Statistical Analysis Plan

Clinical Investigation Plan Title	A Prospective, 5-Year Global Study on MAST™ Minimally Invasive Fusion Procedures for the Treatment of the Degenerative Lumbar Spine
Clinical Investigation Plan Identifier	P15-01
Clinical Investigation Plan Version	Version 4.0
Sponsor/Local Sponsor	Medtronic Bakken Research Center BV Spinal and Biologics Clinical Endepolsdomein 5 6229 GW Maastricht, The Netherlands
Document Version	Version 1.0

Table of Contents

1. Version History

Version	Summary of Changes	Author(s)/Title
1.0	Not Applicable, New Document	[redacted] Senior Statistician

2. List of Abbreviations and Definitions of Terms

Abbreviation	Definition
ADE	Adverse Device Effect
AE	Adverse Event
ALIF	Anterior Lumbar Interbody Fusion
ANCOVA	Analysis of Covariance
CI	Confidence Interval
CIP	Clinical Investigation Plan
DDD	Degenerative Lumbar Disc Disease
DLIF	Direct Lateral Lumbar Interbody Fusion
DN4	Douleur Neuropathique 4
EC	Ethical Committee
FU	Follow-up
IRB	Institutional Review Board
MAST™	Minimal Access Surgical Technologies
MIDLF	Midline Lumbar Fusion
ODI	Oswestry Disability Index
OLIF	Oblique Lumbar Interbody Fusion
PLIF	Posterior Lumbar Interbody Fusion
PRO	Patient Reported Outcomes
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SADE	Serious Adverse Device Effect
TLIF	Transforaminal Lumbar Interbody Fusion

3. Introduction

This prospective 5-year global study on MAST™ (Minimal Access Surgical Technologies) minimally invasive fusion procedures for the treatment of the degenerative lumbar spine is expected to better understand the different minimally invasive MAST™ procedures regarding patient pathology, surgical exposure and expected safety and clinical outcomes, from a short (intra- and peri-operative) to a long-term follow-up (FU) post-operatively, from the standard of care perspective. A single or double level fusion procedure is performed using minimally invasive posterior lumbar interbody fusion (PLIF), transforaminal lumbar interbody fusion (TLIF), direct lateral lumbar interbody fusion (DLIF), oblique lumbar interbody fusion (OLIF), anterior lumbar interbody fusion (ALIF), or midline lumbar fusion (MIDLF) procedure in subjects with degenerative lumbar disc disease (DDD) with or without spondylolisthesis and/or stenosis. Anterolateral approach includes ALIF, DLIF or OLIF, and posterior approach includes PLIF, TLIF or MIDLF.

The study is anticipated to enroll 350 subjects. Assessments will be done pre-operative, during surgical procedure and hospital stay and post-operative at 1, 3 months, 1, 2, 3, 4 and 5 years FU.

In this study, the two-sided 95% confidence interval (CI) will be conducted to test hypothesis for primary and the first secondary objectives. For other analyses without hypothesis testing, the analyses are descriptive in nature. In general, the categorical variables will be summarized using frequency tables and continuous variables will be summarized using mean, standard deviation, median, minimum and maximum.

The approved clinical investigation plan (CIP) version 4.0 provides the guidelines to this statistical analysis plan (SAP). The purpose of this SAP is to define the plan for data analysis to support an interim report when all subjects reach 3-month visit, final report and publication defined with publication plans.

4. Study Objectives and Endpoints

The MASTERS-D 2 clinical study aims:

- To evaluate the effectiveness of MAST techniques for anterior/lateral and posterior approaches in patients with spondylolisthesis ((≥grade I).
- To assess how single or double level MAST fusion procedures PLIF, TLIF, DLIF, OLIF, ALIF, or MIDLF are used in surgical practice and to describe long-term safety and effectiveness in a broad patient population of patients with degenerative lumbar disc disease.

4.1. Primary Objective and Endpoint

The primary objective is to demonstrate that DDD patients operated for spondylolisthesis fare equally well regardless of the surgical procedure (anterolateral or posterior) performed as measured by the primary endpoint which is defined as improvement of Oswestry Disability Index (ODI) at 3 months post-operatively as compared to baseline.

4.2. Secondary Objectives and Endpoints

The secondary objectives include:

1. To evaluate whether DDD patients without spondylolisthesis fare equally well regardless of the surgical procedure (anterolateral or posterior) performed, as measured by endpoint which is defined as improvement of ODI at 3 months as compared to baseline.
2. To observe and document clinical and radiological outcomes through 5-year follow-up (except for ODI at 3 months). Endpoints include improvement of ODI as compared to baseline, improvement of VAS back- and leg pain intensity as compared to baseline, improvement of EQ-5D as compared to baseline, neurological success, fusion success, secondary surgeries at index and /or adjacent level(s) and adverse events (AE).
3. To observe and document the patients' short term recovery after surgery. Endpoints include time needed for first ambulation and surgery recovery day.
4. To document health resources consumption and perform economic analyses.
5. To observe and document the patient profile when choosing a particular minimally invasive fusion procedure.

5. Determination of Sample Size

Taking into account the literature information and the opinion of the Study Advisory Board, the "equally well" is defined as the mean difference in ODI improvement at 3 months between the anterolateral and posterior group is within the equivalence range of (-10, 10) points, which indicates if the mean difference between the anterolateral procedures group and the posterior procedures group is within the range of (-10, 10) points for the ODI improvement at 3 months, the anterolateral and posterior treatment options will be considered equivalent in terms of effectiveness.

Assuming that the standard deviation for ODI improvement at 3 months is 20 for both groups (anterolateral and posterior) and there is no difference in the mean ODI improvement at 3 months, with 80% power and 5% alpha level, the sample size per group is 70 and 140 in total in order to claim equivalence of the primary endpoint between two groups with equivalent margin (-10, 10).

From the MASTERS-D study it is known that approximately 50% of the patients with degenerative disc disease have spondylolisthesis ≥ grade I. Hence the required total sample size is 140/0.5 = 280.

The study is anticipated to enroll 350 patients. The amount of patients with spondylolisthesis versus nonspondylolisthesis will be monitored during the study. In case the study team identifies that the target amount of 70 spondylolisthesis patients per group for the anterolateral procedures and posterior procedures might not be reached, the sites might receive instruction to limit the enrollment of non-spondylolisthesis patients. This action will secure the enrollment of target sample size as calculated for the primary endpoint.

6. Statistical Methods

6.1. Study Subjects

6.1.1. Disposition of Subjects

Patients may exit from the study at any time and for any reason. Stop of data collection for patients in the study may occur due to the following reasons:

- Withdrawal of consent
- Patient lost-to-follow-up
- Patient death
- Patient did not undergo the instrumented lumbar fusion as defined in CIP

Patients who terminate from the study will be summarized by the reasons for termination according to the categories in the CRF at each visit. An accountability table will also be provided to summarize the disposition of patients, including number of patients discontinued from the study, number of patients having follow-up, and follow-up.

6.1.2. Clinical Investigation Plan (CIP) Deviations

In this study, potential CIP deviations are defined as deviations related to:

- Data release or patient informed consent procedure
- Enrolled patient does not meet inclusion/exclusion criteria
- Visit not done
- Assessment(s) at the visit not performed
- Visit outside the CIP defined visit window
- Patient did not undergo the instrumented lumbar fusion as defined in CIP
- Improper Serious Adverse Event (SAE)/Adverse Device Effect (ADE)/Serious Adverse Device Effect (SADE) reporting (please see section Table 5 AEs of this CIP)
    - Any (U)(S) A(D)Es, device deficiencies with SADE potential and serious procedure related AEs, that are not reported timely
    - Any AE or device deficiency without (S)A(D)E potential that is not reported within 20 calendar days
- Ethical Committee (EC)/Institutional Review Board (IRB) approval not obtained, if required

Details of CIP deviations will be summarized descriptively. Subjects with deviations will be summarized by type of deviations listed above and by follow-up visit.

6.1.3. Analysis Datasets

The primary analysis dataset defined in this study comprises of all subjects who are enrolled and receive a single or double level instrumented fusion using MAST™ minimally invasive PLIF, TLIF, DLIF, OLIF, ALIF and MIDLIF surgical procedures with a mandatory posterior fixation. For double levels instrumented fusion, the same procedure must be used for both levels. All the analyses will be carried out on the

primary analysis dataset. For subjects who are enrolled but not included in the primary analysis dataset, their demographic and other baseline characteristics will be summarized separately. The reasons for not using minimally invasive procedure or surgery not done will be listed as well.

In addition to the primary analysis dataset, four subgroups are defined in the following:

1) DDD patients with spondylolisthesis and stenosis
2) DDD patients with spondylolisthesis but without stenosis
3) DDD patients with stenosis but without spondylolisthesis
4) All patients not included in subgroups 1), 2), 3).

The analysis for the primary endpoint and secondary endpoints outlined in secondary objectives 2 and 3 (section 4.2) will be performed on these four subgroups as well. For the primary analysis dataset and each of these four subgroups, the analysis will be carried out on further subgrouped by anterolateral and posterior groups.

Other subgroup analysis such as by number of treated levels may also be performed.

6.2. General Methodology

In this study, the two-sided 95% CI will be conducted to test hypothesis for primary and the first secondary objectives. For other analyses without hypothesis testing, the analyses are descriptive in nature. In general, the categorical variables will be summarized using frequency tables and continuous variables will be summarized using mean, standard deviation, median, minimum and maximum.

Time-to-event (life table method) will be performed to estimate the cumulative event rate up to 5 years for secondary surgery at index levels or adjacent levels and AEs.

Time-to-event analysis will be conducted for fusion success.

6.3. Center Pooling

This study will be conducted in approximately 30 sites located in Europe, Latin America and Asia Pacific. A test of analysis of covariance (ANCOVA) will be performed to check whether data are poolabe across different regions for the primary endpoint. The ANCOVA model includes the change of ODI at 3 months as dependent variable, and region, procedure group (anterolateral versus posterior) and interaction between region and procedure group as factors, and baseline ODI as covariate. We anticipate Latin America has less subjects treated than other regions. If Latin America has less than 10 subjects in each procedure group, subjects in this region will be combined with subjects in Europe. If the region factor is not statistically significant, the pooling data across regions is justified and pooling data from all regions is planned for final statistical analyses. Otherwise, the results from each region will be presented separately.

6.4. Handling of Missing Data and Dropouts

All data analyses will be based on the observed data. Missing data due to lost to follow-ups will not be imputed. For study subjects included in the primary analysis dataset who drop out or withdraw from the study, data collected before the occurrence of dropout or withdrawal will be included in analysis.

6.5. Adjustments for Multiple Comparisons

The two-sided 95% CI will be conducted to test hypothesis for primary and the first secondary objective. For changes from the baseline, paired t-test will be performed to test whether the changes are statistically significant. No adjustments for multiple comparisons will be carried out.

6.6. Demographic and Other Baseline Characteristics

Demographic and other baseline information such as general history, comorbidities, musculoskeletal history, spine surgical history, medication, work status and indication for surgery will be summarized descriptively. In the indication for surgery, instead of summarizing measurements of length of body and slip in millimeter, the percentage of slip or grade/type of spondylolisthesis, and disc relative height for adjacent levels will be summarized due to zooming effect from x-ray and MRI. Continuous variables will be summarized using mean, standard deviation, median, minimum and maximum and categorical variables will be summarized using frequency and percentage tables.

Baseline information will be summarized for both all DDD patients and subgroup datasets defined in section 6.1.3 and/or per surgical procedure (ALIF, OLIF, DLIF, PLIF, TLIF, MIDLF) and per anterolateral and posterior groups.

6.7. Treatment Characteristics

Surgical procedure data include:

1. Duration of the surgery (minutes)
2. Estimated blood loss (ml) and transfusion
3. Total fluoroscopy time (seconds)
4. Levels treated
5. Procedure information (type of approach, access system, decompression, cage/interbody, anterior plate(s), posterior fixation technique, bonegraft and substitutes)

Continuous variables will be summarized using mean, standard deviation, median, minimum and maximum while categorical variables will be summarized using frequency tables.

Concomitant medications including nonopioid analgesics, opioid analgesics and other medications will also summarized at each visit descriptively.

6.8. Interim Analyses

In this study, in addition to final report, an interim report when all subjects reach 3-month visit is anticipated. The methodology described in section 6.2 will be used for the interim analyses. The cumulative event rate for secondary surgery at index levels or adjacent levels and AEs up to cut-off date of interim analyses will be conducted by time-to-event analysis (life table method).

6.9. Evaluation of Objectives

All analyses will be performed based on the scheduled visits. A sensitivity analysis on primary endpoint will be performed as well. In the sensitivity analysis, the visits are defined based on the analysis visit window. In the case that two or more observations fall inside the same visit window, the observation which is closer to the nominal visit date will be analyzed.

The analysis visit windows are defined in Table 1, which is deviated from the scheduled visits defined in the CIP.

For subjects who have additional surgical procedures at the index levels or have implanted device explanted, since these additional surgical procedures are considered to have interfered with clinical outcomes of intended treatments, the last observation taken before the first additional surgical procedures will be carried forward for all future evaluation periods for clinical endpoints and radiographic endpoints.

Table 1: Analysis Visit Window

Study Visit	Analysis Visit Window
Surgery	Day 0
Postop (Day 1 to 2-Week)	Day 1 to 14
4-Week	(>2 to 6 weeks): Day 15 to 42
3-Month	(>6 weeks to 6 months): Day 43 to 182
1-Year	(>6 to 18 months): Day 183 to 547
2-Year	(>18 to 30 months): Day 548 to 912
3-Year	(>30 to 42 months): Day 913 to 1277
4-Year	(>42 to 54 months): Day 1278 to 1642
5-Year	(>54 months): Day 1643 and beyond

6.9.1. Primary Objective

The study hypothesis for the primary objective is to verify whether DDD patients operated for spondylolisthesis (≥ grade I) have equivalent mean improvement of ODI at 3 months regardless of anterolateral or posterior procedure used.

The hypothesis includes:

$$H_0: \triangle_{ODI\_anterolateral} \neq \triangle_{ODI\_posterior}$$
$$H_a: \triangle_{ODI\_anterolateral} = \triangle_{ODI\_posterior}$$

Where $\triangle_{ODI\_anterolateral}$ is the average improvement in ODI score (baseline – 3 months) in spondylolisthesis subjects treated with the anterolateral approach and $\triangle_{ODI\_posterior}$ is the average improvement in ODI score (baseline – 3 months) in spondylolisthesis subjects treated with the posterior approach.

ANCOVA will be performed with baseline ODI as covariate and the two-sided 95% CI for the least square mean difference of ODI improvement between two groups at 3 months will be calculated. If the 95% CI is within the equivalent range of (-10, 10), then the primary objective is met. In addition, the paired t-test will be carried out for each group to see whether the ODI improvement at 3 months from baseline is significant.

6.9.2. Secondary Objectives

6.9.2.1. To evaluate whether DDD patients without spondylolisthesis fare equally well regardless of the surgical procedure (anterolateral or posterior) performed, as measured by ODI at 3 months

The similar method as the one outlined above for the primary endpoint will be used to assess whether the ODI improvement at 3 months is equivalent for the non-spondylolisthesis subjects treated with anterolateral approaches and the ones treated with posterior approaches.

6.9.2.2. To observe and document clinical and radiological outcomes through 5-year follow-up (except for ODI at 3 months)

The following analyses will be carried out for both the primary dataset and subgroup datasets defined in section 6.1.3.

6.9.2.2.1. Improvement of ODI / VAS back and leg pain intensity / EQ -5D as compared to baseline

Summary statistics (mean, standard deviation, median, minimum and maximum) will be presented for the ODI / VAS back and leg pain intensity / EQ-5D at baseline and each follow-up (except for ODI at 3 months), and the improvement of the ODI / VAS back and leg pain intensity / EQ-5D at each follow-up through 5-year visit (except for ODI at 3 months) from baseline. Paired t-test will be performed to see whether the improvement is significant.

6.9.2.2.2. Neurological success

Neurological status is based on four sections: motor, sensory, reflexes, and straight leg raising. Each of the sections is comprised of a number of elements.

6.9.2.2.2.1 Motor Function

For the motor function, the assessments are listed in ascending order of improving as indicated below:

0 = Total Paralysis

1 = Palpable or Visible Contraction

2 = Active Movement, Gravity Eliminated

3 = Active Movement, Against Gravity

4 = Active Movement, Against Some Resistance

5 = Active Movement, Against Full Resistance (full strength) (Normal)

For the time period evaluated, for each of elements, if the score increases from baseline, the subject is considered "improved'; if the score remains the same as baseline, the subject is considered "maintained"; if the score decreases from the baseline, the subject is considered "deteriorated". In the situation where the score at baseline is 0 (Total Paralysis) and a postoperative visit is missing, or the score at baseline is missing and at a postoperative visit is 5 (Normal), the subject is conservatively considered as "maintained". For overall motor status, if at least one element has the status of "improved" and the rest are "maintained", then the overall motor status is considered "improved"; if all elements have the status of "maintained", then the overall motor status is "maintained"; if at least one element has the status of "deteriorated", then the overall motor status is "deteriorated."

6.9.2.2.2.2 Sensory function (Light Touch or Pin Prick L1 to S1)

For the sensory function, the assessments include absent, impaired and present (normal) which are listed in ascending order of improving. Let absent be 0, impaired be 1 and normal be 2. The same logic outlined in the motor function section will be used to first determine the change status of each element from the baseline and then the overall sensory status from the baseline. For each element, in the situation where the score at baseline is 0 (Absent) and a postoperative visit is missing, or the score at baseline is missing and at a postoperative visit is 2 (Normal), the subject is conservatively considered as "maintained".

6.9.2.2.2.3 Reflexes

For the reflexes, the assessments are listed as indicted below:

0 = Absent or Trace

1 = Hyper-reflexic

2 = Normal

For each element, a change from 0 (Absent or Trace) to 2 (Normal) or 1 (Hyper-reflexic) to 2 (Normal) is considered improved while the other way around is considered deteriorated. No change is considered maintained. Any change from 0 (Absent or Trace) to 1 (Hyper-reflexic) is considered "diminished to hyperactive" while the other way around is considered "hyperactive to diminished". In the situation where the score at baseline is missing and at a postoperative visit is 2 (Normal), the subject is conservatively considered as "maintained". If none of element has the situation of "diminished to hyperactive" or "hyperactive to diminished", an overall evaluation of reflex will be summarized using the same logic outlined in the motor function section. Otherwise, there will be no overall evaluation.

6.9.2.2.2.4 Straight leg raise

For the straight leg raise, the assessments include positive and negative (normal). Let positive be 0 and negative be 1. The same logic outlined in the motor function section will be used to first determine the

change status of each element from the baseline and then the overall straight leg raise status from the baseline. For each element, in the situation where the score at baseline is 0 (Positive) and a postoperative visit is missing, or the score at baseline is missing and at a postoperative visit is 1 (Negative), the subject is conservatively considered as "maintained".

Overall neurological success will be defined as maintenance or improvement in all sections (motor, sensory, reflex, and straight leg raising) for the time period evaluated. In order for a section to be considered a success, each element in the section must remain the same or improved from the time of the baseline evaluation to the time period evaluated. If at least one section is deteriorated, the overall neurological success will be defined as failure.

Summary statistics (frequencies) will be presented for status of each section and overall neurological success at each follow-up.

6.9.2.2.3. Fusion success

Fusion will be assessed at 1 year, preferably by collecting a CT-scan, alternatively fusion may also be collected through X-Rays. If no fusion observed during the first fusion assessment, further fusion assessment(s) will be at the discretion of the investigator. The surgeon or hospital radiologist will determine fusion. Time-to-event analysis will be conducted for fusion success. The fusion success rate will be summarized at 1, 2, 3, 4 and 5-year follow-up, if applicable. If a subject shows fusion success at early timepoints and non-fusion at a later timepoint, the fusion status at all earlier points should be considered non-fusion. For 2-level subjects, fusion success will be defined as achieving fusion at both treated levels. In addition to the time-to-event analysis, the last evaluation for fusion will be summarized.

6.9.2.2.4. Secondary surgeries at index and/or adjacent level(s)

For secondary surgery at index levels and/or at adjacent levels, time-to-event (life table method) will be performed to estimate the cumulative event rate up to 5 years.

6.9.2.2.5. Adverse events

Detailed information about how AEs will be summarized can be found in section 6.10.

6.9.2.3. To observe and document the patients' short term recovery after surgery

The following analyses will be carried out for both primary and subgroup datasets defined in section 6.1.3.

6.9.2.3.1. Time needed for first ambulation

Time needed for first ambulation is defined as the days after the initial surgery for patients to get out of bed and ambulate with or without assistance. Summary statistic including mean, standard deviation, median, minimum and maximum for continuous variable will be provided.

6.9.2.3.2. Surgery recovery day

Surgery recovery day is defined as days after the initial surgery, when patients have no need for intravenous infusion of analgesic drugs, no ongoing surgery related AEs impending discharge, and no need for nursing care. Summary statistic including mean, standard deviation, median, minimum and maximum for continuous variable will be provided.

6.9.2.4. To document health resources consumption and perform economic analyses

Healthcare utilization will be measured with the following three cost components and summarized descriptively:

1. Resource: procedure, length of hospital stay, additional medical visit (outpatient consultation, visits to primary care services, etc.)
2. Medication: the number of patients who take pain medications at baseline and the past week before each followup visit will be summarized.
3. Non-pharmacologic therapies: the number of patients who utilize rehabilitation program between the scheduled followup visit and the previous scheduled visit will be summarized.

Economic analyses are not planned in this SAP.

6.9.2.5. To observe and document the patient profile when choosing a particular minimally invasive fusion procedure

The baseline information will be summarized by individual procedures.

6.9.3. Additional Analyses

1. Analyses for baseline information can be found in section 6.6.

2. Determine if there is a correlation between sagittal balance at baseline and the radiological findings (fusion) and the clinical outcomes at the follow-up visit, in those sites where this assessment is standard of care. Pearson correlation coefficient and 95% CI will be provided.

3. Hospital stay in days after the surgery when patients has been discharged from hospital will be summarized.

4. Days to first return to work will be summarized.

5. Sagittal balance parameters and change from the baseline if applicable will be summarized by visits in those sites where this assessment is part of standard of care. The parameters include pelvic incidence, sacral slope, pelvic tilt, and global lordosis angle.

6. Adjacent levels radiographic review will be summarized by visits if it is available. Instead of summarizing measurements of length of body and slip in millimeter, the percentage of slip or grade/type of spondylolisthesis, and disc relative height will be summarized due to zooming effect from x-ray and MRI.

7. Several subgroup analyses will be conducted based on Douleur Neuropathique 4 (DN4) scores (respectively for leg and back DN4 score) at baseline, 3-month and 12-month visit. Subgroups include:

- Subjects with baseline DN4≥4 and DN4 ≥4 at 3-month follow-up
- Subjects with baseline DN4≥4 and DN4 <4 at 3-month follow-up
- Subjects with baseline DN4<4 and DN4 ≥4 at 3-month follow-up
- Subjects with baseline DN4<4 and DN4 <4 at 3-month follow-up
- Subjects with baseline DN4≥4 and DN4 ≥4 at 12-month follow-up
- Subjects with baseline DN4≥4 and DN4 <4 at 12-month follow-up
- Subjects with baseline DN4<4 and DN4 ≥4 at 12-month follow-up
- Subjects with baseline DN4<4 and DN4 <4 at 12-months follow-up

For each subgroup, the number and percentage of subjects will be provided. In addition, ODI, EQ-5D, VAS back and leg pain scores at baseline, at each follow-up and the improvement from baseline will be presented by the subgroups. Statistical comparison may be carried out for these subgroups analyses.

8. Regression analyses may be conducted by taking into consideration of pathology (spondylolisthesis versus non-spondylolisthesis) and surgical approach perspective (anterolateral versus posterior) to examine the influence of pathology or surgical approach on clinical outcomes including neurological status, DN4, EQ-5D and work status. Also analyses my be conducted to examine how the demographic information such as age predicts the clinical outcomes.

6.10. Safety Evaluation

The analyses include:

- Time-to-event (life table method) to assess the cumulative AE rate up to 4-week, 1-year, 2-year, 3-year, 4-year and 5-year visit intervals.
- Summary of number of AEs and number of subjects having AEs up to the 5-year visit.
  - All AEs
  - Serious Adverse Events (SAE)
  - Device related AEs
  - MAST™ Approach/Procedure related AEs
  - General Surgery related or Lumbar Fusion procedure related AEs
  - Disease related AEs.
  - Serious device related AEs
  - Serious MAST™ Approach/Procedure related AEs

A relatedness determination of "possible", "probable" or "causal relationship" is considered "related".

6.11. Health Outcomes Analyses

Please refer to session 6.9.2.4..

6.12. Changes to Planned Analysis

This section is not applicable.

7. Validation Requirements

For interim and final report, level I validation will be applied for all statistical analyses in this study, which means that the peer reviewer will independently write statistical programs, generate results and then compare the output with the one generated by the primary statistician.

Progress reports, such as annual report, are intended to include as much and the latest data as possible, including data that have not been monitored, cleaned or finalized. Thus, these reports will not be validated by using independent programming. Instead level III validation will be used, which means that the primary statistician will perform a visual inspection of the code and output to confirm functionality.

For any manuscript publication, level I validation will be applied with independent programming from peer reviewer.

Please refer to work instruction 056-WI181, "Statistical Programming Code Quality Control and Validation" for details.