C-REGS 2
Cerebrolysin **REG**istry **S**tudy in **S**troke

A registry study to observe clinical practices, safety and effectiveness of routine use of Cerebrolysin in the treatment of patients with moderate to severe neurological deficits after acute ischaemic stroke

STATISTICAL ANALYSIS PLAN (SAP)

Study Identification №: **EVER-AT-0717**

Observational Plan: **Version 3.1 (October 17, 2017)**
Version 3.2 (July 10, 2018, Inclusion of Pharmaeconomics Substudy)

Development Phase: **Phase IV**

Study Design: **Non-interventional, controlled, open-label, prospective, multicentre, restricted cohort, observational registry study**

Sponsor: **EVER Neuro Pharma GmbH**
4866 Unterach, Oberburgau 3, Austria

SAP Version: **Final 1.1, July 22, 2024**

SAP Revision History

Version	Date	Author	Review	Stage I FPFV	Stage II LPLV	Comments
Draft 1.0	17 Oct 2017	J.C. Vester	V.W. Rahlfs	08 May 2018		Initial Version
Final 1.0	24 Oct 2017	J.C. Vester	V.W.Rahlfs			No changes to Draft 1.0
Final 1.1	22 Jul 2024	J.C. Vester	V.W.Rahlfs		26 April 2024	Additional specifications in Section 6.

1. TABLE OF CONTENTS

PRELIMINARY REMARK

Sections 2 to 5 of this Final Statistical Analysis Plan (SAP) display the original sections of the Final SAP Version 1.0 from 24.10.2017.

Section 6 documents additional specifications and further operationalizations to the sections 2 to 5 of the Final Version 1.0 from 24.10.2017.

There are no major changes to the Final SAP Version 1.0 from 24.10.2017.

2. INTRODUCTION

2.1 Background information

Cerebrolysin is a neuropeptide preparation with marketing authorization for the treatment of cerebrovascular disorders and neurodegenerative disease for many years worldwide. Since its first approval stroke therapy has evolved and new treatment concepts have been implemented. In addition, Cerebrolysin treatment in stroke has evolved with different time windows, dosages and lengths of therapy being given in a pragmatic way by physicians. The main aim of this study is to capture these variables of the Cerebrolysin treatment and its comedication in the group of patients with moderate to severe neurological deficits after acute ischemic stroke in order to give guidance to further research.
C-REGS 2 is an international, non-interventional, prospective registry study to observe clinical practices of routine use of Cerebrolysin in patients with moderate to severe neurological deficits after acute ischemic stroke in a controlled and open-labelled manner. All patients receive acute stroke care according to local treatment standards, which will not be amended or influenced by the study in any way. To evaluate the safety and effectiveness of Cerebrolysin in routine practice the outcome of Cerebrolysin-treated patients are compared with control group patients, who do not receive Cerebrolysin

2.2 SAP Related Documents

The statistical analysis plan (SAP) is a detailed technical extension to the Observational Plan Version Revised Final 3.0, following principles of the Guidelines ICH E3/E6R2/E9 and GRACE, as well as relevant idv SOPs and/or guidelines as far as applicable to the C-REGS 2 trial. This plan describes the final statistical analysis planned to be performed after all enrolled patients data are available for analysis of EVER Neuro Pharma Protocol EVER-AT-0717. The analysis plan will be finalized and signed after raw database hard lock, but before generation of the analysis database.

All planned analyses identified in this SAP will be included in the Clinical Study Report (CSR). Exploratory analyses not necessarily identified in this SAP may be performed to support planned analyses. Any post-hoc or unplanned analyses not specified in this SAP will be clearly identified as such in the CSR.

The following documents were reviewed when preparing this SAP:

- Observational Plan, Version Final 3.1, dated October 15, 2017
- C-REGS 2 Annotated eCRF, Version Draft 3, dated July 31, 2017

Readers of this SAP are encouraged to read the Observational Plan for details on the conduct of this study and the operational aspects of clinical assessments and timing for completing a patient in this study.

3. OVERALL STUDY DESIGN AND OBSERVATIONAL PLAN

Title	**C-REGS 2** **C**erebrolysin **REG**istry **S**tudy in **S**troke A registry study to observe clinical practices, safety and efficiency of routine use of Cerebrolysin in the treatment of patients with moderate to severe neurological deficits after acute ischaemic stroke
Name of finished product	Cerebrolysin
Name of active substance	Cerebrolysin Concentrate
Coordinating investigator	Prof. Michael Brainin
Number of sites & countries	Approx. 70 sites in Europe and Asia
Indication	Acute Ischemic Stroke with moderate to severe neurological deficits
Study Design	Non-interventional, controlled, open-label, prospective, multicentre, restricted cohort observational registry study
Study groups	All patients receive acute stroke care according to local treatment standards, not amended or influenced by the study: **Cerebrolysin Group:** Patients who are treated with Cerebrolysin; dosage, frequency and duration follows local clinical practice in accordance with terms of the local marketing authorization (see Appendix 2) **Control group:** Patients who are not treated with Cerebrolysin
Study timelines	Patient recruitment Q1/2018 – Q1/2020 Definition “End of study”: Database closure
Study Duration	Patients are followed over a maximum of 100 days
Sample Size	Approx. 2000 patients within the framework of a two-stage procedure according to Bauer-Köhne. Stage I will be completed after enrollment of approx. 670 patients. Sample size is statistically justified.
Study objectives	Investigation of clinical practices, safety and effectiveness of Cerebrolysin in routine treatment of patients with moderate to severe neurological deficits after acute ischemic stroke

Documented items	**Baseline:** • Patient data (age, gender, ethnicity) • Inclusion/exclusion criteria • Patient logistics • Risk factors • NIHSS • Evidence of dementia (IQCODE) **Treatment:** • Acute interventions • Neurorehabilitation • Cerebrolysin treatment • Other treatments (during hospital stay) **Discharge:** • Patient logistics • Stroke diagnosis (confirmation) • NIHSS • mRS • Neurorehabilitation **Day 21±4:** • NIHSS • mRS **Day 90±10:** • Patient logistics • NIHSS • mRS • Cognitive status (MoCA) • New event (within three months) • Neurorehabilitation **Death:** • Date/time • Cause **Adverse events:** • Date/time • Relationship/Seriousness/Outcome
Eligibility criteria	**Observation criteria:** • Signed Informed Consent • Clinical diagnosis of acute ischemic stroke confirmed by imaging • Moderate to severe neurological deficits with NIH Stroke Scale (NIHSS) 8 to 15, both inclusive • No prior stroke • No prior disability • Patient's independence prior to stroke onset (pre-morbid mRS of 0 or 1) • Reasonable expectation of successful follow-up (max. 100 days)

Statistical Methods	**Sample Size Calculation:** Nonparametric sample size calculation was performed to allow detection of "small" group differences in the ordinal comparative effectiveness evaluation with 90% power. The study will use a two-stage adaptive design according to Bauer-Köhne. The total sample size including compensation for 'usual ambiguities' (dropouts, etc.) results in approx. 2000 patients to be enrolled (stage I and stage II). The first stage will enroll approximately 670 patients ($r_{subsample\ I}$ = 0.3). If there is no rejection after stage I analysis due to success or futility, the trial may continue to stage II. **Bias minimizing measures:** • Enrollment Bias: In order to minimize enrollment bias, the patient groups will be standardized using multilevel stratification procedures in combination with a 'restricted cohort' design. The respective risk factors have been identified from published research results on predictors of stroke outcome, allowing appropriate control for confounders. The pre-specified strategy follows the recommendations of the *Principles for Good Research on Comparative Effectiveness (GRACE)*. • Quality assurance The study shall be conducted in a manner fully consistent with good clinical practice. Data will be captured using an eCRF-system with quality assurance performed by edit checks and frontline risk-based control. In addition, and in order to comply with recent calls for high-quality non-interventional comparative effectiveness research, a risk-based centralized statistical approach to monitoring is introduced in combination with targeted on-site monitoring for ongoing surveillance of study conduct, thus ensuring highest standards of data quality and integrity according to the most recent requirements of the ICH E6 *Guideline for Good Clinical Practice* (GCP, Amendment R2, July 2015), the FDA Guidance for Industry on a *Risk-based Approach to Monitoring*, and the EMA reflection-paper on *risk-based quality management in clinical trials*. • Other sources of bias: Other aspects of care than administration of study drug may vary between the study groups. Analyses will consider these potential sources of variation by appropriate sensitivity analyses.

	Effectiveness analysis: • Primary effectiveness analysis: 1) Ordinal modified Rankin Scale (mRS) at 3 months after stroke onset • Secondary effectiveness analyses include: 1. Ordinal NIH Stroke Scale (NIHSS) at 21 days and 3 months after stroke onset 2. Ordinal modified Rankin Scale (mRS) at 21 days after stroke onset 3. Proportion of patients with excellent recovery (mRS score 0-1) at 3 months after stroke onset 4. Proportion of patients with functional independence (mRS score 0-2) at 3 months after stroke onset 5. Ordinal MoCA at 3 months after stroke **Safety analysis:** • Mortality • AEs, ADRs, SAEs, SADRs • SUSARs to Cerebrolysin
Selection of Patients	**Criteria for observed patient population:** • Signed Informed Consent • Clinical diagnosis of acute ischemic stroke, confirmed by imaging • Moderate to severe neurological deficits (observation window: NIHSS 8-15) • No prior stroke • No prior disability • Patient's independence prior to stroke onset (pre-morbid mRS of 0 or 1) • Reasonable expectation of successful follow-up (max. 100 days)

4. STATISTICAL METHODS

4.1 General Principles

This is a registry-based, non-interventional, controlled, restricted-cohort, observational study to evaluate safety and effectiveness of treatment after acute ischaemic stroke under real-life practice conditions.

Selection of patients for exposure to treatment based on clinical features and physician preference instead of random allocation inevitably introduces opportunities for bias and confounding. According to the principles of *Good Research for Comparative Effectiveness Research* (GRACE)[1], and in line with the HTA recommendations for non-randomized studies[2], appropriate control of confounding variables together with rigorous pre-specification of analytical techniques is one of the primary requirements for high quality effectiveness research.

4.2 Effectiveness Evaluation

Ordinal analysis of the modified Rankin scale (mRS) at 3 months after stroke onset is chosen as clinically relevant primary endpoint for final treatment effects. Leading secondary endpoint is the NIHSS score on day 21 and 90.

The technical operationalizations for the first line analysis of the primary and secondary effectiveness measures, based on observed cases (OC) and target population (see section 4.9) are as follows:

- Primary outcome measures (effectiveness evaluation)
    1) Ordinal modified Rankin Scale (mRS) at 3 months after stroke onset, absolute scores, Mann-Whitney Effect Size (MW)[3,4,5,6,7], OC, Target Population
- Secondary outcome measures (effectiveness evaluation)
    2) Ordinal NIH Stroke Scale (NIHSS) at 21 days and 3 months after stroke onset, absolute scores, Mann-Whitney Effect Size (MW), OC, Target Population
    3) Ordinal modified Rankin Scale (mRS) at 21 days after stroke onset, absolute scores, Mann-Whitney Effect Size (MW), OC, Target Population
    4) Proportion of patients with excellent recovery (mRS score 0-1) at 3 months after stroke onset, Odds Ratio, OC, Target Population
    5) Proportion of patients with functional independence (mRS score 0-2) at 3 months after stroke onset, Odds Ratio, OC, Target Population
    6) Ordinal MoCA at 3 months after stroke, Mann-Whitney Effect Size (MW), OC, Target Population

According to the ICH Guideline E9 (ICH Topic E9, Statistical Principles for Clinical Trials, Step 4, Consensus guideline, 5 February 1998, CPMP/ICH/363/96) the results will be given as P-values as well as effect size measures with their associated confidence intervals (outcome no. 1, 2, 3, 6: Mann-Whitney effect size; outcome no. 4 and 5: odds ratio, supplemented by Mann-Whitney effect size for inter-outcome comparisons[8]), so that the direction and quantity of the treatment effects are determined with their precision.

The Mann-Whitney effect size3-6 is the most valuable effect size measure for nonparametric approaches based on the well-known Wilcoxon framework because it is valid in data situations where the Hodges-Lehmann shift parameter is no longer appropriate. Furthermore, the Mann-Whitney effect size is appropriate for continuous, ordinal and binary data at the same time and represents an ideal effect size measure. Incidentally, the 25th Anniversary of the journal Statistics in Medicine dedicated a whole issue to papers about the Mann-Whitney statistic3.

The Mann-Whitney effect size measure (MW) gives the probability that a randomly chosen subject of the test group is better off than a randomly chosen subject of the comparison group, defined in statistical shortcut: P (X<Y) + 0.5 P (X = Y). Applying the Mann-Whitney effect size measure, the null and alternative hypothesis for the comparison of Cerebrolysin vs. placebo can be formulated as follows (superiority test):

H_0: $MW_{TC} \leq 0.50$
H_A: $MW_{TC} > 0.50$

H_0: Null-hypothesis; H_A: Alternative Hypothesis; T: Test Treatment; C: Control
MW: Mann-Whitney Effect Size Measure

The traditional benchmarks for the Mann-Whitney effect size measure (MW) are as follows[9,10]:

0.29	large inferiority
0.36	medium inferiority
0.44	small inferiority
0.50	equality
0.56	small superiority
0.64	medium superiority
0.71	large superiority

The global alpha of the trial is 0.05 two-sided. The primary outcome measure will be analyzed according to the pre-defined Bauer-Koehne alpha for stage I of the trial (see section 4.5). The secondary outcomes will be analyzed using the same alpha, however, applying the principle of *a priori* ordered hypotheses (fixed sequence) for multiplicity control. If the test for superiority with respect to the primary outcome measure shows statistical significance, the secondary criteria can be tested with the same alpha as the first test with full control of the study-wise type I error. The sequence and nature of the a priori ordered test-statistical hypotheses is as defined above (outcome measures no. 1 - 6). The procedure of *a priori* ordered hypotheses is most powerful with full control of alpha (for control of alpha using stepwise testing see[11]).

4.3 Safety Evaluation

The operationalizations for the evaluation of the pre-defined safety measures, based on observed cases (OC) and intention-to-include (ITI) population, are as follows:

1) Mortality, Odds Ratio, OC, Target Population
2) Serious Adverse Events, Odds Ratio, OC, Target Population
3) Adverse Events, Odds Ratio, OC, Target Population

These safety measures will be used for group comparisons. In addition to the analysis of the Target Population with control of confounders, sensitivity analyses will be performed based on the ITI population (see section 4.7 and 4.9). Adverse drug reactions to Cerebrolysin (ADR), serious adverse drug reactions to Cerebrolysin (SADR) and suspected unexpected serious adverse reactions to Cerebrolysin (SUSAR) will be displayed for the Cerebrolysin treatment group.

4.4 Case-Mix Standardization

In order to minimize enrollment bias, the patient groups will be standardized using nonparametric multilevel stratification procedures in combination with a 'restricted cohort' design. The respective risk factors have been identified from previous research results on NIHSS predictor variables, allowing appropriate control for confounders of outcome after acute ischemic stroke[12,13,14]. The pre-specified case-mix standardization strategy follows the recommendations of the GRACE Principles for Good Research on Comparative Effectiveness1,[15].

Pre-Defined Clinical Predictor Variables[12]:

1. Initial NIHSS
2. Small Vessel Disease (yes-no)
3. Prior Stroke (yes-no)
4. Prior Diabetes (yes-no)
5. Prior Disability (yes-no)
6. Age

The operational definitions of variables no. 2 to 6 for consistent application across all participating sites are provided in the operational manual of the trial. The combination of the above variables has been shown to be a highly efficient predictor for outcome after ischemic stroke, making an additional control of infarct volume dispensible due to comparable areas under the receiver operator characteristic (ROC) curves[12].

The top level case-mix standardization will be based on the initial NIHSS score as one of the strongest predictors for outcome after stroke[12,13,14]. The top level control will be performed by implementing stratification per NIHSS score unit with subsequent meta-analytic pooling of strata (i.e., comparing groups within identical baseline NIHSS score). The eligibility restriction to NIHSS 8-15 allows full stratification for each possible baseline score (leading to a total of eight top level strata).

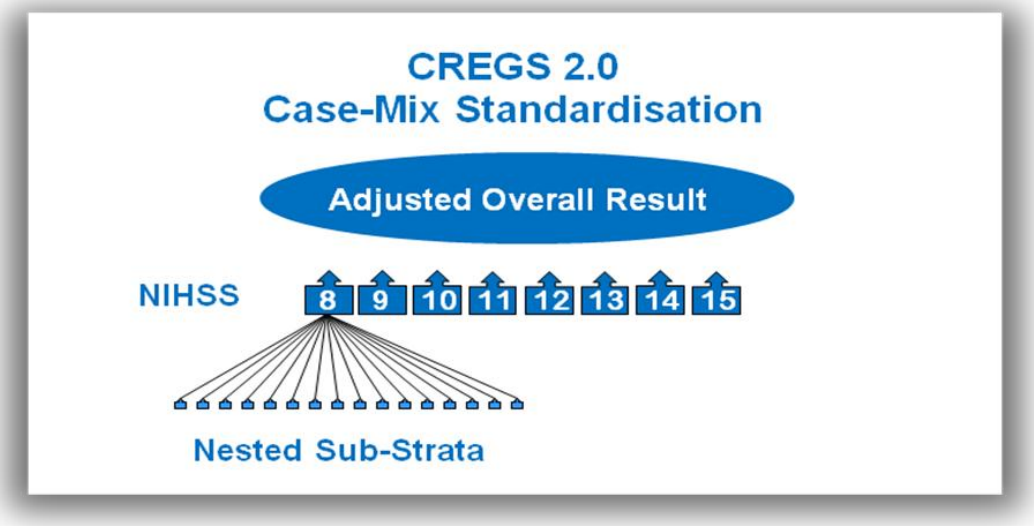
The pre-planned method for synthesis of the strata based on the primary Mann-Whitney (MW) effect size measure3-7 is the Wei-Lachin test of stochastic ordering (one-dimensional test)[16], a maximin-efficient robust test (MERT)[17,18] which provides a combined MW estimate and test of overall treatment effect from the pre-defined ensemble of independent strata (validated software package METASUB, Version 4.1, PROC STOCHASTIC ORDERING).

The second level case-mix standardization is performed for control of further confounders (see pre-defined clinical predictor variables). It is implemented within each of the top level NIHSS strata by means of nested sub-strata and subsequent adjustment by means of the Cochran-Mantel-Haenszel (CMH) pooling procedure (also known as the van-Elteren procedure)[19]. The nested sub-strata are based on the following pre-defined clinical predictor variables (see above):

- Diabetes (yes - no)
- Small Vessel Disease (yes - no)
- Age (< 65 - ≥ 65 years)

The combination of the three binary predictor variables results in a total of eight nested sub-strata. Technically the robust Peto-Wilcoxon test with CMH pooling of sub-strata will be applied, providing adjusted MW across sub-strata with associated confidence intervals (validated software package TESTIMATE, version 6.5.14, PROC PETO-WILCOXON). This procedure allows unbiased adjustment of ordinal or binary data also in the presence of very low sample sizes (only sub-strata with total N<3 are excluded from CMH analysis).

The advantage of the chosen multilevel case-mix standardization as compared to other, model-based approaches such as regression models is that any assumption about the nature of risk-outcome relation is avoided, allowing true like-to-like comparisons. Furthermore, some potential drawbacks of other procedures, as the model-based propensity score matching with its risk of bias due to incomplete matching, are reduced.



In addition to the specified multilevel case-mix standardization, controlling four out of the six pre-defined predictor variables, a specific method to strengthen observational, non-interventional studies is introduced for control of the two remaining confounders ('prior stroke' and 'prior disability'): the 'restricted cohort' design, i.e. patients are only eligible for this trial without prior stroke and without prior disability. This way any risk of bias associated to these two confounders can be avoided.

The described case-mix standardization for control of confounders will be performed for all comparative effectiveness evaluations, resulting in an adjusted overall effect size with associated confidence interval for each defined endpoint. The same applies to the comparative safety evaluations in order to minimize confounding (see, e.g., FDA Guidance to Industry on Reporting Safety Studies)[20]. Unadjusted safety analyses including all available patient data will be performed as additional sensitivity analysis (see section 4.7).

4.5 Two-Stage Procedure

The two-stage adaptive procedure of Bauer P and Köhne K (1994) is chosen as the sequential method.

The two-stage procedure based on Fisher's combination test (Bauer and Köhne) shows only a negligible loss in test power as compared to a fixed sample size study but allows early stopping due to success or failure[21]. Furthermore, assumptions for sample size calculation can be rechecked after stage I. The same applies to design modifications within the framework of the adaptive approach although this is not the rationale for introducing the two-stage procedure in this study. The formal Bauer-Köhne *futility* benchmark is set for this study to $\alpha_0 = 0.3$. It is important to note that this benchmark is taking into account the limited number of available patients for a possible stage II due to the restricted cohort design.

With a global multiple level alpha = 0.05 two-sided, and defined futility level of $\alpha_0 = 0.3$ the following decision structure will be formally established (p_1 = P-value of stage I, p_2 = P-value of stage II):

Decision Structure for Stage I results (two-sided)

$p_1 \geq \alpha_0 = 0.3$	: stop because of futility
$p_1 \in (0.0299; 0.3)$	: continue with stage II
$p_1 \leq \alpha_1 = 0.0299$	: stop with success (rejection of H_0)

Decision Structure for Stage II results (two-sided)

$p_1 p_2 > \alpha_c = 0.0087$	: stop because of futility
$p_1 p_2 \leq \alpha_c$	: rejection of H_o (proof of efficacy)

4.6 Lost to Follow Up

The current guidelines emphasize that every effort should be made to avoid missing data. In order to minimize loss of data and in order to comply with recent calls for high-quality non-interventional comparative effectiveness research, a risk-based centralized statistical approach to monitoring is introduced in combination with targeted on-site monitoring for ongoing surveillance of study conduct.

Regarding patients lost to follow up, a by patient listing of patients lost to follow up will be provided. Descriptive statistics will be presented for all patients lost to follow up as number and proportion of such patients in the Cerebrolysin and control group along with baseline demographics. The analyses will be performed for both, the ITI and the target population.

4.7 Sensitivity Analyses

Regarding the ordinal analysis of mRS and NIHSS, a missing value imputation will be performed as sensitivity analysis, using the last observation carried forward (LOCF) approach. Baseline values will not be carried forward.

With respect to time to hospital, a sensitivity analysis will be performed for the primary endpoint (mRS Day 90) using the quartiles of the distribution as sub-strata with adjustment by means of CMH procedure and top level stratification for initial NIHSS.

With respect to the comparative safety evaluation, a non-adjusted analysis based on the ITI population (see section 4.9), including all available patient data, will be performed as sensitivity analysis.

Further sensitivity analyses based on comparable procedure may be performed for additional baseline variables of interest.

4.8 Descriptive Statistics

4.8.1 Dichotomous and Categorical Variables

Categorical data will be presented in frequency tables using counts and percentages. Percentages will be based on the total number of patients in the ITT and ITI analysis set, unless otherwise specified.

4.8.2 Continuous and Quasi-Continuous Variables

Standard descriptive summary statistics will be calculated for continuous and quasi-continuous variables: arithmetic mean, standard deviation, minimum value, lower quartile, median, upper quartile, maximum value, number of non-missing values.

4.9 Analysis Sets

Screening failures will be entered into the database if consent was granted to hold their data. All patients, including those technically excluded from the multilevel stratification procedures ($N_{substratum}$ <3), both treated and control, will be included in the patient listing.

All eligible patients who give consent to participate in the study and are entered into the database will be included into the intent-to-include (ITI) population irrespective of the multilevel stratification process. All effectiveness analyses will be summarized by treatment received. Assignment to a treatment group (Cerebrolysin vs. Control) is non-randomized, thus determined on a case by case basis.

The Target population will consist of all patients in the ITI population who can technically be included into the multilevel stratification process. Inclusion is defined by the minimum sub-strata size (sub-strata with total N<3 are excluded from analysis), as well as by the availability of the six stratification variables (pre-defined clinical predictor variables, see section 4.4). The analyses of the target population represents the principal results of the study; ITI analyses are performed for additional description of baseline characteristics and for safety sensitivity analyses.

4.10 Software Utilized

Sample size calculation was performed using the validated software packages Nnpar 1.0 and Bauer-Köhne 4.0 from IDV Data Analysis and Study Planning (Krailling/Munich, Germany). The statistical analyses will be performed using the validated software packages TESTIMATE (Version 6.1.14), ScienceGraph (Version 4.9.39), METASUB (Version 4.1), and ForestPlot (Version 4.1) from IDV Data Analysis and Study Planning (Kralling/Munich, Germany) on high security PCs (HSPC) within a validated working environment at the department 'Clinical Research/Biometry', IDV Data Analysis and Study Planning (Krailling/Munich, Germany), under supervision of Volker W. Rahlfs, PhD., C. Stat. (RSS), with a 'Certificate Biometry in Medicine GMDS'.

4.11 Post Hoc Changes to Planned Analyses

Due to the observational character of this trial, any major changes of the statistical analysis plan after First Patient First Visit (FPFV) will be regarded as *post hoc* and will be specified in the Clinical Study Report of the final analyses with corresponding scientific rationale. The version history of the statistical analysis plan will be documented throughout the whole trial.

5. PROSPECTIVE META-ANALYSIS (PMA)

Meta-analytic techniques are recognized as a useful tool to summarize the overall efficacy results of a drug application (ICH E9 biostatistical guideline, CPMP/EWP/2330/99). An extension of this approach is prospective meta-analysis (PMA) in which studies are identified, evaluated and determined to be eligible before the results of any of the studies become known[22] (see also *Prospective Meta-analysis*, Cochrane Handbook for Systematic Reviews of Interventions; Part 3, Chapter 19. http://handbook.cochrane.org).

Currently two registry trials on Cerebrolysin after stroke are planned (C-REGS 2) or ongoing (CREGS-S) with similar endpoints and a 90 day follow-up:

- C-REGS 2 (EVER-AT-0717)
- CREGS-S (EVER-GB-0514)

A meta-analytic combination of these two registry studies after is regarded as useful complement to the individual study analyses.

Therefore, after study terminations, the data of the present C-REGS 2 registry trial will be combined with the data of the CREG-S registry trial by formal meta-analysis procedures in order to gain further insight into effectiveness of Cerebrolysin after stroke. For this purpose, and for ensuring consistent analysis data, all C-REGS 2 definitions for effectiveness analysis, operationalized in this SAP, will be applied to the CREGS-S data. This is achieved by using individual patient (IPD) data analysis of both trials, the gold standard for meta-analytic pooling[23].

The principle features of the planned meta-analysis reflect the blinded *a priori* definitions in this SAP for the analysis of the C-REGS 2 study. The meta-analysis will be conducted on the following endpoints:

- Primary outcome measures (effectiveness evaluation)
    1) Ordinal modified Rankin Scale (mRS) at 3 months after stroke onset, absolute scores, Mann-Whitney Effect Size (MW)[3-7], OC, Target Population as operationalized in the C-REGS 2 SAP
- Secondary outcome measures (effectiveness evaluation)
    1) Ordinal NIH Stroke Scale (NIHSS) at 3 months after stroke onset, absolute scores, Mann-Whitney Effect Size (MW), OC, Target Population as operationalized in the C-REGS 2 SAP
    2) Proportion of patients with excellent recovery (mRS score 0-1) at 3 months after stroke onset, Odds Ratio, OC, Target Population as operationalized in the C-REGS 2 SAP

3) Proportion of patients with functional independence (mRS score 0-2) at 3 months after stroke onset, Odds Ratio, OC, Target Population as operationalized in the C-REGS 2 SAP
4) Ordinal MoCA at 3 months after stroke, Mann-Whitney Effect Size (MW), OC, Target Population as operationalized in the C-REGS 2 SAP

These endpoints are identical with the definitions in this SAP for C-REGS 2, except Day 21 endpoints, since the CREG-S trial does not involve any Day 21 assessments.

The pre-planned method of synthesis for the primary Mann-Whitney (MW) effect size measure[3-7] is the Wei-Lachin test of stochastic ordering (one-dimensional test)[16], a maximin-efficient robust test (MERT)[17,18] which provides a combined MW estimate and test of overall treatment effect from an ensemble of independent studies. This approach is 'assumption-free' and has been shown to be robust also with respect to presence of heterogeneity[16]. Qualitative interaction will be tested by means of the Gail-Simon test[24], with P-values < 0.10 preventing formal combination of studies.

As sensitivity analysis the "classic" approaches based on fixed effects model (Hedges-Olkin)[25] and random effects model (DerSimonian-Laird)[26] will be calculated. Associated tests for quantitative heterogeneity will be performed using standard chi-square statistic[27] and I^2 statistic[28].

The meta-analyses will be performed using the software packages METASUB (Version 4.1), and ForestPlot (Version 4.1) from IDV Data Analysis and Study Planning (Krailling/Munich, Germany) on high security PCs (HSPC) within a validated working environment at the department 'Clinical Research/Biometry' in the institute IDV Data Analysis and Study Planning (Krailling/Germany) under supervision of Volker W. Rahlfs, PhD., C. Stat. (RSS), with a 'Certificate Biometry in Medicine GMDS'.

.

6. ADDITIONAL SPECIFICATIONS TO THE STATISTICAL ANALYSIS PLAN VERSION 1.0 (2017)

There are no major changes to the version Final 1.0 of the Statistical Analysis Plan (SAP) from 24.10.2017.

Additional specifications and further operationalizations are documented in the following sections 6.1 to 6.5.

6.1 Sensitivity Analyses

6.1.1 Specifications in Final SAP Version 1.0 from 24.10.2017

Regarding the ordinal analysis of mRS and NIHSS, a missing value imputation will be performed as sensitivity analysis, using the last observation carried forward (LOCF) approach. Baseline values will not be carried forward.

With respect to time to hospital, a sensitivity analysis will be performed for the primary endpoint (mRS Day 90) using the quartiles of the distribution as sub-strata with adjustment by means of CMH procedure and top level stratification for initial NIHSS.

With respect to the comparative safety evaluation, a non-adjusted analysis based on the ITI population (see section 4.9), including all available patient data, will be performed as sensitivity analysis.

Further sensitivity analyses based on comparable procedure may be performed for additional baseline variables of interest.

6.1.2 Additional Specifications in Final SAP Version 1.1 from 22.07.2024

As unanimously decided on the 16th SAB Meeting from December 3, 2023, the following additional sensitivity analyses are implemented in accordance with the requirements of the ICH E9 Biostatistics Guidance[29] to evaluate the robustness of the results and primary conclusions of the trial.

Pre-specified subgroups

1) High enrollers – enrolment excess

   High enrolling countries might dominate the primary results of the study by substantially higher sample sizes as compared to other countries. In order to ensure the robustness and generalizability of the primary results, 'leave-one-out' analyses are performed for countries with substantial "enrolment excess" (sample size >50% more than any other country).
   This is the case for Ukraine and Russia as each of these countries is contributing over 50% more patients than any other single country.

In order to capture the full range of possible interference, the following sensitivity analyses are defined with respect to the two countries (exploratory interpretation):

- RU stand alone
- UA stand alone
- Non-Russia countries leave-one-out
- Non-Ukraine countries leave-one-out
- Non-Russia/non-Ukraine leave-two-out

Additional SAB comment

Since the two exceeding countries are currently at war, the sensitivity analyses as defined above shall additionally exclude any potential bias by special circumstances in these countries.

2) Age

< 65 years*
≥ 65 years*

**This is the pre-specified age cut-off as defined in the Statistical Analysis Plan Version Final 1.0 from 2017.*

< 75 years*
≥ 75 years*

**Upon SAB recommendation, an additional cut-off at ≥ 75 years is introduced to consider eventual changes due to the time passed since the study planning and improved stroke care.*

3) Cerebrolysin Dosage

Dose of < 30 ml per day*
Dose of ≥ 30 ml per day*

**cutoff at overall median*

The comparisons will be executed between each dose group and standard of care, as well as between the two Cerebrolysin dose groups as such.

4) Baseline NIHSS

Baseline NIHSS ≤ 9*
Baseline NIHSS > 9*

**cutoff at overall median*

Baseline NIHSS ≤ 12*
Baseline NIHSS > 12*

**Upon SAB recommendation, the cutoff of NIHSS = 12 is introduced for additional severity alignment with the pre-specified cut-off of the CASTA[30] trial.*

Additional Sensitivity Analyses

An analysis of non-fatal mRS (mRS scores 0 to 5) will be performed in addition to the primary full scale mRS (mRS scores 0 to 6) to provide insight on the outcome of stroke survivors.

Non-fatal Serious Adverse Events (non-fatal SAE) will be evaluated in addition to Serious Adverse Events (SAE) due to recent discussions based on non-fatal SAE.

With respect to the primary ordinal effectiveness outcomes, an ANCOVA with adjustment for baseline NIHSS scores will be performed as a sensitivity analysis. This sensitivity analysis will be implemented for the target population as well as for the ITI population.

For binary outcomes, analyses based on risk ratios (RR) will be provided in addition to the pre-specified analyses based on odds ratios (OR). Rationale: alignment with publications/guidelines based on RRs only.

6.2 Safety Analyses

6.2.1 Specifications in Final SAP Version 1.0 from 24.10.2017

Section 4.3 Safety Evaluation:
The operationalizations for the evaluation of the pre-defined safety measures, based on observed cases (OC) and intention-to-include (ITI) population, are as follows:

1. Mortality, Odds Ratio, OC, Target Population
2. Serious Adverse Events, Odds Ratio, OC, Target Population
3. Adverse Events, Odds Ratio, OC, Target Population

These safety measures will be used for group comparisons. In addition to the analysis of the Target Population with control of confounders, sensitivity analyses will be performed based on the ITI population (see section 4.7 and 4.9). Adverse drug reactions to Cerebrolysin (ADR), serious adverse drug reactions to Cerebrolysin (SADR) and suspected unexpected serious adverse reactions to Cerebrolysin (SUSAR) will be displayed for the Cerebrolysin treatment group.

Section 4.7 Sensitivity analyses:
With respect to the comparative safety evaluation, a non-adjusted analysis based on the ITI population (see section 4.9), including all available patient data, will be performed as sensitivity analysis.

6.2.2 Additional Specifications in Final SAP Version 1.1 from 22.07.2024

As defined in the SAP Version 1.0, the first line safety analysis will be based on the target population, a sensitivity analysis will be performed on the ITI population. The ITI analysis remains unadjusted for case-mix standardization substrata, as the latter are only available in the target population (see Section 6.3.2). However, due to highly "zero-dominated" data (low overall rate of reported side effects), a case-mix standardization using substrata as pre-specified for effectiveness evaluation will technically also not be feasible for the target population. Thus, in agreement with the decision on the 16th SAB Meeting from December 3, 2023, stroke severity will be used as stratifying factor to ensure case-mix standardization for both, target and ITI analysis:

> *AEs will be adjusted for stroke severity at baseline as usually patients with higher stroke severity experience more events than patients with less severe stroke. Adjustment will be performed by means of stratification with subsequent meta-analytic pooling (top level adjustment as predefined for the efficacy criteria).*

Technically, the meta-analytic pooling of the pre-specified NIHSS severity strata will be performed for all safety event data by means of the Mantel-Haenszel (M-H) procedure as established in the validated computer program RevMan[31] (The Cochrane Collaboration). Primary analysis of the safety events will be based on odds ratio (OR). Analysis based on risk ratios (RR) will be provided in addition as sensitivity analysis.

In case of indication for heterogeneity among adverse event results, the random effects model (DerSimonian-Laird)[26] will be used for all safety outcomes, else the fixed effects model (Hedges-Olkin)[25]. The benchmarks for determining heterogeneity are pre-specified as follows: $I^2 > 0.3$ and/or $P_{Heterogeneity} < 0.1$.

6.3 Analysis Sets

6.3.1 Specifications in Final SAP Version 1.0 from 24.10.2017

Section 4.9 Analysis Sets:
Screening failures will be entered into the database if consent was granted to hold their data. All patients, including those technically excluded from the multilevel stratification procedures ($N_{substratum}$ <3), both treated and control, will be included in the patient listing.

All eligible patients who give consent to participate in the study and are entered into the database will be included into the intent-to-include (ITI) population irrespective of the multilevel stratification process. All effectiveness analyses will be summarized by treatment received. Assignment to a treatment group (Cerebrolysin vs. Control) is non-randomized, thus determined on a case by case basis.

The Target population will consist of all patients in the ITI population who can technically be included into the multilevel stratification process. Inclusion is defined by the minimum sub-strata size (sub-strata with total N<3 are excluded from analysis), as well as by the availability of the six stratification variables (pre-defined clinical predictor variables, see section 4.4). The analyses of the target population represents the principal results of the study; ITI analyses are performed for additional description of baseline characteristics and for safety sensitivity analyses.

6.3.2 Additional Specifications in Final SAP Version 1.1 from 22.07.2024

The intent to include (ITI) population includes all eligible subjects who give consent to participate in the study, have at least one baseline assessment entry, and provide any follow-up effectiveness or safety data (including death), irrespective of the multilevel stratification process. The ITI analyses are unadjusted for case-mix standardization substrata, as the latter are only available in the target population (see below).

The target population, on which the primary analysis of the C-REGS 2 study relies, consists of all subjects in the ITI population with existing data for the multilevel case-mix standardization process.

6.4 Two-Stage Procedure

6.4.1 Specifications in Final SAP Version 1.0 from 24.10.2017

Section 4.5 Two-Stage Procedure:

The two-stage adaptive procedure of Bauer P and Köhne K (1994) is chosen as the sequential method.

The two-stage procedure based on Fisher's combination test (Bauer and Köhne) shows only a negligible loss in test power as compared to a fixed sample size study but allows early stopping due to success or failure[21]. Furthermore, assumptions for sample size calculation can be rechecked after stage I. The same applies to design modifications within the framework of the adaptive approach although this is not the rationale for introducing the two-stage procedure in this study. The formal Bauer-Köhne *futility* benchmark is set for this study to $\alpha_0 = 0.3$. It is important to note that this benchmark is taking into account the limited number of available patients for a possible stage II due to the restricted cohort design.

With a global multiple level alpha = 0.05 two-sided, and defined futility level of $\alpha_0 = 0.3$ the following decision structure will be formally established (p_1 = P-value of stage I, p_2 = P-value of stage II):

Decision Structure for Stage I results (two-sided)

$p_1 \geq \alpha_0 = 0.3$	: stop because of futility
$p_1 \in (0.0299; 0.3)$	: continue with stage II
$p_1 \leq \alpha_1 = 0.0299$	: stop with success (rejection of H_0)

Decision Structure for Stage II results (two-sided)

$p_1 p_2 > \alpha_c = 0.0087$	: stop because of futility
$p_1 p_2 \leq \alpha_c$	: rejection of H_o (proof of efficacy)

6.4.2 Additional Specifications in Final SAP Version 1.1 from 22.07.2024

Due a slow initial enrolment progress and the necessity to included additional countries, the supervising C-REGS2 Enrollment Working Group (EWG) decided in 2020 that the Stage I will be prematurely terminated in favour of a seamless transition to an enlarged Stage II (without interim analysis, i.e. omitting the decision structure for Stage I). Excerpt from the EWG Report:

> *In order to avoid an excessive Stage I duration and to allow inclusion of a substantial amount of additional countries and sites without late stage I mix-up, a premature Stage I termination is recommended with seamless transition to stage II. This way the enlarged country/site ensemble is separated from the current stage I ensemble, allowing more balanced within-stage comparisons for confirmatory proof.*

The pre-specified Bauer-Köhne decision structure, based on Fisher's combination test for the pooling of the Stage I and Stage II results (see section 4.5), will be applied as planned for the first line analysis of the *a priori* ordered effectiveness outcomes (see Section 4.2). Due to the omittance of any interim analysis, Stage I and II will be evaluated for all other purposes as combined data set.

6.5 Unblinding

6.5.1 Specifications in Final SAP Version 1.0 from 24.10.2017

None.

6.5.2 Additional Specifications in Final SAP Version 1.1 from 22.07.2024

Due to the high-quality comparative effectiveness research approach (HQCER) of the trial[1], the independent study biostatisticians remained blinded to treatment throughout the whole course of the study, not receiving any treatment-related individual or aggregate outcome information. All pre-planned risk-based monitoring (RBM) procedures[32,33] (see SAP section 3., 'Quality Assurance') were executed by separate technical staff not involved in statistical issues or the operationalizations of the Statistical Analysis Plan, respectively.

After final database lock and SAP finalization, the technical unblinding procedures with subsequent generation of the analysis database will be executed with audit trail records, applying the recommendations of the ICH E9 Biostatistics Guideline[29]:

> *The plan should be reviewed and possibly updated as a result of the blind review of the data (see 7.1 for definition) and should be finalised before breaking the blind. Formal records should be kept of when the statistical analysis plan was finalised as well as when the blind was subsequently broken.*

7. SIGNATURES

Statistician	Prof. Johannes Vester Idv Data Analysis & Study Planning Department of Biometry and Clinical Research Gauting Germany
	_____ Date Signature
Review	Volker W. Rahlfs, PhD, CStat (RSS) Idv Data Analysis & Study Planning Department of Biometry and Clinical Research Gauting Germany
	_____ Date Signature

Coordinating investigator	Univ. Prof. Dr. Michael Brainin Donau-Universität Krems Department for Clinical Neuroscience and Preventive Medicine Dr.-Karl-Dorrek-Straße 30 3500 Krems Austria _____ Date Signature
Sponsor	EVER Neuro Pharma GmbH Oberburgau 3 4866 Unterach Austria
Sponsor's representative	Stefan Winter, PhD Oberburgau 3 4866 Unterach Austria +43 7665 20555 422 stefan.winter@everpharma.com _____ Date Signature

8. REFERENCES

[1] Dreyer NA, Schneeweiss S, McNeil B, et al. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. Am J Manag Care. 2010;16(6):467-71.

[2] Deeks JJ, Dinnes J, D'Amico R, et al. Evaluating non-randomised intervention studies. Health Technology Assessment 203; Vol. 7: No. 27.

[3] D'Agostino RB, Campbell M, Greenhouse J. The Mann–Whitney statistic: continuous use and discovery. Statist Med. 2006;25:541-542. doi: 10.1002/sim.2508.

[4] Rothmann MD, Wiens BL, Chan IS. Design and Analysis of Non-Inferiority Trials. Boca Raton, Fla.: Chapman & Hall/CRC; 2011.

[5] Munzel U, Hauschke D. A nonparametric test for proving non-inferiority in clinical trials with ordered categorical data. Pharmaceut Statist. 2003;2:31-37. doi: 10.1002/pst.17.

[6] Kieser M, Friede T, Gondan M. Assessment of statistical significance and clinical relevance. Stat Med. 2013;32:1707–1719.

[7] Rahlfs VW, Zimmermann H, Lees KL. Effect size measures and their relationships in stroke studies, Stroke. 2014;45:627-633

[8] Rahlfs, V. W., Zimmermann, H., & Lees, K. R. (2014). Effect size measures and their relationships in stroke studies. Stroke, 45(2), 627-633.

[9] Cohen, J., Statistical power analysis for the behavioral sciences, Lawrence Erlbaum Associates. Hillsdale, NJ 1988.

[10] Colditz, G.A., Miller, J.N., Mosteller F., Measuring gain in the evaluation of medical technology. The probability of a better outcome. Int J Technol Assess Health Care 1988; 4(4):637-42.

[11] Maurer W, Hothorn LA, Lehmacher W (1995). Multiple comparisons in drug clinical trials and preclinical assays: a-priori ordered hypotheses. In: Testing Principles in Clinical and Pre-clinical Trials, ed. Von J. Vollmar, Biometrie in der chemisch-pharmazeutischen Industrie 6, Gustav Fischer, Stuttgart, 3-18

[12] Johnston KC, Connors AF, Wagner DP, et al. A Predictive Risk Model for Outcomes of Ischemic Stroke. Stroke. 2000;31:448-455.

[13] DeGraba TJ, Hallenbeck JM, Pettigrew KD, et al. Progression in Acute Stroke – Value of Initial NIH Stroke Scale Score on Patient Stratification in Future Trials. Stroke. 1999;30:1208-1212.

[14] Adams HP, Davis PH, Leira EC, et al. Baseline NIHS Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999; 53(1):126-131.

[15] Dreyer NA, Velentgas P, Westrich K et al. The GRACE Checklist for Rating the Quality of Observational Studies of Comparative Effectiveness: A Tale of Hope and Caution. JMCP. 2014;20(3):301-308d.

[16] Lachin JM. Biostatistical methods. The assessment of relative risks, New York, Wiley 2000.

[17] Frick H (1994) A maximin linear test and its application to Lachin's Data. Commun Statist - Theory and Methods 23: 1021–1029.

[18] Frick H (1995) Comparing trials with multiple outcomes: The multivariate one-sided hypothesis with unknown covariances. Biom J 8: 909–917.

[19] Van Elteren PH. On the combination of independent two-sample tests of Wilcoxon. Bulletin of the Institute of International Statistics. 1960;37:351-361.

[20] Guidance for Industry and FDA Staff - Best Practices for Conductiong and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data.. FDA May 2013. Rockville, USA.

[21] Banik N, Köhne K, Bauer P. On the Power of Fisher's combination test for two stage sampling in the presence of nuisance parameters, Biometrics 1996;38:25-37

[22] Roetzheim RG, Freund KM, Corle DK, Murray DM, Snyder FR, Kronman AC. Analysis of Combined Data from gheterogeneous Study Designs; A Methodlogical Proposal from the patient Navigation Research Program. Clin Trials 2012;9(2):176-187.

[23] Riley, R. D., Lambert, P. C., Staessen, J. A., Wang, J., Gueyffier, F., Thijs, L., & Boutitie, F. (2008). Meta-analysis of continuous outcomes combining individual patient data and aggregate data. Statistics in medicine, 27(11), 1870-1893.

[24] Gail M, Simon R. Testing for qualitative interactions between treatment effects and patient subsets. Biometrics 1985; 41: 361-372.

[25] Hedges, L.V., Olkin, J. (1985). Statistical Methods for Meta-Analysis, Academic Press Inc., San Diego.

[26] DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. Controlled clinical trials, 7(3), 177-188.

[27] Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. BMJ: British Medical Journal, 327(7414), 557.

[28] Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in medicine, 21(11), 1539-1558.

[29] ICH Harmonised Guideline. ICH Topic E9: Statistical Principles for Clinical Trials - Step 5. CPMP/ICH/363/96. September 1998.

[30] Heiss WD, Brainin M, Bornstein NM, Tuomilehto J, Hong Z, Cerebrolysin Acute Stroke Treatment in Asia (CASTA) Investigators. Cerebrolysin in patients with acute ischemic stroke in Asia: results of a double-blind, placebo-controlled randomized trial. Stroke 2012;43(3):630–6

[31] Review Manager (Revman) [Computer program]. Version 5.3, Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

[32] European Medicine Agency. Reflection Paper on risk based quality management in clinical trials, EMA/269011/2013

[33] FDA. Guidance for Industry: Oversight of Clinical Investigations - A Risk-Based Approach to Monitoring, U.S. Department of Health and Human Services, Food and Drug Administration. August 2013