREG DATA=burn_data;
CLASS Subject_id treatment_group centergroup numtw;
MODEL time*status(0) = treatment_group centergroup age tbsa pctFT numtw;
RANDOM Subject_id;
HAZARDRATIO "Frailty Model" treatment_group;
RUN;
```

where subject_ID is a unique identification number for each patient, time is the time to complete eschar removal on a wound level, status is an indicator whether the wound has achieved complete eschar removal or not (censoring indicator, 0 for censor and 1 for complete ER event), treatment_group gives the treatment group, centergroup is the stratification factor defined in Section 11.1.11, age is the patient’s age, TBSA is the overall percent TBSA of a patient, pctFT is the proportion of the FT area of a patient, numtw is the number of TWs (1,2,3-4, ≥5) of the patient and burn_data is a SAS data set containing the required information. These methods include information on all the target wounds of each patient and account for any

within-patient correlation. Missing information due to incomplete follow-up is incorporated naturally into such analyses as censored observations.

Kaplan-Meier curves will be presented graphically to display the distribution of time to complete eschar removal (on a wound level) under the two treatments.

11.1.4. Secondary Efficacy Analyses

The secondary endpoints will be analyzed as described below. Missing data will be handled by the methods described in Section 10.1. All secondary endpoints will be analyzed using the FAS. Sensitivity analyses will include analyzing the endpoints using the PP as analysis population.

11.1.4.1. Incidence of Surgical Excision

This is a binary yes/no variable and the proportion of patients who need excision will be compared using logistic regression. Missing data for the main analysis and for the two sensitivity analyses will be handled by the methods described in Section 10.1.3.1. The explanatory variables in the model will include treatment, center, age, %TBSA, proportion of the FT area of a patient and number of TWs (1,2,3-4, ≥5). The odds ratio of requiring surgery for NexoBrid versus SoC will be estimated from the model, as well as a 95% confidence interval and the level of statistical significance.

As sensitivity analysis, the procedure outlined above will be repeated using the PP instead of the FAS as analysis population.

A supportive analysis will be performed adjusting for other baseline variables that are imbalanced, based on the analyses described in Section 11.1.1.

As an additional analysis, the incidence of surgical excision will be compared on a wound level using a mixed logistic regression model with a random effect for patient.

11.1.4.2. Blood Loss

As defined in Section 10.3.3.2.1, the measure of blood loss per patient will be computed using two different methods:

- Method 1: treating the whole eschar removal process as one continuous procedure

- Method 2: summing the blood loss computed per procedure over all procedures carried out to remove eschar

Descriptive statistics of the blood loss values will be provided by treatment group.

The imputation of missing values, involving the multiple imputations procedure, and further statistical analysis are detailed in Section 10.1.3.2. An overall estimate of the treatment difference with its standard error will be reported, together with the final combined p-value.

The analysis with less missing values will be regarded as more reliable and is therefore considered the main analysis. The other analysis is considered supportive.

As sensitivity analysis, the analysis outlined above will be repeated using the PP instead of the FAS as analysis population.

A supportive analysis will be performed adjusting for other baseline variables that are imbalanced, based on the analyses described in Section 11.1.1.

Note: No wound level analysis of this endpoint is possible, since blood loss cannot be attributed to single target wounds.

11.1.4.3. Area of Autograft in DPT Wounds

For the purpose of the main analysis, the mean percent area of autograft in DPT wounds will be compared on a subject level between NexoBrid and SoC. This analysis will be restricted to subjects having at least one DPT target wound. Descriptive statistics of the observed values on a subject level (including the number of missing values) will be provided by treatment group. The imputation of missing values, involving the multiple imputations procedure, and further statistical analysis are detailed in Section 10.1.3.3. An overall estimate of the treatment difference with its standard error will be reported, together with the final combined p-value.

As sensitivity analysis, the main analysis outlined above will be repeated using the PP instead of the FAS as analysis population, also restricted to subjects having at least one DPT target wound.

Supportive analyses on a subject level will include adjustments for additional baseline variables which are found to be imbalanced, based on the analyses described in Section 11.1.1.

As an additional analysis, the percent area autografted in DPT target wounds will be analyzed on a target wound level. This analysis will be restricted to DPT target wounds only. Descriptive statistics will be provided on a target wound level for each treatment group. The differences in the distribution of percent area autografted in DPT wounds between the treatment groups will be tested using a mixed linear regression model with patient as the random effect that accounts for within-subject correlation between target wounds.

11.1.4.4. Incidence of Autograft in DPT Wounds

This is a binary yes/no variable and the proportion of Deep Partial Thickness TWs that were autografted will be compared (on a wound level). Missing data for the main analysis and for the two sensitivity analyses will be handled by the methods described in Section 10.1.3.4. The comparison of NexoBrid versus SoC will be conducted using logistic regression within the generalized estimating equations (GEE) framework that accounts for within-subject correlation between target wounds. The explanatory variables in the model will include treatment, center, age, %TBSA (wound level), proportion of the FT area of a patient and number of TWs (1,2,3-4, ≥5). The odds ratio of requiring Autograft for NexoBrid versus SoC will be estimated from the model, as well as 95% confidence intervals and the level of statistical significance.

As sensitivity analysis, the procedure outlined above will be repeated using the PP instead of the FAS as analysis population.

Supportive analyses will include adjustments for baseline variables which are found to be imbalanced, based on the analyses described in Section 11.1.1.

These analyses will only take into account DPT target wounds.

11.1.5. Safety Analyses

All safety variables except for the first safety endpoint “time to reach complete wound closure”, which will be evaluated using the FAS, will be evaluated on the safety set.

11.1.5.1. Time to Reach Complete Wound Closure

For the purpose of the main analysis, time to reach complete wound closure will be compared between the NexoBrid and SoC at a wound level using a method of survival analysis with clustered data that is based on appropriate assumptions. By “clustered data”, we refer to the

multiple target wounds that can occur in a patient. As stated in the protocol, a non-inferiority margin will be incorporated into the analysis that will represent a 7-day advantage to the SoC group. This will be done by adding 7 days to the wound closure times in the SoC group. After that, the proportional hazards assumption will be checked in the same way as in the analysis of the timely eschar removal endpoint. I.e. a Cox regression model will be fitted with covariates treatment, center, age, %TBSA, proportion of the FT area of a patient, number of TWs (1,2,3-4, ≥5) and a variable representing the interaction between time since randomization and treatment group. The latter coefficient will be tested to be nonzero. Additionally, a log likelihood test comparing the model fit of a model with all time by covariate interactions added with a model without these interactions will be computed. The proportionality assumption will be rejected if one of these tests yields a significant result at the 10% level. If the proportional hazards assumption is appropriate we will use a marginal Cox regression analysis with a robust sandwich estimator. If not, then we will use a parametric frailty model. Either method can be implemented in SAS (see Gharibvand and Liu, 2009 for details). The marginal Cox regression model is implemented in the PHREG procedure and the robust variance option is implemented by specifying the COVS(AGGREGATE) option. This induces the use of the robust variance method of Binder et. al. [22]. The parametric frailty model will be implemented using the following SAS-code:

```
PROC PHREG DATA=burn_data;
CLASS Subject_id treatment_group centergroup numtw;
MODEL time*status(0) = treatment_group centergroup age tbsa pctFT numtw;
RANDOM Subject_id;
HAZARDRATIO "Frailty Model" treatment_group;
RUN;
```

where subject_ID is a unique identification number for each patient, time is the time to wound closure on a wound level, status is an indicator whether the wound has achieved complete wound closure or not (censoring indicator), treatment_group gives the treatment group, centergroup is the stratification factor defined in Section 11.1.11, age is the patient's age, tbsa is the overall percent TBSA of a patient, pctFT is the proportion of the FT area of a patient, numtw is the number of TWs (1,2,3-4, ≥5) of the patient and burn_data is a SAS data set containing the required information. These methods include information on all the target wounds of each patient and account for any within-patient correlation. Missing information due to incomplete follow-up is incorporated naturally into such analyses as censored observations.

In case the proportional hazards assumption is rejected and the parametric frailty model is used, the underlying proportional hazards assumption (conditional on the frailty term) will be checked by adding a variable representing the interaction between time since randomization and treatment group to the model. In case the coefficient of the latter term is significantly different from zero (at the 5% level) a further sensitivity analysis using an accelerated failure time (AFT) model with shared frailty will be performed. This model will be implemented using the following SAS-code:

```
PROC NLMIXED DATA=burn_data;
BOUNDS gamma>0;
linp = b0 + b1*treatment_group + b2*age1 + b3*age2 + b4*age3 + b5*tbsa1 +
b6*ft1+ b7*number_tw2+ b8*number_tw3 + b9*number_tw4 + b10*US + b11*EEU +
b12*WEU + z;
alpha = exp(-linp);
G_t = 1/(1+alpha*time**gamma);
g = alpha*gamma*time**(gamma-1)/(1+alpha*time**gamma)**2;
ll = (censor=0)*log(g) + (censor=1)*log(G_t);

model time~general(ll);

RANDOM z ~ normal(0,sigma) subject=subject_id;

RUN;
```

where subject_ID is a unique identification number for each patient, time is the time to wound closure on a wound level. Censor is a censoring indicator, treatment_group gives the treatment group, age1-age3 are indicator variables for a subject being in age group A1 – A3 respectively (per age groups defined in Section 11.1.11), ft1is an indicator for a subject with proportion of the FT area <20% out of the total burned area (group C1, defined in Section 11.1.11), TBSA1 is an indicator for a subject with overall %TBSA in (1% to 15%), inclusive (group B1, defined in Section 11.1.11), number_tw2 is an indicator variable for the subject having 2 target wounds (patient level), number_tw3 is an indicator variable for the subject having 3-4 target wounds (patient level), number_tw4 is an indicator variable for the subject having ≥5 target wounds (patient level), US, EEU and WEU are indicator variables for the individual being in region US, EEU and WEU (patient level) and burn_data is a SAS data set containing the required information.

Kaplan-Meier curves will be presented graphically to display the distribution of time to complete wound closure under the two treatments.

If any of the baseline factors are found to be significantly different between the treatment groups (based on the analyses described in Section 11.1.1), then the factor will be included as an extra adjusting covariate in a supportive analysis of time to complete wound closure.

As an additional analysis, the time to complete wound closure will be analyzed on a subject level. Kaplan-Meier curves will be presented graphically to display the distribution of time to complete eschar removal under the two treatments on a subject level. Median time to complete eschar removal will be estimated for each treatment group with a 95% confidence interval. Furthermore, descriptive statistics for time to complete eschar removal will be presented. The treatment groups will be compared using either a Cox regression model (if the proportional hazards assumption appears to hold), or a generalized Wilcoxon-Gehan test (if the proportional hazards assumption does not hold), as described in Section 11.1.3 for the main analysis of the primary efficacy endpoint.

As supportive analyses, the time to reach 100% wound closure will be analyzed on a wound level (similar to the main analysis outlined above) and on a subject level (similar to the additional analysis outlined above).

11.1.5.2. Cosmesis and Function at 12 Months from Wound Closure

The MVSS score for a patient will be calculated as described in Section 10.3.4.2. Descriptive statistics of MVSS score at 12 months will be provided by treatment group, for both observed and imputed data (for handling of missing data see Section 10.1.4.2.) The treatment groups will be compared using a linear model with imputed MVSS score at 12 months as the dependent variable. The explanatory variables in the model will include treatment, and the following variables: treatment center, age, %TBSA, proportion of the FT area of a patient and number of TWs (1,2,3-4, ≥5). The coefficient corresponding to the treatment group will be estimated and will represent the estimated mean difference in MVSS score at 12 months between NexoBrid and SoC adjusted for any imbalance in the stratification factors. The results of the regression analysis will be provided.

A clinically meaningful difference will be incorporated into the analysis that will represent a 1.9 or more units advantage to the SoC group, so that the analysis will show whether the NexoBrid group is estimated to have on average a MVSS score that is not worse by more than 1.9 units than the score for the SoC group. That is, the following hypothesis are about to be tested:

$$H_0: \Delta \geq 1.9 \quad vs. \quad H_1: \Delta < 1.9$$

Where Δ is the mean difference in MVSS score between NexoBrid and SoC adjusted for any imbalance in the stratification factors. These hypotheses will be tested by comparing the 95% confidence interval for the coefficient of the treatment group from the linear model outlined above to the clinically meaningful difference 1.9. The null hypothesis will be rejected if the upper bound of this interval is smaller than 1.9.

The main analysis will be performed as described above using both on-site and remotely collected evaluations, with missing values imputed as described in SAP Section 10.1.4.2.

In addition, a sensitivity analysis will be performed (similar to the main analysis described above) to explore the possible influence of remotely captured data. For this sensitivity analysis, the remote evaluations will be "corrected", as described in SAP Section 10.5.2.1, and missing values at 12 months will be imputed applying the imputation procedure from SAP Section 10.1.4.2 on observed *on-site* evaluations and "corrected" *remote* evaluations.

11.1.5.3. Cosmesis and Function at 24 Months from Wound Closure

Not applicable for this stage of analysis since the relevant information for all patients is not available yet. This endpoint will be analyzed in the second stage analysis (see 11.2.4.3).

11.1.5.4. Adverse Events

In this analysis, only the treatment emergent adverse events occurring until the 12 months follow-up visit will be analyzed.

All adverse events will be listed. Furthermore, the following tables will be produced separately for the following sub-periods: 1) until the 12 weeks follow-up, 2) between the 12 weeks follow-up and the 12 months follow-up (including AEs from the first sub-period that are still ongoing), 3) combined until the 12 months follow-up:

- Summary of Treatment Emergent Adverse Events (TEAEs) on a subject level and on an event level
- Summary of TEAEs by System Organ Class (SOC), and Preferred Term (PT) separately for:
    - All TEAEs

  - Treatment related TEAEs
  - General and local, not TW-related TEAEs
  - Local, TW-related TEAEs
- Summary of Special Interest TEAEs by PT (for sub-period 1), separately for:
  - TEAEs with PTs associated with fever (Pyrexia, Body Temperature Increased, Hyperthermia)
  - TEAEs associated with target wound infections (only if Local, TW-related: relevant preferred terms will be reviewed after finalization of coding and can include for example preferred terms as: Wound Infection Bacterial, Staphylococcal Skin Infection, Wound Infection, Wound Infection Staphylococcal if reported as local TW-related)
  - TEAEs associated with target wound pain (relevant preferred terms will be reviewed after finalization of coding)
- Summary of TEAEs by SOC, PT, and time of onset (for sub-period 1) separately for:
  - All TEAEs
  - General and local, not TW-related TEAEs
  - Local, TW-related TEAEs
  - All Treatment Emergent Serious Adverse Events (TESAEs)
- Summary of TEAE by SOC, PT and maximum intensity. In this table, patients who experienced multiple AEs with the same PT, but different intensities will be counted only once per PT with their maximum intensity. Subjects with multiple AEs with the same SOC but different intensities will be counted only once per SOC with their maximum intensity in the SOC row. This table will be produced separately for:
  - All TEAEs
  - Treatment related TEAEs
  - General and local, not TW-related Treatment related TEAEs
  - Local, TW-related Treatment related TEAEs
  - All TESAEs
  - Treatment related TESAEs

11.1.5.5. Central Laboratory Values

Laboratory measurements obtained prior and after each eschar removal procedure as well as pre-post differences in laboratory values for each eschar removal procedure will be presented by treatment group using descriptive statistics. In addition, laboratory measurements obtained prior the first eschar removal procedure and after the last eschar removal procedure, as well as the corresponding pre-post difference, will be presented by treatment group using descriptive statistics.

As described in Section 10.3.6, laboratory values will be categorized as “abnormal (low)”, “normal” and “abnormal (high)” based on the reference limits of the labs. This categorization will be analyzed by displaying counts and percentages of each category by treatment group and by procedure, as well as at the last eschar removal procedure, for each laboratory value separately.

Shift tables for central laboratory values indicating shifts from baseline to post every type of eschar removal procedure (first application, additional application, surgical procedure, additional surgical procedure, and nonsurgical procedure), as well as to post last eschar removal procedure, will be produced.

Furthermore, shift tables indicating shifts from baseline (low/normal/high) to any abnormal (low/high) value will be presented.

Additional tables will indicate shifts from baseline (low/normal/high) to any post treatment potentially clinically significant result. A listing of subjects experiencing such shifts will be provided.

The denominator for the shift tables will be the patients that had both pre and post results for the respective procedure.

The above tables will be presented for biochemistry parameters, hematology parameters and urine analysis parameters (where applicable) separately, and for parameters with numeric values only.

11.1.5.6. Vital Signs

All vital signs (body temperature, heart rate, respiratory rate, systolic blood pressure, diastolic blood pressure) will be presented by treatment group using descriptive statistics.

The analyses will be done for the vital signs assessments at baseline, at the first 7 daily assessments as well as for the pre-post differences of the daily vital signs assessments described in Section 10.3.4.4 separately.

Note: Vital sign assessments will only be recorded while a subject is hospitalized. It may however be the case that certain subjects will be discharged during the first week, making some of the measurements not available. The respective subjects will be excluded from all vital sign analysis referring to time points after hospital discharge.

Furthermore, for the vital signs assessment at baseline and at the first 7 daily assessments the information whether the overall vital signs investigator's assessment was normal, abnormal (NCS) or abnormal (CS) is given in the eCRF. For these overall vital signs investigator's assessment classifications, a shift table will be produced indicating shifts from baseline to each of the daily assessments from day 1 to day 7 and to the last daily assessment within the first 7 days.

Additionally, for individual vital signs parameters a shift table will be produced indicating shifts from baseline to post treatment potentially clinically significant (pCS) results at each of the daily assessments from day 1 to day 7 and at the last daily assessment within the first 7 days.

Furthermore, a shift table indicating shifts from baseline to any post treatment potentially clinically significant result within the first 7 daily assessments will be provided. A subject listing of patients experiencing such shifts will be provided. As additional analysis, this shift analysis will be repeated excluding subjects with hospital discharge before day 7.

11.1.5.7. Pain Assessment

The pain assessment score (FPS-R) will be analyzed descriptively by treatment group with counts and percentages and compared between groups using a chi square test (or Fisher’s exact test in case of small estimated cell counts). Additionally, the pain score will be evaluated as a metric variable and presented using numeric descriptive statistics by treatment group.

The analyses will be done for the pain assessments at screening, pre and post eschar removal procedures as well as at the first 7 daily assessments separately. Changes from baseline per time point as calculated in Section 10.3.4.4, will be presented by treatment group using descriptive statistics.

Note: Pain assessments will only be recorded while a subject is hospitalized. It may however be the case that certain subjects will be discharged during the first week, making some of the measurements not available. The respective subjects will be excluded from all pain analysis referring to time points after hospital discharge.

Furthermore, for the pain assessment at screening, pre and post eschar removal procedures and at the first 7 daily assessments the information whether the pain assessment was normal, abnormal (NCS) or abnormal (CS) is given in the eCRF. These classifications will be analyzed separately for each time point of pain assessment by displaying shifts from baseline by treatment group.

Furthermore, a shift table indicating shifts from baseline to any abnormal (NCS) or any abnormal (CS) result will be provided. As additional analysis, this shift analysis will be repeated excluding subjects with hospital discharge before day 7.

The Pain scale (FPS-R) was validated only for children above the age of 4. The analyses described above will be therefore done for the children above the age of 4 and below the age of 4 separately.

11.1.5.8. Blood Transfusions

The volume of blood transfusions in milliliters (as calculated in Section 10.3.4.4) will be presented by treatment group using descriptive statistics. The mean volume of blood transfusions between treatment groups will be compared by a one-way analysis of variance.

The volume of blood transfusions will be presented overall (during hospitalization), as well as separately for volumes transfused before start of treatment, during the ER period, within 1 week after the ER period and later than 1 week after the ER period.

The volume of blood transfusions will be analyzed as described above separately for calculated transfusion RBC volume, for the volume of packed red blood cells transfused and the volume of whole blood transfused.

11.1.5.9. Immunogenicity and PK Evaluation

The immunogenicity and PK evaluation will be done separately and is described in a separate document.

11.1.5.10. Antibiotic Use

The number of days of exposure to systemic antibiotic drugs will be presented by treatment group using descriptive statistics and compared between groups with a one-way analysis of variance. This analysis will be conducted for overall systemic antibiotic use, for antibiotic use due to AE involvement (as indicated in the eCRF) and for prophylactic antibiotic use (as indicated in the eCRF) separately.

11.1.5.11. Medical History

Medical history events will be classified using the MedDRA dictionary. They will be tabulated with reference to system organ class, preferred term and frequency of occurrence. The system organ classes will be sorted alphabetically, the preferred terms within the system organ classes will be displayed in decreasing order of incidence in the NexoBrid group.

11.1.5.12. Concomitant Medications

In this analysis, only the concomitant medications administered during the first phase (until the 12 months follow-up for each patient) will be analyzed. The number of patients involved will be tabulated with overall numbers and with reference to ATC coding using the Level 2 and Level 4 codes. The summary tables will be produced separately for the following sub-periods: 1) until the 12 weeks follow-up, 2) between the 12 weeks follow-up and the 12 months follow-up (including CM with a start date before the 12 weeks follow up but an end date after that or still ongoing), 3) combined until the 12 months follow-up.

11.1.5.13. Hospital Readmission Rates

Hospital admission rates will be analyzed on a patient level by displaying counts and percentages by treatment group. The proportion of patients needing readmission to a hospital will be compared between treatment groups using a chi square test (or Fisher’s exact test in case of small estimated cell counts).

Furthermore, the number of planned hospital readmissions per patient will be analyzed by displaying counts and percentages by treatment group. The number of planned hospital readmissions per patient will be compared between treatment groups using a chi square test (or Fisher’s exact test in case of small estimated cell counts).

Additionally, the number of unplanned hospital readmissions per patient will be analyzed by displaying counts and percentages by treatment group. The number of unplanned hospital readmissions per patient will be compared between treatment groups using a chi square test (or Fisher's exact test in case of small estimated cell counts).

Hospital readmission will also be analyzed on a readmission level. Counts and percentages of planned/unplanned readmissions will be displayed by treatment group. The proportion of planned/unplanned readmissions to hospitals will be compared between treatment groups using a chi square test (or Fisher's exact test in case of small estimated cell counts).

11.1.5.14. Change in PTT/INR

Change in INR/PTT will be evaluated at 4h post first treatment. For the classification of normal and abnormal INR/PTT values, the reference limits of the local labs will be used (see also Section 10.3.4.4).

The following tables will be produced for INR and PTT:

- Shift tables indicating shifts from baseline (low/normal/high) to post first treatment (low/normal/high) as well as a listing of subjects experiencing shifts from low/normal to high
- Shift tables indicating shifts from baseline to post first treatment potentially clinically significant result as well as a listing of subjects experiencing such shifts. For the definition of pCS high and pCS low, see Section 13.2.2.

Only subjects with non-missing pre and post values will be considered for the analysis.

11.1.5.15. Lower Extremity Function Scale (LEFS)

The analysis of LEFS is restricted to subjects with at least one TW with lower extremity involvement. The LEFS scores at 12 months follow up assessment will be presented by treatment group using descriptive statistics. The mean LEFS scores between treatment groups will be compared by a one-way analysis of variance.

11.1.5.16. Disabilities of the Arm, Shoulder and Hand (QuickDASH)

The analysis of QuickDASH is restricted to subjects with at least one TW with arm, shoulder or hand involvement. The QuickDASH scores at 12 months follow up assessment will be

presented by treatment group using descriptive statistics. The mean QuickDASH scores between treatment groups will be compared by a one-way analysis of variance.

11.1.5.17. Range of Motion (ROM)

The ROM will be evaluated using a variable with 2 levels (all measurements normal / at least one abnormal finding). Summary tables will be provided giving sample size, absolute and relative frequency and 95% CI (Confidence Interval) for the proportion with a normal score, separately for each study treatment group.

A logistic regression model will be applied for testing the statistical significance of the difference in ROM measurements between NexoBrid and SoC treatments at 12 months. Adjustment will be made for any imbalance between the groups in treatment center, overall treated %TBSA and % of FT per patient.

Only subjects with joint injury will be regarded for this analysis. A footnote to the results will indicate this.

11.1.5.18. EQ-5D (Quality of Life)

The 12 months follow up assessment of the EQ-5D will be analyzed as follows.

The EQ-5D consists of 5 items with 3 possible answers each and an EQ visual analogue scale. For each item (Mobility, Looking after myself, Doing usual activities, Having pain or discomfort, Feeling worried, sad or unhappy) the count and percentages of each possible answer will be presented per age group (the analysis will be done per age at randomization) and per treatment group. The difference between treatment groups will be tested using a chi square test (or Fishers' exact test in case of small expected cell counts).

The values of the EQ visual analogue scale (EQ VAS) will be presented by treatment group using descriptive statistics. The mean EQ VAS values between treatment groups will be compared by a one-way analysis of variance.

Additionally, the single summary index will be analyzed analogously as the EQ VAS.

The linear regression model with the single summary index as target variable will be applied for testing the statistical significance of the difference in scores between NexoBrid and SoC treatments at 12 months. Adjustment will be made for any imbalance between the groups in treatment center and overall treated %TBSA.

11.1.5.19. Burn Outcome Questionnaire (BOQ)

The BOQ will be analyzed separately. Its analysis will be detailed in a separate document.

11.1.6. Exploratory Analyses

11.1.6.1. Incidence of surgical Escharotomy procedures on circumferential extremities target wounds

Two analyses will be done for this exploratory endpoint. The number and percentage of circumferential extremities target wounds that went through escharotomy will be presented by treatment group. The proportion of circumferential extremities target wounds that went through escharotomy will be compared between treatment groups using a chi square test (or Fisher's exact test in case of small estimated cell counts).

The number and percentage of patients having at least one circumferential extremities target wound, which went through escharotomy, will be presented by treatment group. The proportion of patients having at least one circumferential extremities target wound which went through escharotomy will be compared between treatment groups using a chi square test (or Fisher's exact test in case of small estimated cell counts). The reference count will be the number of subjects with at least one circumferential extremities target wound.

11.1.6.2. Incidence of reduction in interstitial/compartment pressure in circumferential extremities target wounds

This is a binary (yes/no) variable indicating whether the patient experienced reduction in interstitial/compartment pressure in circumferential extremities target wounds for each target wound and each procedure separately. Rates of patients with this event will be analyzed by displaying counts and percentages by treatment group. The proportion of patients with reduction in interstitial/compartment pressure in circumferential extremities target wounds will be compared between treatment groups using a chi square test (or Fisher's exact test in case of small estimated cell counts).

11.1.6.3. % wound area surgically excised for eschar removal

The % wound area surgically excised for eschar removal per patient (computed as in section 10.3.5.3) will be presented by treatment group using descriptive statistics and compared between groups with a one-way analysis of variance.

11.1.6.4. Incidence of surgically harvested donor site wounds

This is a counting variable, giving the number of donor site wounds per patient. Rates of patients with different numbers of donor site wounds will be analyzed by displaying counts and percentages by treatment group. The number of donor site wounds will be compared between treatment groups using a chi square test (or Fisher's exact test in case of small estimated cell counts).

11.1.6.5. % area of surgically harvested donor site wounds

The % wound area of surgically harvested donor site wounds per patient (computed as in section 10.3.5.5) will be presented by treatment group using descriptive statistics and compared between groups with a one-way analysis of variance.

11.1.6.6. Blood loss using changes in hemoglobin following eschar removal procedures

The measure of blood loss defined in Section 10.3.5.6 will be computed for each patient, presented by treatment group using descriptive statistics and compared between groups with a one-way analysis of variance.

11.1.6.7. POSAS

The POSAS will be evaluated on a subject level. The POSAS based on target wounds will be calculated as outlined in Section 10.3.5.7 and will be presented by treatment group using descriptive statistics. The mean POSAS values between treatment groups will be compared by a one-way analysis of variance.

The analysis will be repeated for the patient scale and for the observer scale separately.

The analyses will be conducted for the POSAS assessments (both on-site and remotely collected) of 6 and 12 weeks and 6 and 12 months.

In addition, the analysis will be repeated excluding subjects with remotely collected data.

11.1.6.8. MVSS (target wounds)

The MVSS for target wounds (analyzed on a subject level) at 6 and 12 weeks and 6 months (both on-site and remotely collected) will be calculated as outlined in Section 10.3.4.2 and will be presented by treatment group using descriptive statistics. The mean MVSS values between treatment groups will be compared by a one-way analysis of variance.

In addition, the analysis will be repeated excluding subjects with remotely collected data.

11.1.6.9. Duration of hospitalization

The duration of hospitalization will be analyzed descriptively with Kaplan-Meier plots and compared between groups with a log rank test.

11.1.7. Pharmacokinetic Variables

Pharmacokinetic variables will be analyzed separately. The analysis is described in a separate document.

11.1.8. Multicenter Data

Randomization will be stratified by center as detailed in the Section 15.7.2 of the protocol.

The analyses of the primary and secondary endpoints of this trial will adjust for possible center effects by including the grouped center as covariate in the analyses. The grouping will be done using the geographical regions: US (United States), EEU (Eastern Europe), WEU (Western Europe) and India.

11.1.9. Handling of Multiple Comparisons

Unless otherwise specified, all statistical tests will be conducted against a two-sided alternative hypothesis, employing a significance level of 0.05. The primary and secondary endpoints of efficacy will be tested in a hierarchy, so as to preserve the nominal significance levels:

1. Primary: time to complete eschar removal
2. Secondary: Incidence of surgical excision
3. Secondary: Blood loss
4. Secondary: Area of Autograft in DPT wounds

5. Secondary: Incidence of Autograft in DPT wounds

11.1.10. Interim Analysis

No interim analyses will be performed.

11.1.11. Stratification

The following factors are used to stratify the study design:

A. Age group:

A1: 0 to 23 month

A2: 24 months to 3 years

A3: 4 to 11 years

A4: 12 to 18 years

B. Overall total area of burns measured by % of TBSA

B1: (1% to 15%), inclusive

B2: (> 15% to ≤ 30%)

C. The proportion of the FT area out of the total burned area of a patient

C1: < 20% of the burned area defined as FT

C2: ≥ 20% of the burned area defined as FT

D. Clinical center, so as to address possible differences in SoC procedures between sites (if any), to allow comparison of the results within center and to reduce the likelihood of a center including patients in only one arm of the study.

D_1, D_2,…,D_m (m being the number of participating clinical centers).

Using this stratification, patients will be randomized in a 1:1 ratio (NexoBrid: SoC).

11.1.12. Analysis of Subgroups

The following subgroup analyses were pre planned and will be conducted to explore efficacy and safety within certain pre-defined subgroups of the population. All analyses presented here are of exploratory nature.

11.1.12.1. Age group 0-23 months

A subgroup analysis will be performed on the subject aged 0-23 months.

The primary endpoint time to eschar removal will be analyzed with Kaplan-Meier curves to display the distribution of time to complete eschar removal under the two treatments. Time to complete eschar removal will be compared between treatment groups by a log rank test.

For the secondary endpoint “incidence of surgical excision” counts and percentages of patients needing surgical excision will be presented. The treatment groups will be compared using a chi square test (of Fisher’s exact test in case of small expected cell counts).

For the secondary endpoint “blood loss” descriptive statistics will be presented by treatment group. Treatment groups will be compared using a t-test.

For the secondary endpoint “area of Autograft in DPT wounds” descriptive statistics (on a subject level) will be presented by treatment group. Treatment groups will be compared using at-test.

For the secondary endpoint “incidence of Autograft in DPT wounds” counts and percentages on a target wound level will be presented by treatment group. The treatment groups will be compared using a logistic regression within the generalized estimating equations (GEE) framework that accounts for within-subject correlation between target wounds.

The analysis of Adverse Events for the sub-period until the 12 weeks follow-up, (see Section 11.1.5.4) will be repeated, as deemed relevant, on the subgroup of patients aged 0-23 months.

The analyses of PTT/INR (see Section 11.1.5.14) will be repeated on the subgroup of patients aged 0-23 months.

The analyses of central laboratory parameters (see Section 11.1.5.5) will be repeated, as deemed relevant, on the subgroup of patients aged 0-23 months. The following parameters will be analyzed:

- Biochemistry: Creatinine, Glucose, Total Bilirubin, Albumin, Calcium (Ca++), Sodium (Na+), Potassium (K+), Alkaline Phosphatase, SGOT (ASAT), SGPT (ALAT), LDH.
- Hematology: Hemoglobin, Hematocrit, Platelets, Total WBC, Neutrophils, Lymphocytes.

11.1.12.2. Age group 24 months – 3 years

The same analyses as outlined in Section 11.1.12.1 will be performed but on the subgroup of patients aged 24 months – 3 years.

11.1.12.3. Age group 0 – 3 years

The same analyses as outlined in Section 11.1.12.1 will be performed but on the subgroup of patients aged 0 – 3 years.

11.1.12.4. Age group 4 – 11 years

The same analyses as outlined in Section 11.1.12.1 will be performed but on the subgroup of patients aged 4 – 11 years.

11.1.12.5. Age group 12 – 18 years

The same analyses as outlined in Section 11.1.12.1 will be performed but on the subgroup of patients aged 12 – 18 years

11.2. Second stage analysis – 24 months long-term follow-up (LTFU24)

11.2.1. Demographics Data and Other Baseline Characteristics

Not applicable for this stage of analysis since all information on demographics data and other baseline characteristics was already analyzed in the first stage analysis (see 11.1.1).

11.2.2. Primary Efficacy Analysis

Not applicable for this stage of analysis since all information on the primary efficacy variable was already analyzed in the first stage analysis (see 11.1.3).

11.2.3. Secondary Efficacy Analyses

Not applicable for this stage of analysis since all information on the secondary efficacy variables was already analyzed in the first stage analysis (see 11.1.4).

11.2.4. Safety Analyses

11.2.4.1. Time to Reach Complete Wound Closure

Not applicable for this stage of analysis since all information on this endpoint was already analyzed in the first stage analysis (see 11.1.5.1).

11.2.4.2. Cosmesis and Function at 12 Months from Wound Closure

Not applicable for this stage of analysis since this endpoint was already analyzed in the first stage analysis (see 11.1.5.2).

11.2.4.3. Cosmesis and Function at 24 Months from Wound Closure

The MVSS score for a patient will be calculated as described in Section 10.3.4.2. Descriptive statistics of MVSS score at 24 months will be provided by treatment group, for both observed and imputed data (for handling of missing data see Section 10.1.4.2). The treatment groups will be compared using a linear model with imputed MVSS score at 24 months as the dependent variable. The explanatory variables in the model will include treatment, and the following variables: treatment center, age, %TBSA, proportion of the FT area of a patient and

number of TWs (1,2,3-4, ≥5). The coefficient corresponding to the treatment group will be estimated and will represent the estimated mean difference in MVSS score at 24 months between NexoBrid and SoC adjusted for any imbalance in the stratification factors. The results of the regression analysis will be provided.

A clinically meaningful difference will be incorporated into the analysis that will represent a 1.9 or more units advantage to the SoC group, so that the analysis will show whether the NexoBrid group is estimated to have on average a MVSS score that is not worse by more than 1.9 units than the score for the SoC group. That is, the following hypothesis are about to be tested:

$$H_0: \Delta \geq 1.9 \quad vs. \quad H_1: \Delta < 1.9$$

Where Δ is the mean difference in MVSS score between NexoBrid and SoC adjusted for any imbalance in the stratification factors. These hypotheses will be tested by comparing the 95% confidence interval for the coefficient of the treatment group from the linear model outlined above to the clinically meaningful difference 1.9. The null hypothesis will be rejected if the upper bound of this interval is smaller than 1.9.

The main analysis will be performed as described above using both on-site and remotely collected evaluations, with missing values imputed as described in SAP Section 10.1.4.2.

In addition, a sensitivity analysis will be performed (similar to the main analysis described above) to explore the possible influence of remotely captured data. For this sensitivity analysis, the remote evaluations will be "corrected", as described in SAP Section 10.5.2.2, and missing values at 24 months will be imputed applying the imputation procedure from SAP Section 10.1.4.2 on observed *on-site* evaluations and "corrected" *remote* evaluations.

11.2.4.4. Adverse Events

The analyses of Adverse Events described in Section 11.1.5.4 will be repeated, as deemed relevant, on the following additional sub-periods:

- adverse events occurring after the 12 months follow up (including AEs from the first stage that are still ongoing);
- adverse events occurring during the entire study period

11.2.4.5. Concomitant Medications

In this analysis, only the concomitant medications given after the 12 months follow up (including CM with a start date before the 12 months follow up but an end date after that or still ongoing) will be analyzed.

The number of concomitant medication and patients involved will be tabulated with overall numbers and with reference to ATC coding using the Level 2 and Level 4 codes.

11.2.4.6. Lower Extremity Function Scale (LEFS)

The LEFS scores at 24 months follow up assessment will be presented by treatment group using descriptive statistics. The mean LEFS scores between treatment groups will be compared by a one-way analysis of variance.

11.2.4.7. Disabilities of the Arm, Shoulder and Hand (QuickDASH)

The QuickDASH scores at 24 months follow up assessment will be presented by treatment group using descriptive statistics. The mean QuickDASH scores between treatment groups will be compared by a one-way analysis of variance.

11.2.4.8. Range of Motion (ROM)

The ROM will be evaluated using a variable with 2 levels (all measurements normal / at least one abnormal finding). Summary tables will be provided giving sample size, absolute and relative frequency and 95% CI (Confidence Interval) for the proportion with a normal score, separately for each study treatment group.

A logistic regression model will be applied for testing the statistical significance of the difference in ROM measurements between NexoBrid and SoC treatments at 24 months. Adjustment will be made for any imbalance between the groups in treatment center, overall treated %TBSA and % of FT per patient.

Only subjects with joint injury will be regarded for this analysis. A footnote to the results will indicate this.

11.2.4.9. EQ-5D (Quality of Life)

The 24 months follow up assessment of the EQ-5D will be analyzed as follows.

The EQ-5D consists of 5 items with 3 possible answers each and an EQ visual analogue scale. For each item (Mobility, Looking after myself, Doing usual activities, Having pain or discomfort, Feeling worried, sad or unhappy) the count and percentages of each possible answer will be presented per age group (the analysis will be done per age at randomization) and per treatment group. The difference between treatment groups will be tested using a chi square test (or Fishers' exact test in case of small expected cell counts).

The values of the EQ visual analogue scale (EQ VAS) will be presented by treatment group using descriptive statistics. The mean EQ VAS values between treatment groups will be compared by a one-way analysis of variance.

Additionally, the single summary index will be analyzed analogously as the EQ VAS.

The linear regression model with the single summary index as target variable will be applied for testing the statistical significance of the difference in scores between NexoBrid and SoC treatments at 24 months. Adjustment will be made for any imbalance between the groups in treatment center and overall treated %TBSA.

11.2.5. Exploratory Analyses

11.2.5.1. POSAS

The POSAS will be evaluated on a subject level. The POSAS for target wounds will be calculated as outlined in Section 10.3.5.7 and will be presented by treatment group using descriptive statistics. The mean POSAS values between treatment groups will be compared by a one-way analysis of variance.

The analysis will be repeated for the patient scale and for the observer scale separately.

The analyses will be conducted for the POSAS assessments (both on-site and remotely collected) of 18 and 24 months.

In addition, the analysis will be repeated excluding subjects with remotely collected data.

11.2.5.2. MVSS (target wounds)

The MVSS for target wounds (analyzed on a subject level) at 18 months (both on-site and remotely collected) will be calculated as outlined in Section 10.3.4.2 and will be presented by

treatment group using descriptive statistic. The mean MVSS values between treatment groups will be compared by a one-way analysis of variance.

In addition, the analysis will be repeated excluding subjects with remotely collected data.

11.2.6. Pharmacokinetic Variables

Not applicable. All pharmacokinetic data was already analyzed in the first stage analysis (see Section 11.1.7).

11.2.7. Multicenter Data

See Section 11.1.8.

11.2.8. Handling of Multiple Comparisons

Not applicable. All testing done during this second stage analysis is exploratory in nature. So no adjustment for multiplicity is needed.

11.2.9. Interim Analysis

No interim analyses will be performed.

11.2.10. Stratification

See Section 11.1.11.

11.2.11. Analysis of Subgroups

No subgroup analyses are planned in this stage.

12. Reference List

[1] B. Atiyeh, G. S.W. and S. Hayek, "State of the art in burn treatment," *World J Surg,* vol. 29, no. 2, pp. 131-48, 2005.

[2] J. Barret and D. Herndon, "Modulation of inflammatory and catabolic responses in severely burned children by early burn wound excision in the first 24 hours," *Arch Surg,* vol. 138, no. 2, pp. 127-32, Feb 2003.

[3] J. Barret et al., "Total Burn Wound Excision Of Massive Paediatric Burns in The First 24 Hours Post-Injury," *Annals of Burns and Fire Disasters,* vol. 12, no. 1, pp. 25-27, 1999.

[4] Y. Ong, M. Samuel and C. Song, "Meta-analysis of early excision of burns," *Burns,* vol. 32, no. 2, pp. 145-50, 2006.

[5] W. Xiao-Wu et al., "Effects of delayed wound excision and grafting in severely burned children," *Arch Surg,* vol. 137, no. 9, pp. 1049-54, 2002.

[6] D. Young, "Burn and Electrical Injury," *Second ed. Plastic Surgery,* vol. 1, pp. 811-833, 2006.

[7] S. Hindocha, "Burns eschar debridement: A review.," *Journal of Wound Technology,* vol. 21, pp. 16-19, 07 2013.

[8] S.-J. Edmondson, I. Jumabhoy and A. Murray, "Time to start putting down the knife: A systematic review of burns excision tools of randomised and non-randomised trials," *Burns,* vol. 44, 02 2018.

[9] American Burn Association, "National Burn Repository® Version 12.0," 2016.

[10] L. Rosenberg, Y. Shoham, Y. Krieger, G. Rubin, F. Sander, J. Koller, K. David, D. Egosi, R. Ahuja and A. Singer, "Minimally invasive burn care: a review of seven clinical studies of rapid and selective debridement using a bromelain-based debriding enzyme (Nexobrid®)," *Ann Burns Fire Disasters,* vol. 28, no. 4, pp. 264-274, 2015.

[11] L. Rosenberg, Y. Krieger, A. Bogdanov-Berezovski, E. Silberstein, Y. Shoham and A. Singer, "A novel rapid and selective enzymatic debridement," *Burns,* vol. 40, pp. 466-47, 2014.

[12] Shoham et al, "The Use of Rapid Enzymatic Debridement in the 2015 Romanian Burn Mass Casualty Incident: Ground Breaking Experience," in *SAWC Spring meeting*, San Diego, 2017.

[13] S. Fischer et al., "Feasibility and safety of enzymatic debridement for the prevention of operative escharotomy in circumferential deep burns of the distal upper extremity," *Surgery,* vol. 165, no. 6, pp. 1100-1105, June 2019.

[14] American Burn Association, "American Burn Association Consensus Statements," *Journal of Burn Care & Research,* July 2013.

[15] D. Hart, S. Wolf, R. Beauford, S. Lal, D. Chinkens and D. Herndon, "Determinants of blood loss during primary burn excision," *Surgery,* vol. 130, no. 2, pp. 396-402, Aug 2001.

[16] C. Licht, New methods for generating significance levels from multiply-imputed data, 2011.

[17] G. Molenberghs and M. G. Kenward, Missing Data in Clinical Studies, John Wiley & Sons, Ltd , 2008.

[18] C. Mallinckrodt, „Missing Data in Clinical Trails,“ in *Handbook of Missing Data Methodology*, CRC Press, 2015, pp. 499-522.

[19] M. Levine and W. K. D. Shao, "Monitoring of international normalized ratios: comparison of community nurses with family physicians," *Can Fam Physician,* vol. 58, no. 8, p. e465-e471, 2012.

[20] M. van Reen, B. Janssen, M. Oppe, S. Kreimeier and W. Greiner, "EQ-5D-Y user guide," August 2014. [Online]. Available: https://euroqol.org/wp-content/uploads/2019/10/EQ-5D-Y-User-Guide.pdf. [Accessed 22 April 2020].

[21] J. W. Shaw, J. A. Johnson and S. J. Coons, "US Valuation of the EQ-5D Health States - Development and Testing of the D1 Valuation Model," *Medical Care,* vol. 43, no. 3, March 2005.

[22] D. Binder, "Fitting Cox's proportional hazards models from survey data," *Biometrika,* vol. 79, pp. 139-147, 1992.

[23] T. K. Mattoo, "Definition and diagnosis of hypertension in children and adolescents (UpToDate)," [Online]. Available: https://www.uptodate.com/contents/definition-and-diagnosis-of-hypertension-in-children-and-adolescents?search=blood%20pressure%20children&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1. [Accessed 1 June 2020].

[24] N. Nishimura and N. Hiranuma, "Respiratory changes after major burn injury," *Crit Care Med,* vol. 10, no. 1, pp. 25-28, 1982.

[25] Mayo Clinic, "Hypothermia.," [Online]. Available: https://www.mayoclinic.org/diseases-conditions/hypothermia/symptoms-causes/syc-20352682. [Accessed 17 April 2020].

[26] E. Mann-Salinas, M. Baun, J. Meininger and e. al., "Novel predictors of sepsis outperform the American Burn Association sepsis criteria in the burn intensive care unit patient," *J Burn Care Res,* vol. 34, no. 1, pp. 31-43, 2013.

[27] D. Greenhalgh, J. Saffle, J. 4. Holmes and e. al., "American Burn Association consensus conference to define sepsis and infection in burns," *J Burn Care Res,* vol. 28, no. 6, pp. 776-790, 2007.

[28] National Cancer Institute (NCI), "Common Terminology Criteria for Adverse Events (CTCAE) version 5.0," 2017. [Online]. Available: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf. [Accessed 31 May 2020].

[29] R. Bellomo, C. Ronco, J. Kellum, R. Mehta and P. Palevsky, "Acute Dialysis Quality Initiative workgroup. Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus

Conference of the Acute Dialysis Quality Initiative (ADQI) G," *Crit Care,* vol. 8, no. 4, pp. R204-R212, 2004.

[30] P. Thornton, C. Stanley, D. De Leon and e. al., "Recommendations from the Pediatric Endocrine Society for Evaluation and Management of Persistent Hypoglycemia in Neonates, Infants, and Children," *J Pediatr,* vol. 167, no. 2, pp. 238-245, 2015.

[31] C. Holm, F. Hörbrand, M. Mayr, G. von Donnersmarck und W. Mühlbauer, „Acute hyperglycaemia following thermal injury: friend or foe?," *Resuscitation. ,* Bd. 60, Nr. 1, pp. 71-77, 2004.

[32] A. Laird, P. Miller, P. Kilgo, J. Meredith and M. Chang, "Relationship of early hyperglycemia to mortality in trauma patients," *J Trauma,* vol. 56, no. 5, pp. 1058-1062, 2004.

[33] G. Mecott, A. Al-Mousawi, G. Gauglitz, D. Herndon and M. Jeschke, "The role of hyperglycemia in burned patients: evidence-based studies," *Shock,* vol. 33, no. 1, pp. 5-13, 2010.

[34] M. Jeschke, R. Micak, C. Finnerty and D. Herndon, "Changes in liver function and size after a severe thermal injury," *Shock,* vol. 28, no. 2, pp. 172-77, 2007.

[35] O. Aguayo-Becerra, C. Torres-Garibay, M. Macías-Amezcua and e. al., "Serum albumin level as a risk factor for mortality in burn patients," *Clinics (Sao Paulo),* vol. 68, no. 7, pp. 940-945, 2013.

[36] M. Çakırca, İ. Sözen, G. Tozlu Bindal, M. Baydar und A. Yastı, „Relationship between the albumin level and the anesthesia method and the effect on clinical course in patients with major burns," *Ulus Travma Acil Cerrahi Derg,* Bd. 25, Nr. 1, pp. 55-59, 2019.

[37] G. Liumbruno, F. Bennardello, A. Lattanzio, P. Piccoli and G. Rossettias, "Italian Society of Transfusion Medicine and Immunohaematology (SIMTI). Recommendations for the use of albumin and immunoglobulins," *Blood Transfus,* vol. 7, no. 3, pp. 216-234, 2009.

[38] G. Lindner and G. Funk, "Hypernatremia in critically ill patients," *J Crit Care,* vol. 28, no. 2, p. 216.e11-216.e2.16E20, 2013.

[39] P. Hébert, G. Wells, M. Blajchman and e. al., "A multicenter, randomized, controlled clinical trial of transfusion requirements in critical care," *N Engl J Med,* vol. 340, no. 6, pp. 409-17, 1999.

[40] T. Palmieri, "Burn injury and blood transfusion," *Curr Opin Anaesthesiol,* vol. 32, no. 2, pp. 247-251, 2019.

[41] J. Teruya, "Red blood cell transfusion in infants and children: Indications (UpToDate)," [Online]. Available: https://www.uptodate.com/contents/red-blood-cell-transfusion-in-infants-and-children-indications. [Accessed 31 March 2020].

[42] ,. E. Demidowicz, J. Moppert, A. Nowacka, J. Styczynski and M. Wysocki, "Essential thrombocythemia and reactive thrombocytosis in children," *EC Paediatrics,* vol. 2, no. 2, pp. 107-115, 2016.

[43] R. Dinsdale, A. Devi, P. Hampson and e. al., "Changes in novel haematological parameters following thermal injury: A prospective observational cohort study," *Sci Rep,* vol. 7:3211, 2017.

[44] S. Sevitt, "Eosinophil and other leucocyte changes in burned patients, with special reference to adreno-cortical activity," *Br Med J,* vol. 1, no. 4713, pp. 976-983, 1951.

[45] N. Cosgriff, E. Moore, A. Sauaia, M. Kenny-Moynihan, J. Burch and B. Galloway, "Predicting life-threatening coagulopathy in the massively transfused trauma patient: hypothermia and acidoses revisited," *J Trauma,* vol. 42, no. 5, pp. 857-862, 1997.

[46] J. McCullogh, Transfusion Medicine, 4th Edition ed., Chincester, West Sussex: John Wiley & Sons, 2017.

[47] R. Kagan, D. Arenholtz, W. Hickerson, J. Holmes, R. Korentager, J. Kraatz and G. Kotoski, "American Burn Association white paper Surgical Management of the Burn Wound and Use of Skin Substitutes," 2009.

[48] G.D. Warden, J.R. Saffle and M. Kravitz, "A Two-stage Technique for Excision and Grafting of Burn Wounds," *J Trauma,* vol. 22, no. 2, pp. 98-103, Feb 1982.

13. Appendices

13.1. Randomization Algorithm

Treatment allocation is performed using the method of minimization with stratification by the factors A, B, C and D described in Section 11.1.11 above: age group (A), overall %TBSA of all TW wounds (B), overall % of FT area (C) and clinical center (D).

At any point in the trial a new patient is to be entered into the trial, the four above described factors have to be collected for that particular patient. The value for each factor of this patient is denoted by Ai, Bi, Ci and Di, respectively, where i is corresponding to the value of each factor, which is i = 1, 2, 3, 4 for factor A, i = 1, 2 for factors B and C and i = 1,..., m for factor D, with m being the number of participating clinical centers.

For example: a patient, of 25 months of age with overall total area of burns of 4% - 9% and a proportion of the FT area of ≥20% is to be recruited in the 5th clinical center. This would lead to (A2, B1, C2, D5).

The different treatment groups are denoted by j where j=1 corresponds to the NexoBrid group, j=2 to the SoC group. The numbers of patients in each stratum A, B, C, D receiving the j-th treatment and having the i-th factor are denoted by Aij, Bij, Cij and Dij.

When a new patient is to be allocated a treatment, a score for each treatment group j is determined based on the patient's specific factors and on the number of patients already allocated to the treatments. This score is denoted by Nj and is calculated by the sum of Aij, Bij, Cij and Dij. In our example the new patient has the factors A2, B1, C2, D5, hence the numbers of patients for each of these factors in treatment group j=1 which is the NexoBrid treatment are A21, B11, C21, and D51 respectively and the score N1 can be calculated as N1 = A21 + B11 + C21+D51. The score for the SoC treatment N2 can be calculated similarly: N2 = A22 + B12 +C22+D52. To determine which treatment has to be allocated to the next patient, these two scores are compared.

The trial started with the following method of treatment allocation. Since a 1:1 randomization is used in this trial, the patient is allocated to the treatment group with the smallest value among these two score values (strict minimization). Thus, the allocation decision is based on min (N1,

N2). If N1 equals N2, then simple randomization between the treatments is used with probabilities p1=p2=0.5 for each treatment.

On 19th May 2017 it was decided to modify the method of treatment allocation in the following manner. The strict minimization method (described above) assigned the treatment to a patient so as to reduce the imbalance between the treatment groups at any stage. Only if allocating to either treatment yielded the same level of imbalance would the treatment be assigned at random (with a probability of 0.5). A modification is that the algorithm would henceforth assign with probability 0.9 the treatment that reduces the imbalance and with probability 0.1 the treatment that increases the imbalance (minimization with a random element). As before, if both treatment allocations lead to the same imbalance, randomization with probability 0.5 will be used. Thus, using the same notation as above, if N1 < N2 then treatment 1 is allocated with probability 0.9 and treatment 2 with probability 0.1. Conversely, if N1 > N2 then treatment 2 is allocated with probability 0.9 and treatment 1 with probablitiy 0.1. If N1 = N2, then treatment 1 is allocated with proability 0.5 and treatment 2 with probability 0.5.

13.2. Potentially Clinically Significant Thresholds

13.2.1. Vital Signs

The burn injury and any treatments administered before study treatment may affect body systems and their regulatory functions, including heart rate, blood pressure, fluid balance, and temperature. Therefore, measurements of vital signs might be abnormal prior to study treatment administration. In addition, only a few hours elapse between baseline and the following time point when vital signs are measured.

In order to capture significant changes in vital signs that might be related to treatment, potentially clinically significant (PCS) thresholds for vital signs were defined using both the numerical value of the parameter and a clinically significant change from baseline.

Vital Sign (unit)	**Criterion value**	**Explanation**
Systolic Blood Pressure (mmHg)	Low PCS: < age defined threshold and decrease >10%	Low PCS is based on PALS Guidelines depending upon age: Term neonates (0 to 28 days): The systolic BP is < 60 mmHg Infants (1 to 12 months): Systolic BP is < 70 mmHg Children 1 to 10 years (5th BP percentile): Systolic BP is < 70 mmHg + (age in years x 2) Children > 10 years: Systolic BP is < 90 mmHg
	High PCS:≥ 20% above the 95th percentile for age/gender and increase >10%	High PCS is based on the normal blood pressure levels for boys and girls by age and height percentile. The 95th percentile is taken for the 50th height percentile [23]. 95th percentile SBP for Girls (mmHg): 1y-103 2y-106 3y-108 4y-109 5y-110 6y-111 7y-112 8y-113 9y-114 10y-116 11y-118 12y-122 13y-124

		14y-125 15y-126 16y-127 17y-127 95th percentile SBP for Boys (mmHg): 1y-103 2y-106 3y-107 4y-108 5y-109 6y-111 7y-112 8y-114 9y-115 10y-116 11y-118 12y-121 13y-125 14y-130 15y-132 16y-134 17y-135 The increase from baseline for High PCS is used to capture significant changes, since elevated blood pressure might be

		present prior to study treatment due to effects of the burn injury and any pre-study treatments.
Respiratory Rate (breaths/min)	Low PCS: <age defined threshold and decrease >10%	Low Respiratory Rate PCS is based on PALS Guidelines depending upon age: Infant, 30 Toddler, 24 Preschooler, 22 School-aged, 18 Adolescent, 12 Respiration can be compromised following burn injury [24].
	High PCS: > age defined threshold and increase >10%	High Respiratory Rate PCS is based on PALS Guidelines depending upon age: Infant, 60 Toddler, 40 Preschooler, 34 School-aged, 30 Adolescent, 16
Temperature (°C)	Low PCS: <36 and decrease >0.5	Temperature <36°C was used in the NexoBrid protocols (eg, CIDS) to define hypothermia. A temperature of 35°C is commonly used in the definition of hypothermia [25], which is lower than the PCS threshold and therefore encompassed by it. Temperature <36°C was found to be useful as a predictor for sepsis [26]

	High PCS: >39 and increase >0.5	Burn patients are very often hyperthermic. Pyrexia was defined as >39°C in the NexoBrid protocols (eg, CIDS). Temperature >39°C was found to be useful as a predictor for sepsis [27]
Heart rate (beats/min)	Low PCS: < lower threshold for age while awake and decrease >10%	Normal Pulse Rate according to PALS will be defined depending upon age, while awake: Newborns up to 3 months old, 85–205 Age 3 months to 2 years old, 100–190 Ages 2 to 10, 60–140 Children aged 10 years and above, 60–100
	High PCS: > upper threshold for age while awake and increase >10%	Normal Pulse Rate according to PALS will be defined depending upon age, while awake: Newborns up to 3 months old, 85–205 Age 3 months to 2 years old, 100–190 Ages 2 to 10, 60–140 Children aged 10 years and above, 60–100

13.2.2. Laboratory Parameters

Burn injury affects almost all body systems. Therefore, a wide range of laboratory parameters may be affected, which may be abnormal prior to study treatment. In addition, only a few hours elapse between baseline and the following time point when laboratory parameters are measured.

For these reasons, in order to capture significant changes in laboratory parameters that might be related to treatment, PCS thresholds for laboratory parameters were defined using both the numerical value of the parameter and a clinically significant change from baseline.

Many of the PCS definitions for laboratory values below are based on Grade 2 definitions from the CTCAE version 5.0 [28]. As in the CTCAE, multiplication of the baseline value is used to define high PCS for several parameters in order to capture significant changes from baseline, as the lab result may be abnormal at baseline.

Test (unit)	**Criterion value**	**Explanation**
Serum Chemistry		
Creatinine (umol/L)	High PCS: >1.5x ULN if Baseline ≤ ULN or >1.5x Baseline value if Baseline > ULN.	Acute kidney injury is a common complication in burn patients. Early kidney damage in the burn patient might be due to hypovolemia, inflammatory processes, tissue destruction and release of denaturated proteins, iatrogenic causes (eg, nephrotoxic agents), and/or cardiac dysfunction. The High PCS definition is based on the definition of Stage 1 acute kidney injury and on the lower limit of the CTCAE Grade 2 definition [28]. The approach of using multiplication of baseline value when baseline is already elevated is used in the CTCAE Grade 2 definition and in the RIFLE risk definition in the Acute Dialysis Quality Initiative (ADQI) criteria for the definition and classification of AKI [29].
Glucose (mmol/L)	Low PCS: <3.00 mmol/L	Low PCS is based on the upper limit of CTCAE definition of Grade 2 hypoglycemia [28] [30]
	High PCS: >11.10 mmol/ L and increase ≥2.78 mmol/L	Hyperglycemia is common in burn patients [31]. High glucose values >200 mg/dL (>11.10 mmol/L) are related to higher rates of infection and mortality among burn and trauma patients [32] [33]. In addition, glucose >200 mg/dL is one of

		the triggers for the definition of sepsis according to the American Burn Association [27]
Total Bilirubin (umol/L)	High PCS: >1.5x ULN if Baseline ≤ ULN or >1.5x Baseline value if Baseline > ULN	Based on the lower limit of the CTCAE Grade 2 definition [28].
Alkaline Phosphatase (U/L)	High PCS: >2.5x Age/ Gender ULN if Baseline ≤ ULN or >2.5x Baseline value if Baseline > Age/Gender ULN	Based on the lower limit of the CTCAE Grade 2 definition [28], and age normal range U/L of study laboratory: 7-365 days: 50-270 1-2y: 50-270 3-12y: 60-415 13-15y male: 60-500, 13-15y female: 60-350 16-19y male: 30-225, 16-19y female: 30-165
SGPT (ALT) (U/L)	High PCS: >3x ULN if Baseline ≤ ULN or >3.0x Baseline value if	Based on the lower limit of the CTCAE Grade 2 definition [28]. Alanine aminotransferase (ALT) levels can increase immediately after burn injuries [34] and therefore might be increased already at baseline.

	Baseline > ULN	
SGOT (AST) U/L	High PCS: >3x ULN if Baseline ≤ ULN or >3.0x Baseline value, if Baseline > ULN	Based on the lower limit of the CTCAE Grade 2 definition. Aspartate aminotransferase (AST) levels can increase immediately after burn injuries [34] and therefore might be increased already at baseline.
Albumin (g/L)	Low PCS: ≤25 g/L and Decrease ≥10%	Hypoalbuminemia is common in burn patients. Hypoalbuminemia might occur in burn patients due to protein loss, an inflammatory process that affects plasma protein synthesis in the liver, and/or higher vascular permeability in burn wounds that causes exudation with protein loss. The CTCAE definition of Grade 2 hypoalbuminemia is albumin levels between 2 to 3 g/dL (20-30 g/L) [28]. Since the lower limit of normal laboratory ranges of the study is around 3g/dL (7days-2y: 3-5g/dL, 3-199y: 3.2-5g/dL), PCS value lower than normal was taken. In the literature, morbidity corresponds to albumin <20 to 25 g/L [35] [36]. Oncotic activity remains physiologically adequate at albumin ≥2 g/dL. An albumin level ≤20 g/L is commonly agreed by practitioners to trigger administration of intravenous albumin [37].

Calcium (Ca^{++}), (mmol/L)	Low PCS: <2 mmol/L and decrease ≥10%	Please refer to corrected calcium below
	High PCS: >2.9 mmol/L and increase ≥10%	Please refer to corrected calcium below
Corrected Calcium (Ca^{++}), (mmol/L)	Low PCS: <1.75 mmol/L and decrease ≥10%	Low calcium levels are seen in burn patients and might be due to extracellular-intracellular shifts, calcium accumulation in erythrocytes, increased urinary excretion, and/or calcium loss in tissue exudates. Calcium levels are related to albumin levels, which are frequently decreased in burn patients. Low PCS is based on the upper limit of CTCAE definition of Grade 3 hypocalcemia [28].
	High PCS: >2.9 mmol/L and increase ≥10%	High PCS is based on the lower limit of CTCAE definition of Grade 2 hypercalcemia [28].
Sodium (Na^{+}) (mmol/L)	Low PCS: ≤129 and decrease >4	Low PCS is based on the upper limit of CTCAE definition of Grade 2 hyponatremia [28].
	High PCS: >150 and increase ≥5	High PCS is based on the lower limit CTCAE definition of Grade 2 hypernatremia [28]. Studies have shown increased morbidity and mortality for patients with hypernatremia defined as >150 meq/L (mmol/L) [38].

Potassium (K^+) (mmol/L)	Low PCS: <3 and decrease ≥10%	Low PCS is based on the upper limit of CTCAE definition of Grade 3 hypokalemia [28]. Since the normal laboratory ranges of the study are between 3.5-5.5mmol/L (0-2y 3.8-5.5mmol/L, 3-12y 3.5-5.5mmol/L, 13-199y 3.5-5.3mmol/L), PCS value lower than normal corresponding to Grade 3 was taken.
	High PCS: >6.0 and increase ≥10%	High PCS is based on the lower limit of CTCAE definition of Grade 3 hyperkalemia [28]. Since the normal laboratory ranges of the study are between 3.5-5.5mmol/L (0-2y 3.8-5.5mmol/L, 3-12y 3.5-5.5mmol/L, 13-199y 3.5-5.3mmol/L), PCS value higher than normal corresponding to Grade 3 was taken.
Hematology		
Hemoglobin (g/L)	Low PCS: <80 and decrease ≥10%	Low PCS was set to be in line with the limit at which blood transfusion is recommended [39]. According to current clinical recommendations for burn patients, blood transfusion is administered when hemoglobin is below 7.0 to 8.0 g/dl (70-80 g/L) [40]. For most patients, restrictive RBC transfusion strategy (ie, Hgb of 7 g/dL (70g/L) as the transfusion threshold) rather than more liberal transfusion strategies is recommended. Slightly higher limit was taken as PCS since higher Hgb levels may be appropriate in certain conditions [41]
Platelets (10^9/L)	Low PCS: <75 and decrease ≥10%	Low PCS is based on the upper limit of CTCAE Grade 2 definition [28].

	High PCS: >600 and increase >10%	Platelet levels greater than the upper limit of normal but below 700 are regarded as mild thrombocytosis and do not require treatment. Platelet levels >700 are defined as moderate thrombocytosis [42] Increased platelet turnover might be seen in burn patients due to a fall in platelet count followed by an increase in platelet production and a rebound thrombocytosis [43]
Total Leukocyte (10^9/L)	Low PCS: <=20% age related lower threshold	The Low PCS is based on the normal limits of total leukocytes depending upon age; 60-179days: 5-18 180days-1y: 5-17, 2-5y: 5-16, 6-11y: 4.5-13.5, 12-17y: 4.5-13
	High PCS: ≥ 20% age related threshold	The high PCS is based on the normal limits of total leukocytes depending upon age; 60-179days: 5-18 180days-1y: 5-17, 2-5y: 5-16, 6-11y: 4.5-13.5, 12-17y: 4.5-13
Neutrophil (10^9/L)	Low PCS: <1.5	Low PCS is based on the common definition of neutropenia and the upper limit of CTCAE Grade 2 definition [28].
	High PCS: ≥20% upper limit of normal White Cell	Neutrophil count might increase following burn trauma, surgical procedures, or stress [44]

	Count upon age group and increase ≥ 10% from baseline	Normal White Cell Count upon age group: 90-179days: 1-8.5 180-364days: 1.5-8.5 1-199y: 1.8-8
Coagulation		
International Normalized Ratio (INR) ratio	High PCS >1.5 if Baseline is ≤1.2 >1.5x Baseline if Baseline is >1.2	A common normal range of INR was set to be up to 1.2, since normal ranges for local labs were not available for all sites. The PCS is based on the lower limit of the CTCAE Grade 2 definition [28].
Activated Partial Thromboplastin Time (PTT) (sec)	>1.5x ULN if Baseline < ULN >1.5 x Baseline if Baseline > ULN	Derived from the lower limit of CTCAE Grade 2 definition [28]. This limit was used in the DETECT protocol to define a significant increase [45]

13.3. Update to Study Sample Size Due to COVID-19 Pandemic

According to the study protocol version 9.01 dated 29 July 2019, 160 patients were to be enrolled into the study according to the defined age group distribution. Prior to the outbreak of COVID-19 pandemic and until March 2020, 145 patients were enrolled into the study.

The COVID-19 pandemic significantly affected all study sites' ability to enroll additional patients. There was a great uncertainty regarding when restrictions would be lifted and study sites would be able to resume enrollment. Therefore, MediWound explored whether the recruitment to the study can be completed at this stage, with 145/160 (91%) of patients

enrolled, in line with FDA guidance document: “Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency”, issued on 17 June 2020.

The original sample size and statistical power calculations for this study were based on effects reached during the first Phase 3 study (MW2004-11-02) which was completed in 2009. Since then MediWound had also completed the recruitment and the 12 month follow up of 175 adult patients randomized to the Phase 3 study MW2010-03-02 (DETECT). Power calculations were performed also using data from this latest phase 3 study.

The updated sample size calculations, presented below, confirmed that based on a sample size of 145 patients, the study is sufficiently powered to meet the primary and main secondary and safety endpoints as originally planned. Thus, the study provides all required data to evaluate NexoBrid’s safety and efficacy in pediatric patients while enabling the early availability of the product to address an unmet need in the pediatric population. These power calculations were submitted to the pediatric committee (PDCO) within the EMA and to the FDA to support MediWound’s request for early completion of enrollment.

The request for early study completion was endorsed by the study’s DSMB, relying on the confirmation that study endpoints can be met with the sample size of 145 patients and pending FDA and EMA approval.

The request for early completion of enrollment was accepted by both the FDA and the PDCO (EMA) and therefore the enrollment of patients into the study was completed with a total of 145 patients, with the following age distribution:

- 45 patients ≥0 months and ≤23 months old
- 30 patients ≥24 months and ≤3 years old
- 50 patients ≥4 years and ≤11 years old
- 20 patients ≥12 years and <18 years old

Power calculations for MW2012-01-01 study considering a reduced sample size of 145 patients

Sample sizes needed to achieve a power of 90% for the statistical tests in the analysis of MW2012-01-01 (CIDS) are estimated. The calculations are based on the 12 months’ data of

the DETECT study (MW2010-03-02), as well as on data from the previous study MW2004-11-02.

Note, that the power calculated for each test does only refer to the power of that specific test. When conducting multiple tests in a hierarchical testing procedure, endpoints which are positioned at later points of the hierarchy, might not be tested if earlier tests do not reject the corresponding null hypotheses. Therefore, the power of the tests in a hierarchical testing procedure also depends on the other endpoints and will be lower.

Time to Complete Eschar Removal

Estimate based on study MW2004-11-02

The sample size calculation based on the results of study MW2004-11-02 assumes that the proportional hazards assumption will be judged appropriate and consequently a Cox model will be used for the primary analysis:

Estimated log hazard ratio from study MW2004-11-02 (NexoBrid versus SoC) = -1.39, using a Cox proportional hazards model with treatment as only covariate.

Standard error = 0.28

Target log hazard ratio = -1.39 + 0.28/2 = -1.25; target hazard ratio λ= 0.29.

The formula for total number of events D is: $D = \frac{4(z_\alpha + z_\beta)^2}{(\ln \lambda)^2}$

(Piantadosi, Clinical Trials: A Methodologic Perspective, p.169)

z_α = 1.96;

z_β = 1.28 for 90% power.

If the proportion of events in group 1 is denoted by p1 = 0.59

and the proportion of events in group 2 is denoted by p2 = 0.70

which are the observed proportions in the study MW2004-11-02, the number of patients required per group is n = $\frac{2D}{p_1+p_2} = 42$ (to the nearest even number), 21 in each group.

Estimate based on study MW2010-03-02

Since the proportional hazards assumption was rejected in study MW2010-03-02 trial, the sample size calculation based in this study assumes that this will also be the case in the CIDS study and consequently a generalized Wilcoxon-Gehan test will be used in the analysis. Based on the data from MW2010-03-02, simulations were conducted, using the generalized Wilcoxon-Gehan-Test for the primary analysis, resulting in a total sample size of 94 (47 per group) needed to achieve a power of 90%.

% Wound Area Surgically Excised for Eschar Removal

Estimate based on study MW2004-11-02

Estimated difference in mean percent wound excised from study MW2004-11-02 (NexoBrid vs. SoC) = 15.1 – 59.6 = –44.5.

Standard error = 5.3

Target difference Δ = –44.5 + 5.3/2 = –41.85.

The formula for the number of patients per group n is: $n = \frac{2(z_\alpha + z_\beta)^2 \sigma^2}{\Delta^2}$

where

z_α = 1.96;

z_β = 1.28 for a power of 90%;

σ = the estimated common standard deviation of percent area of wound excised = 34.

For 90% power, number per group = 14

Estimate based on study MW2010-03-02

Estimated mean % area surgically excised from study MW2010-03-02: NexoBrid 3.66; SoC 58.74

Standard errors: NexoBrid 2.09; SoC 5.78

Difference: -55.08; Standard error of difference = 5.89

Target difference Δ = -55.08 + 5.89/2 ≈ -52.13.

Using a two-sided two-sample t-test with a two-sided significance level of 5%, the total sample size required to achieve a power of 90% is 24 (12 per group).

Incidence of Surgical Excision

Estimate based on study MW2004-11-02

Estimated proportions having surgical excision from study MW2004-11-02: NexoBrid 0.42; SoC 0.80

Standard errors: NexoBrid 0.057; SoC 0.044

Difference = 0.38; Standard error of difference = 0.073

Anticipated difference: 0.38-0.073/2 ≈ 0.34

We therefore take the anticipated proportion with surgical excision to be: NexoBrid 0.44 and SoC 0.78.

Using Fisher's exact test with a two-sided significance level of 5%, the total sample size required to achieve a power of 90% is 96 (48 per group).

Estimate based on study MW2010-03-02

Estimated proportions having surgical excision from study MW2010-03-02: NexoBrid 0.04; SoC 0.72

Standard errors: NexoBrid 0.023; SOC 0.052

Difference = 0.68; Standard error of difference = 0.057

Anticipated difference: 0.68-0.057/2 ≈ 0.652

We therefore take the anticipated proportion with surgical excision to be: NexoBrid 0.05 and SOC 0.70.

Using Fisher's exact test with a two-sided significance level of 5%, the total sample size required to achieve a power of 90% is 26 (13 per group).

Incidence of Autograft in DPT Wounds (wound level analysis)

Estimate based on study MW2004-11-02

Only subjects with at least one DPT target wound will be eligible for this analysis. The expected proportion of subjects with at least one DPT target wound estimated from study MW2004-11-02 is 62.6%.

Estimated log OR using GEE approach from study MW2004-11-02 (NexoBrid vs SoC): -0.676

Standard Error: 0.40

Anticipated log OR: -0.676+0.40/2=-0.476

Using the normal approximation to the Wald test statistics (see for example Shih (1997)) we obtain that a total sample size of 1150 patients (720 evaluable) is needed to achieve a power of 90% using a two-sided significance level of 5%. This means that a sample size of 575 subjects per arm would be required to obtain the minimum number of evaluable subjects for 90% power in this endpoint.

Estimate based on study MW2010-03-02

Only subjects with at least one DPT target wound will be eligible for this analysis. The expected proportion of subjects with at least one DPT target wound estimated from study MW2010-03-02 is 58%.

Estimated log OR using GEE approach from study MW2010-03-02 (NexoBrid vs SoC): -0.5078

Standard Error: 0.4988

Anticipated log OR: -0.5078+0.4988/2=-0.5078

Using the normal approximation to the Wald test statistics (see for example Shih (1997)) we obtain that a total sample size of 5880 patients (3410 evaluable) is needed to achieve a power of 90% using a two-sided significance level of 5%. This means that a sample size of 2940 subjects per arm would be required to obtain the minimum number of evaluable subjects for 90% power in this endpoint.

Area of Autograft in DPT Wounds (wound level analysis)

Estimate based on study MW2004-11-02

Only subjects with at least one DPT target wound will be eligible for this analysis. The expected proportion of subjects with at least one DPT target wound estimated from study MW2004-11-02 is 62.6%.

Estimated difference in mean area Autograft in DPT wounds using GEE approach from study MW2004-11-02 (NexoBrid vs SoC): -11.41

Standard Error: 4.89

Anticipated difference: -11.41+4.89/2=-8.97

Using the normal approximation to the Wald test statistics (see for example Shih (1997)) we obtain that a total sample size of 486 patients (304 evaluable) is needed to achieve a power of 90% using a two-sided significance level of 5%. This means that a sample size of 243 subjects per arm would be required to obtain the minimum number of evaluable subjects for 90% power in this endpoint.

Estimate based on study MW2010-03-02

Only subjects with at least one DPT target wound will be eligible for this analysis. The expected proportion of subjects with at least one DPT target wound estimated from study MW2010-03-02 is 58%.

Estimated difference in mean area Autograft in DPT wounds using GEE approach from study MW2010-03-02 (NexoBrid vs SoC): -9.39

Standard Error: 6.28

Anticipated difference: -9.39+6.28/2=-6.25

Using the normal approximation to the Wald test statistics (see for example Shih (1997)) we obtain that a total sample size of 1590 patients (922 evaluable) is needed to achieve a power of 90% using a two-sided significance level of 5%. This means that a sample size of 795 subjects per arm would be required to obtain the minimum number of evaluable subjects for 90% power in this endpoint.

Blood Loss

Estimate based on study MW2004-11-02

Estimated mean blood loss from study MW2004-11-02: NexoBrid 421.93 ml; SoC 1024.27 ml

Standard errors: NexoBrid 129.52; SoC 240.98

Difference: -602.33; Standard error of difference: 273.58

Anticipated difference: -602.33 + 267.75 /2 = -468.47

Using a two-sided two-sample t-test with a two-sided significance level of 5%, the total number of evaluable patients needed to achieve a power of 90% is 346 (173 per group). Incorporating a rate of missing values of 30%, a total of 494 patients (247 per group) needs to be enrolled to achieve a power of 90%.

Estimate based on study MW2010-03-02

Note: In the MW2010-03-02 study, a different formula for the calculation of blood loss, than will be applied to the CIDS study, was used. However, both formulas try to estimate the same quantity and are therefore considered comparable.

Estimated mean blood loss from study MW2010-03-02: NexoBrid 14.17 ml; SoC 814.51 ml

Standard errors: NexoBrid 66.15; SOC 167.74

Difference: -800.34; Standard error of difference: 156.08

Anticipated difference: -800.34 + 156.08 /2 = -722.30

Using a two-sided two-sample t-test with a two-sided significance level of 5%, the total number of evaluable patients needed to achieve a power of 90% is 56 (28 per group). Incorporating a rate of missing values of 30%, a total of 80 patients (40 per group) needs to be enrolled to achieve a power of 90%.

Time to Complete Wound Closure

In the following, a non-inferiority margin of 7 days is already implemented by adding 7 days to the wound closure times of the SoC group.

Estimate based on study MW2004-11-02

Estimated log hazard ratio from study MW2004-11-02 (NexoBrid versus SoC) = 0.49, using a Cox proportional hazards model with shared frailty adjusted for center, age, %TBSA area of the wound and number of TWs (1,2,3-4, ≥5).

Standard error = 0.63

Target log hazard ratio = 0.49 - 0.63/2 = 0.18; target hazard ratio λ = 1.2.

The formula for total number of events D is: $D = \frac{4(z_\alpha + z_\beta)^2}{(\ln \lambda)^2}$

(Piantadosi, Clinical Trials: A Methodologic Perspective, p.169)

z_α = 1.96;

z_β = 1.28 for 90% power.

If the proportion of events in NexoBrid group is denoted by p_1 = 0.64

and the proportion of events in SoC group is denoted by p_2 = 0.56,

which are the observed proportions in the study MW2004-11-02, the number of subjects to be enrolled to achieve a power of 90% for the non-inferiority test is n = $\frac{2D}{p_1+p_2}$= 2164 (1082 in each group).

Estimate based on study MW2010-03-02

Estimated log hazard ratio from study MW2010-03-02 (NexoBrid versus SoC) = 1.56, using a Cox proportional hazards model with shared frailty adjusted for center, age, %TBSA, proportion of the FT area of a patient and number of TWs (1,2,3-4, ≥5).

Standard error = 0.41

Target log hazard ratio = 1.56 - 0.41/2 = 1.36; target hazard ratio λ = 3.90.

The formula for total number of events D is: $D = \frac{4(z_\alpha + z_\beta)^2}{(\ln \lambda)^2}$

(Piantadosi, Clinical Trials: A Methodologic Perspective, p.169)

z_α = 1.96;

z_β = 1.28 for 90% power.

If the proportion of events in NexoBrid group is denoted by p_1 = 0.62

and the proportion of events in SoC group is denoted by p_2 = 0.73,

which are the observed proportions in the study MW2010-03-02, the number of subjects to be enrolled to achieve a power of 90% for the non-inferiority test is n = $\frac{2D}{p_1+p_2}$= 34 (17 in each group).

Cosmesis and Function: MVSS at 12 months

Estimate based on study MW2004-11-02

Estimated difference in MVSS at 24 months from previous study (NexoBrid vs. SoC) = 3.32 – 3.83 = –0.51.

Standard error = 0.50

Target difference Δ = –0.51 + 0.50/2 = –0.26.

The formula for the number of patients per group n is: $n = \frac{2(z_\alpha + z_\beta)^2 \sigma^2}{q(\Delta - d_0)^2}$

where

z_α = 1.96;

z_β = 1.28 for a power of 0.9;

σ = the estimated common standard deviation= 2.33 (taken from previous study);

d_0 = clinical equivalence level = 1.9;

q = proportion who are retained in follow-up to 12 months = 0.7 (assuming 30% loss to follow-up over 1 year)

For 90% overall power, the required sample size per group is 35 (70 overall).

Estimate based on study MW2010-03-02

Estimated mean MVSS at 12 months from study MW2010-03-02: NexoBrid 3.70; SoC 5.08

Standard errors: NexoBrid 0.24; SoC 0.38

Difference: -1.38; Standard error of difference: 0.44

Anticipated difference: -1.38 + 0.44 /2 = -1.16

The non-inferiority margin employed for this analysis is 1.9 score points.

Using a one-sided two-sample t-test with a significance level of 2.5%, 36 evaluable subjects are needed (18 in each group) to achieve a power of 90%. Using a conservative rate of missing values of 30%, the total sample sizes needed to achieve a power of 90% in the analysis is 52 (26 per group).

Note: In the analysis missing data will be imputed, leading to more evaluable subjects, increasing the power.

Cosmesis and Function: MVSS at 24 months

Estimate based on study MW2004-11-02

Estimated difference in MVSS at 24 months from previous study (NexoBrid vs. SoC) = 3.32 – 3.83 = –0.51.

Standard error = 0.50

Target difference Δ = –0.51 + 0.50/2 = –0.26.

The formula for the number of patients per group n is: $n = \frac{2(z_\alpha + z_\beta)^2 \sigma^2}{q(\Delta - d_0)^2}$

where

z_α = 1.96;

z_β = 1.28 for a power of 0.9;

σ = the estimated common standard deviation= 2.33 (taken from previous study);

d_0 = clinical equivalence level = 1.9;

q = proportion who are retained in follow-up to 24 months = 0.5 (assuming 50% loss to follow-up over 2 years)

For 90% overall power, the required sample size per group is 49 (98 overall).

Estimate based on study MW2010-03-02

Since at this time the 24 months MVSS data of the MW2010-03-02 study is not yet available, the 12 months data will be used for the calculation.

Estimated mean MVSS at 12 months from study MW2010-03-02: NexoBrid 3.70; SoC 5.08

Standard errors: NexoBrid 0.24; SoC 0.38

Difference: -1.38; Standard error of difference: 0.44

Anticipated difference: -1.38 + 0.44 /2 = -1.16

The non-inferiority margin employed for this analysis is 1.9 score points.

Using a one-sided two-sample t-test with a significance level of 2.5%, 36 evaluable subjects are needed (18 in each group) to achieve a power of 90%. Using a conservative rate of missing values of 50%, the total sample sizes needed to achieve a power of 90% in the analysis is 72 (36 per group).

13.4. Update Memo to Sites Due to COVID-19 Pandemic

DATE: 09 April 2020

TO: Investigators and Study Coordinators

FROM: CIDS (MW2012-01-01) Study Team

RE: COVID-19 Update

Protocol: A multicenter, multinational, randomized, controlled, open label study, performed in children with thermal burns, to evaluate the efficacy and safety of NexoBrid as compared to standard of care (SOC) treatment

Purpose: Guidance for Protocol Modifications and Documentation During COVID-19 Pandemic

---

Dear CIDS Investigators and Site Study Teams,

The global situation with regard to COVID-19 is rapidly evolving and we appreciate that you are adapting your activities to ensure the safety of your team and study patients. We empathize the challenges faced by you as well as your patients.

The purpose of this document is to provide guidance for you to consider with regards to study conduct knowing the current situation impacts your patients and study staff, also in line with the FDA and EMA special guidance documents on the conduct of clinical trials of medicinal products during COVID-19 pandemic issued in March 2020. We stress that these points are for study guidance only, and that the institutional and local government directives on operational measures that may already be in place should be prioritized.

You must immediately submit a copy of this protocol memo to your respective local Institutional Review Board/Ethics Committee. In parallel, MediWound will ensure regulatory agencies are notified of protocol modifications.

The following guidance and modifications should be taken into consideration:

- **Enrollment of new patients:**
  As of today, the recruitment remains open for patients in the age of 2-4 years old. We encourage sites to only screen patients that can comply with the required study procedures through the end of the study with a few modifications described below.

  Potential patients should be screened per your hospital protocol for potential signs and/or symptoms of COVID-19 to avoid risk for both health care providers and the patient. If a COVID-19 diagnosis is suspected, the patient should be excluded due to exclusion criteria #16; “Any known conditions that would preclude safe participation in the study or add further risk to the basic acute burn trauma”.

- **Follow-up of Enrolled Study Subjects:**
  We will rely heavily on the investigator’s best judgement and their institutional policies to implement and conduct study procedures as conditions vary across countries. We consider completion of the study visits important to the safety of the patients and they should be

encouraged to continue study participation, as long as the patient's safety, welfare and rights are best served by continuing, as determined by the investigator.

The following are some points by study visit for the investigator to consider:

- If an onsite visit is not feasible and cannot be rescheduled, to assure the safety of trial participants and continued investigator oversight, one or more of the following alternative follow-up methods can be implemented to conduct Long-Term Follow-up Visits. These methods include:
    - Patient interview via telephone- at a minimum to collect adverse events and concomitant medication information; however, if possible, this should be combined with a wound assessment via photo/video.

  Additional options:

    - Telemedicine
    - HIPAA compliant Telecommunication (US only)
    - Wound assessment via photograph and/or videos that are taken by the patient and provided to the site. Note: In these cases, a specific photograph label does not need to be utilized. The patient should be reminded not to send any photographs of faces or tattoos to protect patient health information (PHI). If photographs are transmitted by the patient and they contain PHI, they should not be made available to the sponsor.

If a follow-up visit is conducted remotely via one of the methods listed above, the reason should be thoroughly documented in the source documents. Additionally, any assessments that were completed remotely should also be clearly documented within the study source documents.

- For patients that have been discharged from the hospital and have not yet reached wound closure confirmation-

  All wound assessments included in the weekly follow-up visits and the wound closure confirmation can be performed via photograph and/or videos that are taken by the patient/parent and provided to the site, as described above.
- For patients that are in the long term follow-up stage, post wound closure confirmation:
    - If a Long Term Follow-Up visit will be completed remotely, cosmesis assessments such as POSAS and MVSS should be completed to the best of the blinded assessor's ability through patient and parent interview and photo/video assessment. A live video is preferred as the assessor can request the patient or family member to pinch/press the scars to assess pliability and vascularity while moving the camera around the area of the scar. Additionally, through live video, the height of the scars can be measured by a patient or family member with a ruler while moving the camera to a tangential angle in order to be able to see exactly where to measure and the height on the ruler. Note: Only a designated blinded assessor should complete POSAS and MVSS wound assessments.
    - Additionally, if a Long Term Follow-Up visit will be completed remotely, the patient/parent should be mailed or emailed a copy of the applicable self-completed questionnaires (EQ5D-Y, BOQ (for USA sites), POSAS (patient scale), Lower Extremity Function Scale, and QuickDASH) and patient diary. These should be completed by the patient/parent, initialed and dated, and returned back to the

site by mail or email. If it is not possible for the patient/parent to complete the questionnaires and return them, the questionnaires can be read to the patient/parent and responses can be documented by site personnel. If the questionnaires were not self ccompleted, it should be clearly documented in the patient source documents.

- If a patient is unable to return back to the site or complete a visit via telecommunication then the missed visit shall be documented in the source documents and the reason clearly noted to be due to COVID-19.
- Immunogenicity sample collection cannot be performed remotely. Therefore, once the situation normalizes, any patient previously identified for immunogenicity screen should arrive to the site on the earliest possible day for collection of the immunogenicity blood sample.
- A table of the approved alternative methods for remote assessments is attached to this memo (Table 1).

After COVID-19 hospital restrictions have been lifted, sites should continue to complete all remaining study visits and assessments per the approved study protocol.

- **Notification of Study Subjects:**

  Trial participants should be kept informed of changes to the study that could impact them.
  A patient memo clarifying the possible changes during the COVID-19 pandemic is provided to your site together with this memo and should be submitted to your IRB/EC and provided to all patients' parents. This memo will inform the parents that they may be asked to complete follow-up visits remotely via telemedicine/telecommunicatio or phone interview and they may be asked to provide photos of their child's wounds and complete questionnaires at home.

  This memo should also be provided to parents of new patients that are enrolled during the COVID-19 pandemic, as an addendum to the approved Informed Consent Form, so that they are aware of the changes that may occur in follow up procedures until the situation normalizes and on-site follow up visits are possible.

- **Study Monitoring Procedures:**

  Many sites/institutions are no longer allowing onsite routine monitoring visits. If this is the case at your site, or there is a change to your current situation, please inform your CRA. The onsite monitoring visit can be re-scheduled and replaced by an offsite monitoring visit as needed.

Thank you for your continuous support with the study. We understand the inconvenience associated with this situation and realize that we cannot anticipate all site-specific issues. As always, please do not hesitate to reach out to any one of us if you have any questions. Please contact your CRA, or MediWound PM, Aya Ben Yaakov (email: ayaby@mediwound.com), with any questions you may have. We will respond promptly by email or telephone as you would prefer.

We will alao provide further communications as new information arises.

Please file this communication in your Investigator Site File or appropriate site regulatory binders.

Sincerely,

The CIDS MW2012-01-01 Study Team

Table 1 - Optional Methods for Completion of Remote Long Term Follow-up Visits, During COVID-19 Pandemic

Assessment		Method of Contact			Comments
		Photo sent by Patient	Phone conversation	Video call/Telemedcine	
Wound photographs		X	N/A	N/A	Photos/Videos that will be sent by the patients will not include photo label.
Documentation	**Adverse events**	X Photo Assessment- Local TW related only	X	X	Any SAEs should be reported within 24 hours of knowledge of the event.
	Concomitant Medication review	N/A	X	X	
	Scar modulation procedures	N/A	X	X	
POSAS Observer Scale		X		X	Video method preferred, if available
MVSS		X		X	Video method, if available
ROM		N/A	N/A	N/A	Can be completed during an on-site visit only.
Patient questionnaires	**POSAS Patient Scale**	X Photocopy sent back to site	X Subject interview	X Subject interview	It is preferred to have questionnaires sent to subject via mail/email, completed, and returned to site, if possible.
	QOL (EQ5D,BOQ)	X Photocopy sent back to site	X Subject interview	X Subject interview	It is preferred to have questionnaires sent to subject via mail/email, completed, and returned to site, if possible.
	LEFS/ QuickDASH	X Photocopy sent back to site	X Subject interview	X Subject interview	It is preferred to have questionnaires sent to subject via mail/email, completed, and returned to site, if possible.
	Subject Diary	X Photocopy sent back to site	X Subject interview	X Subject interview	It is preferred to have questionnaires sent to subject via mail/email, completed, and returned to site, if possible.
Immunogenicity		N/A	N/A	N/A	Can be taken during an on-site visit only. May be performed in delay