I.R.I.S.

INSTITUT DE RECHERCHES INTERNATIONALES SERVIER

Document title	**STATISTICAL ANALYSIS PLAN (SAP) ATTACHED TO CL3-95008-002 STUDY**
Study title	**Efficacy and safety of Bumetanide oral liquid formulation in children aged from 2 to less than 7 years old with Autism Spectrum Disorder**
Test drug code	**Bumetanide – S95008**
Indication	**Autism Spectrum Disorder**
Development phase	**Phase III**
Protocol code	**CL3-95008-002**
EudraCT Number	**2017-004420-30**
ClinicalTrial.gov Number	**NTC03715153**
Universal Trial Number	**Not applicable**
Sponsor	**Institut de Recherches Internationales Servier (I.R.I.S.) 50 Rue Carnot 92284 Suresnes Cedex - France**
Date of the document	**21 July 2021**
Version of the document	**Version 3.0**
	CONFIDENTIAL

Signatories

Prepared by:	Signature	Date
[redacted]	[redacted]	
I.R.I.S. - FRANCE		
Approved by:	Signature	Date
[redacted]	[redacted]	
I.R.I.S. - FRANCE		
[redacted]	[redacted]	
I.R.I.S. - FRANCE		

Follow up of versions

Version	Release date (dd/mm/yyyy)	Key modifications (*)	Impact
1.0	12/11/2018	Definition of the primary estimand added	3.4.2.1. Primary analysis
		Additional sensitivity analysis	3.4.2.2 Sensitivity analyses
		MMRM moved from sensitivity analyses to supplementary analyses	3.4.2.2 Sensitivity analyses 3.4.2.3 Supplementary analyses
		Missing data handling, sensitivity analysis and supplementary analyses added for CGI-I analysis	3.4.3.2 CGI-I
		Main analysis performed on the Adaptative Behavior Composite score instead of the domain scores	3.4.3.3 VABS II
		Supplementary analyses added for VABS II analysis	3.4.3.3 VABS II
		Subgroup analyses added	3.4.4 Subgroups analysis
		Safety analyses on open-label period replaced by analyses on combined periods (double-blind + open-label) for adverse events and laboratory parameters	3.6. Safety analysis
		Analysis of C-SSRS added	3.6.2 C-SSRS
		Additional listings and analyses for potassium parameter	3.6.3 Clinical laboratory evaluation
		Additional ECG analyses	3.6.4.2 Electrocardiogram
2.0	09/12/2019	Add justifications on the choice of the estimand based on comments from FDA.	3.4.2.1. Primary analysis
3.0	21/07/2021	Additional analysis set and periods Additional analyses related to Covid-19 pandemic. Additional analyses related to the optional extension part Additional CARS2 responders analysis.	All sections impacted

		Additional analysis for the VABS II domain Additional analysis for quality of life questionnaires Subgroups analysis were deleted Details on analysis period for safety were added.	

(*) Main changes as compared to the statistical analyses planned in the protocol

Table of contents

List of tables

List of figures

List of abbreviations

%	:	percentage
µmol	:	micromole
AE	:	Adverse Event
b.i.d.	:	bis in die (twice a day)
BMI	:	Body Mass Index
BP	:	Blood Pressure
b.p.m	:	beats per minute (heart rate unit)
CARS2	:	Childhood Autism Rating Scale, Second Edition
CGI	:	Clinical Global Impression Scale
CGI–I	:	Clinical Global Impression Scale – Global Improvement
CGI–S	:	Clinical Global Impression Scale – Severity
CHMP	:	Committee for Medicinal Products for Human Use
CI	:	Confidence Interval
cm	:	centimetre
CPK	:	Creatine PhosphoKinase
CPMP	:	Committee for Proprietary Medicinal Products
C–SSRS–C	:	Columbia Suicide Severity Rating Scale Children's version
DBP	:	Diastolic Blood Pressure
DMC	:	Data Monitoring Committee
e.g.	:	exempla gratia (for example)
EAE	:	Emergent Adverse Event
ECG	:	ElectroCardioGram
e–CRF	:	electronic-Case Report Form
EMA	:	European Medicines Agency
EQ–5D–3L	:	EuroQol – 5-Dimension – Three-level
g	:	gram
G/L	:	Giga (109) per litre
GLM	:	General Linear Model
h	:	hour
HR	:	Heart Rate
i.e.	:	id est
I.R.I.S.	:	Institut de Recherches Internationales Servier
ICH	:	International Conference on Harmonization
IME	:	Important Medical Event
IMP	:	Investigational Medicinal Product
IRS	:	Interactive Response System
IS	:	Included Set
IU	:	International Unit
IVRS	:	Interactive Voice Response System
IWRS	:	Interactive Web Response System
kg	:	kilogram
L	:	Litre
LLN	:	Lower Limit of Normal laboratory reference range
LLS	:	Lower Limit used to define potentially clinically Significant abnormal values
m	:	metre
Max	:	Maximum

MedDRA	:	Medical Dictionary for Regulatory Activities
mg	:	milligram
MI	:	Multiple Imputation
min	:	minute
Min	:	Minimum
mL	:	millilitre
mm	:	millimetre
mmHg	:	millimetre of mercury
mmol	:	millimole
MMRM	:	Mixed-effects Model for Repeated Measures
msec	:	millisecond
NA	:	Not Applicable
Na+	:	Sodium
NAE	:	Number of Adverse Events
NEAE	:	Number of Emergent Adverse Events
ng	:	nanogram
NPD	:	Number of Protocol Deviations
PCSA	:	Potentially Clinically Significant Abnormal value
PedsQL	:	Paediatric Quality of Life Inventory
po	:	per os (orally)
PPS	:	Per Protocol Set
PT	:	Preferred Term
PV	:	PharmacoVigilance
QTc	:	QT interval corrected for heart rate
RGLM	:	Robust General Linear Model
RS	:	Randomised Set
s	:	second
SAE	:	Serious Adverse Event
SAP	:	Statistical Analysis Plan
SBP	:	Systolic Blood Pressure
SD	:	Standard Deviation
SE	:	Standard Error
SEAE	:	Serious Emergent Adverse Event
SOC	:	System Organ Class
SRS–2	:	Social Responsiveness Scale, Second Edition
SS	:	Safety Set
T/L	:	Tera (1012) per litre
TLG	:	Tables, Listings and Graphs
TU	:	Therapeutic Unit
ULN	:	Upper Limit of Normal laboratory reference range
ULS	:	Upper Limit used to define potentially clinically Significant abnormal values
WHO	:	World Health Organization
WHO–DRL	:	World Health Organization, Drug Reference List
WHOQOL–BREF	:	Abbreviated World Health Organization Quality of Life

1. INTRODUCTION

This Statistical Analysis Plan details the planned analyses to be performed, in accordance with the main characteristics of the amended study protocol. The templates for Tables, Listings and Graphs (TLG) are described in a separate document.

The study was divided into three periods:

- Double-blind treatment period of 6 months between inclusion (week 0) and month 6 (week 26).
- Open label active treatment period of 6 months between month 6 (week 26) and month 12 (week 52).
- 6-month optional extension period with bumetanide proposed in some countries where it was not possible to provide the treatment at the end of the main periods *via* a Named Patient Basis (NPB) access or a Post-Access Study Program.

This Statistical Analysis Plan detailed the analysis associated to these three periods. The main analysis will be the first one based on the double-blind period. The blind will be broken as soon as all efficacy and safety data collected during the double-blind are available and all the mandatory steps required for study unblinding have been performed. Then, only data of the double-blind period will be analysed following the first database-lock (and unblinding), the analyses of the other periods (open-label and extension) will be performed at the end of each period.

Nevertheless, although the blind will be broken before the end of the study, neither the investigators, nor the raters, nor the patients, nor the parents/legal representatives/caregivers, nor the monitors will be informed of the study treatment taken during the double-blind treatment period before the database lock at the end of the open-label period.

Of note, this SAP does not cover the pharmacokinetic data analyses described in the study protocol. These analyses are covered in separate analysis plans written by I.R.I.S. Clinical Pharmacokinetics.

According to the EMA "*Guidance on the management of clinical trials during the COVID-19 (coronavirus) pandemic*", a risk assessment was performed by the Sponsor on aggregated and blinded data to evaluate the implications on recruitment, loss of patients during the trial, ability to record data and ability to interpret the treatment effect. The risk assessment is described in a separate document.

1.1. Study objectives

The primary objective is to demonstrate the superiority of bumetanide (0.5mg BID) oral liquid formulation compared to placebo in the improvement of ASD core symptoms after 6 months of treatment in ASD children and adolescents aged from 2 to less than 7 years old.

The secondary objectives are:

- To assess the effect of bumetanide on the other efficacy endpoints.
- To assess the safety of bumetanide.
- To confirm the acceptability and palatability of the oral liquid formulation.
- To describe the bumetanide effects on patients quality of life.
- To improve existing pharmacokinetic model of bumetanide in this population.

1.2. Study design

The study CL3-95008-002 is a phase III, international, multicentre, 6-month, randomised, double-blind, placebo-controlled study with two parallel groups (bumetanide or placebo) followed by an open-label active 6-month treatment period with bumetanide and a 6-week discontinuation period after treatment stop.

1.2.1. Study plan

The study is divided into the following periods:

- A run-in period without treatment up to 4 weeks (from ASSE to W0) to allow investigators to perform ECG, renal ultrasound, laboratory examinations and to respect the wash-out period of any unauthorised treatment before inclusion.
- A double-blind period of 6 months (from W000 to W026): at W000, patients are randomised (1:1) in double-blind conditions to one of the two treatment groups: bumetanide or placebo.
- An open-label period of 6 months (from W026 to W052):
  - Patients randomised in the S95008 arm during the double-blind period will continue to be treated by S95008 during the whole open-label period.
  - Patients randomised in the placebo arm during the double-blind period will receive bumetanide during the whole open-label period.
- A follow-up period of 6 weeks after treatment discontinuation (from W052 to WEND). A WEND visit should be planned 6 weeks after Investigational Medicinal Product (IMP) stop (W052 or premature withdrawal).
- A 6-month optional extension period with Bumetanide is proposed in some countries where it is not possible to supply Bumetanide (study formulation) at the end of the follow up period via a Named Patient Basis (NPB) access or a Post-Access Study Program. This extension is optional and proposed to the patients based on the investigator's clinical judgement (clinical benefit observed during the main study without no major safety concern).

IMPs are supplied as two oral administrations per day with 1 mL graduated pipette upon waking in the morning and in the afternoon, 3 hours before going to bed at the latest.

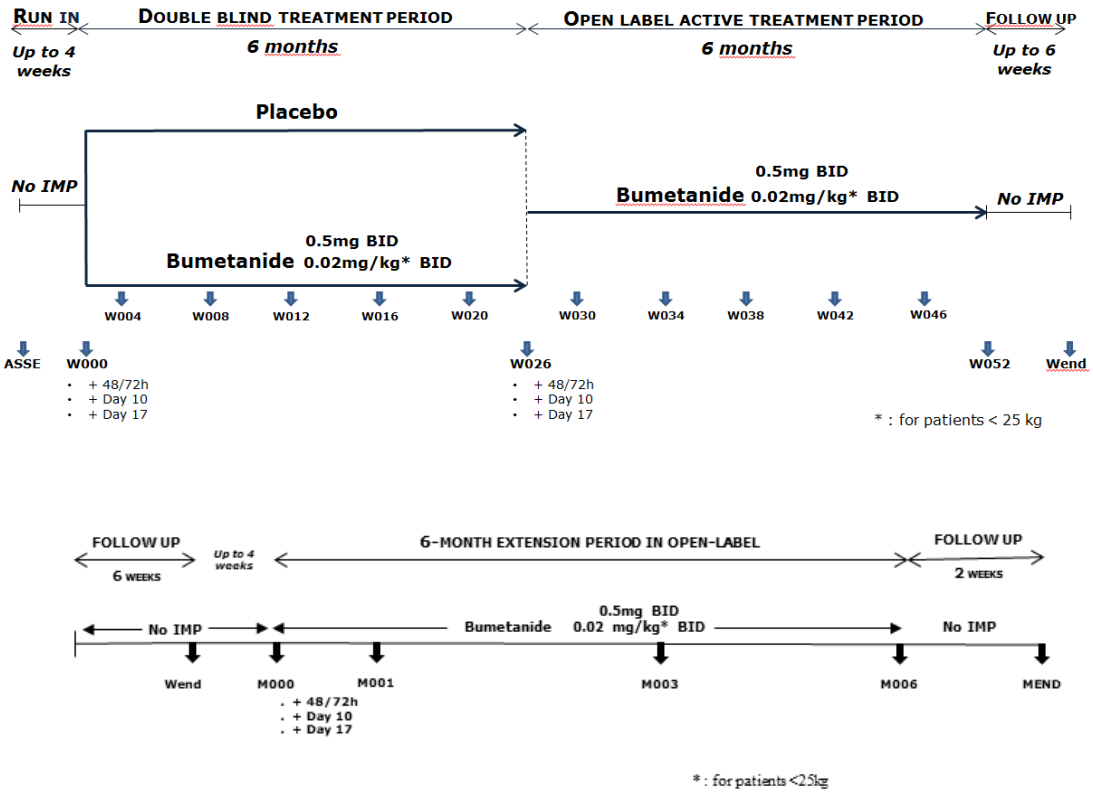
The volume of the oral solution will be adapted according to a body-weight basis for patients with a weight lower than 25 kg.

The appearance and taste of the solution are the same for all IMPs throughout the study. The first IMP intake takes place on the day of the inclusion visit in the afternoon.

An adaptation of the oral solution volume according to the weight of the patients will be performed during the study at visits W012, W026 and W038.

The study plan is shown in Figure (1.2.1) 1.

Figure (1.2.1) 1 - Study plan



1.2.2. Type of randomisation

The treatment randomisation and allocation are centralised using an Interactive Response System (IRS) procedure. The treatment (bumetanide or placebo) is assigned at inclusion visit by a balanced, non-adaptive randomisation with stratification on country and gender.

1.3. Determination of sample size

The sample size was estimated to meet the primary endpoint defined as the change in CARS2 from baseline to 6 months using a two-sided Student's test.
With a total of 170 patients, statistical significance should be established with a power of 90% and a type-one error of 0.05 for an effect size of 0.5 between groups.
Assuming around 15% prematurely withdrawal rate, the total number of patients to be included is 200.
Missing data will be imputed using a reference based multiple imputation with a jump-to-reference approach or a standard multiple imputation depending on the reasons of withdrawal. Simulations were carried out to evaluate the impact of this procedure on power under various scenarios. The results on 1000 simulations have shown that power does not drop under 80% in the worst-case scenario.

2. ANALYSIS SETS / TREATMENT GROUPS

2.1. Analysis sets

- The Randomised Set (RS) is constituted of included and randomised (*i.e.* for whom a therapeutic unit was randomly assigned using IWRS) patients.
- The Safety Set (SS) is constituted of all patients having taken at least one dose of IMP.
- Randomised Set Open (RSO) patients of the RS having taken at least one dose of Bumetanide during the open-label period.
- Safety Set Open (SSO): patients of the SS having taken at least one dose of Bumetanide during the open-label period.
- Randomised Set Combined (RSC): patients of the RS having taken at least one dose of Bumetanide during the open-label period and with a delay between end of double-blind period and start of open-label period less than 30 days.
- Safety Set Combined (SSC): patients of the SS having taken at least one dose of Bumetanide during the open-label period and with a delay between end of double-blind period and start of open-label period less than 30 days.

The size of each analysis set, and reasons for exclusion will be described.
Listings of patients with their membership, or not, to each analysis set and of excluded patients with reasons for exclusion will be provided.

2.2. Treatment arms

For the 26-week double-blind period, treatment arms will be defined as:
- S95008 arm.
- Placebo arm.

For the open-label period, treatment group will be defined as:
- Placebo/ S95008 arm: patients assigned to placebo group at W000 and treated by bumetanide 0.5 mg BID in the open-label period.

For the combined period (double-blind+open label periods), treatment group will be defined as:
- S95008 / S95008 arm: patients assigned to bumetanide 0.5 mg BID group at W000 and treated by bumetanide 0.5 mg BID in the open-label period.

It will correspond to allocated treatment by IRS except for safety analyses for which treatment received at W000 will be considered.

3. STATISTICAL METHODS

3.1. General considerations

3.1.1. Descriptive statistics

For **qualitative data**, number of observed values, number and percentage of patients per class will be presented. Unless otherwise specified in the TLG, no class "Missing" is considered.

For **quantitative data**, number of observed values, mean, standard deviation, median, first and third quartiles, minimum and maximum will be presented.

For **events**: number of patients having experienced the event (n), number of events that occurred and number of patients at risk for the event (N) will be presented as well as the patients-years at risk for the event (NPY), the crude incidence rate and the annual incidence rate calculated as follows:

- NPY = sum of patient's period duration in years.
- Crude incidence rate (*i.e.* percentage): % = 100 x (n / N).
- Annual incidence rate (*i.e.* exposure-adjusted incidence rate): npy = 100 x (n / NPY). Of note that the annual incidence rate is interpreted as the number of patients with events per 100 patients-years.

3.1.2. General definitions

Unless specified otherwise in Sections 3.2 to 3.6, the following definitions will be considered:

Three analysis periods will be defined:

- Double-blind period.
- Open-label period.
- Combined period.

For each, start date, stop date, first and last IMP intake date, baseline and under treatment period will be specified.

- Analysable value will be defined as any non-missing value.
- Baseline value will be defined as the last analysable value prior to the first IMP intake of the considered analysis period (*i.e.* before or the same* date as the first IMP intake date).
  Note:
  - In case of patient included and/or randomised but not treated (*i.e.* patients with treatment duration equal to 0): value at baseline (for the double-blind or combined period) is defined as the last analysable value prior or equal to date of inclusion visit.
  - * For parameters measured the day of inclusion for the double-blind or combined period (resp. the day of W26 or W26B for open-label period) but before the first IMP intake of the considered period.

- Post-baseline value will be defined as any value recorded at a given timepoint after baseline.
- Change from baseline will be defined as the arithmetic difference between a post-baseline value and the baseline value of the considered period.

3.2. Disposition and baseline characteristics

Disposition of patients and baseline characteristics will be described by randomised treatment group, to assess their comparability, and overall.

3.2.1. Disposition of patients

Disposition of patients, including reasons for study*/treatment withdrawal**, will be summarized during the double-blind period, during the open-label period, and during the study (excluding extension period) overall as well as during the double-blind period and during the open-label period by visit, in the RS. The number and percentage of patients who completed the double-blind (resp. open-label) period before (resp. after) the initiation of COVID-19 Pandemic will be also provided.

Note:
- * Study withdrawal related to COVID-19 pandemic will be also described.
- ** Information related to treatment withdrawal will only be described for the double-blind period.

A listing of patients with treatment withdrawal but having completed the double-blind period as well as a listing of patients with a delay between the double-blind and open-label periods will be provided.

The number and percentage of patients entering in the extension period will be described in the RS.

In order to assess the drop-out pattern between the treatment groups, the time to premature IMP withdrawal in the double-blind period will be described overall and by reason (lack of efficacy or AE; other reasons of withdrawal), in the RS using a Kaplan-Meier analysis. Withdrawn and completed patients' characteristics during the double-blind will be described overall, in the RS, and their comparability at inclusion visits assessed.

3.2.2. Protocol deviations

Protocol deviations before or at inclusion, as well as after inclusion, will be described by category of important deviations (based on ICH E3 guideline and ICH E3 Q&A) during the double-blind period, during the open-label period, and during the study (excluding extension period but including the deviations occurring between double-blind and open-label periods when there is a delay between them) in the RS.

The same analysis will be performed for protocol deviations after the initiation of COVID-19 Pandemic (*i.e. after start of COVID-19 impact*) (before or at inclusion, as well as after inclusion).

3.2.3. Demographic data and other baseline characteristics

Demographic data and other baseline characteristics such, vital signs, history and diagnosis of the disease will be described in the RS.

Baseline value of efficacy endpoints will also be described in the RS:
- CARS2: total raw score, version of the questionnaire used (ST or HF), severity of ASD.
- SRS-2 total raw score and core symptoms A and B.
- CGI-S score.
- Vineland adaptative behaviour scale: VABS II Adaptative Behavior Composite score, standard domains scores.

The Columbia-suicide severity rating scale children (C-SSRS-C) base during lifetime and during the past 6 months will be also described in the RS.

Moreover all specific (resp. non-specific) previous treatments taken within the 6 months before selection will be described in the RS, by ATC code; as well as all medical history other than studied disease and surgical or medical procedures history, by primary system organ class and preferred term. Previous therapies by ATC code for the disease will also be described in the RS.

3.3. Treatments of patients

3.3.1. Extent of exposure and treatment compliance

Extent of exposure (follow-up duration (month), treatment duration (month), treatment exposure (month)) and treatment compliance (%) will be described in the SS for double-blind period, in the SSO for open-label period and in the SSC for the combined period.

It is of note that treatment duration will also be described in classes:
- For double-blind period and open-label period: (<4 weeks, [4 weeks; 12 weeks], > 12 weeks).
- For combined period: (< 26 weeks, [26 weeks; 38 weeks], and > 38 weeks).

3.3.2. Concomitant treatments and therapies

All specific (resp. non-specific) concomitant treatments taken at inclusion, during the double-blind (resp. open-label, combined) period, at the first IMP intake of the open-label period and after the last IMP intake of the double-blind (resp. open-label) period will be described in the SS (SSO when open-label and SSC when combined), by ATC code. Therapies for autism spectrum disorder will also be described at inclusion and during the double-blind period in the SS and during the open-label in the SSO and during the combined period in the SSC. The number and percentage of patients with at least one therapy of ASD will be described overall and by type of therapy at inclusion in the SS. The number and percentage of patients with at least one new added therapy / one modification in current therapy / one clinically relevant modification in therapy will be described overall and by type of therapy in each period.

3.3.3. Acceptability and palatability questionnaires

The acceptability and palatability questionnaires will be presented using descriptive statistics for the visit at which it has been collected (W026 or before in case of premature withdrawal) in the RS by treatment group (S95008 VS placebo) for the double-blind period.

3.4. Efficacy analysis

All efficacy analyses will be performed on the randomised set.

3.4.1. Statistical hypotheses

The hypotheses to be tested are:
Let μ_0 and μ_1 be the population means of the change from baseline of the CARS2 at W026 (primary endpoint) under placebo and bumetanide respectively. The statistical hypotheses that will be tested are:

H0: $\mu_0 = \mu_1$ (no difference between bumetanide and placebo)

versus

H1: $\mu_0 \neq \mu_1$ (difference between bumetanide and placebo)

The type I error of the statistical tests will be set at 5% (bilateral situation), which is consistent with the objective of demonstrating the superiority versus placebo (unilateral situation at 2.5%).

Similar hypotheses will also be considered for the secondary efficacy parameters. However, for all other analyses than the primary analysis, no p-value will be provided.

3.4.2. Primary efficacy endpoint

3.4.2.1. Primary analysis

Definition:
The primary estimand is defined according to the primary objective of the trial, which is to evaluate treatment effect:

- Taking into account the unfavorable outcome when patients are unable to continue taking the study drug due to an adverse event or for lack of efficacy.
- Independently of treatment discontinuations for non-medical reasons because those patients would have theoretically continued to be treated as planned in clinical practice.

The attributes of the primary estimand will be defined as following:

- Population: children and adolescents aged from 2 to less than 7 years old with Autism Spectrum Disorder.
- Variable: change in CARS2 total score from baseline to W026.
- Summary measure: difference in means.
- Intercurrent events:
  - Treatment discontinuation due to lack of efficacy or AE (*hypothetical strategy)*:
    - For S95008 arm, the quantification of the treatment effect cannot ignore the situation where a patient can no longer tolerate or benefits from the treatment, from whom a continuation of the treatment would not be conceivable. Bumetanide is a treatment with a rapid onset and short duration of action. So, the assumption that the treatment benefit in patients who discontinue the active arm disappears immediately upon discontinuation is clinically meaningful. CARS2 values after this intercurrent event will be considered as missing and imputed using a reference based multiple imputation ("jump-to-reference" approach).
    - For placebo arm, it will be considered as if patients were stayed under placebo, and CARS2 values after the intercurrent event will be considered as missing and imputed using a multiple imputation (MAR approach).
    - Treatment discontinuation for other reason *(hypothetical strategy)*: a reasonable question is what difference is attributable to treatment if no such events occurred namely if patients were stayed under their randomised treatment. CARS2 values after the intercurrent events will be considered as missing and imputed in both arms using multiple imputations (MAR approach).

Primary endpoint: the primary efficacy endpoint is defined as the change from baseline to W026 of the CARS2 total score.

Primary analysis: bumetanide will be compared to placebo on the primary efficacy endpoint in the RS, using a General Linear Model (see Appendix 5.1.1.1) including the fixed, categorical effect of treatment, gender and country as well as the continuous fixed covariate of baseline value.

GENERAL LINEAR MODEL (ANCOVA): change = baseline country gender treatment [1]

The assumptions underlying the model, as for instance, the normality and homoscedasticity of residuals and detection of outliers, will be checked.

Missing data handling:

- For all study premature withdrawals due to other reasons as well as for study premature withdrawals due to lack of efficacy or AE in placebo arm, missing data will be imputed in both arms using a multiple imputation based on similar patients in the same treatment arm (missing at random assumption).
- For study premature withdrawals due to lack of efficacy or AE in S95008 arm, missing data will be imputed using a reference based multiple imputation with a jump-to-reference approach (*missing not at random assumption*).
  Missing data as well as data considered as missing for the primary analysis due to the strategy used to handle intercurrent event will be imputed in the same way.

Multiple imputation inference involves 3 consecutive phases.

1/ Imputation step:

A total of 100 imputed complete data sets will be generated.

The imputation step of missing data with a monotone pattern will be separated in several sequences:

- For all patients with study or drug premature withdrawals due to other reasons as well as for patients with study or drug premature withdrawals due to lack of efficacy or AE in placebo arm, the regression method (adjusted on the same covariates as the model [1] and considering all longitudinal data of planned visits) will be used to impute missing data (*i.e.* based on patient completers under corresponding treatment arm).
- For patients with study or drug premature withdrawals due to lack of efficacy or AE in S95008 arm, the imputation step will be done sequentially for each time point t separately (W4, W12 and W26) based on placebo patients with non missing data at the timepoint t. The regression method will be adjusted on the same covariates as model [1] except the treatment variable (which is a modified variable in order to consider the placebo data as reference for imputation) and considering all longitudinal data of planned visits up to timepoint t.

It is of note that for missing data with an arbitrary missing pattern, the first imputation step might be preceded by one MI approach based on MCMC method using treatment groups, baseline and variable of interest at the planned visit.

2/ Analysis step:

The same model as [1] (described above) will be applied to each of the 100 imputed datasets.

3/ Combination step:

Statistical inferences will be generated by combining results from the 100 analyses using Rubin's formulae. The multiple imputation estimator of the difference between bumetanide and placebo is the average of the individual 100 estimators. The variance of the estimator is the combination of the between- and within-imputation variability (Carpenter and Kenward, 2007).

Statistical elements: Finally, the following elements will be provided in a summary table:

- Estimate (standard error) of the difference between adjusted treatment arm means.
- Two-sided 95% CI of the estimate.
- Two-sided p-value.

3.4.2.2. Sensitivity analyses

- Unadjusted analysis:

The same primary analysis will be performed for the primary endpoint using a General Linear Model (ANCOVA) without any covariates. The following model will be considered instead of model [1]:

GENERAL LINEAR MODEL (ANCOVA): change = baseline treatment [2]

3.4.2.3. Supplementary analyses

- Treatment policy estimand:

The primary analysis will be repeated for the primary endpoint using all CARS2 values reported regardless of occurrence of treatment discontinuation.
In case of CARS2 values are still missing, the same imputation process as in the primary analysis will be performed for these individuals.

- Hypothetical estimand based on Mixed-effects Model for Repeated Measures

A Mixed-effects Model for Repeated Measures (MMRM) (see Appendix 5.1.3) using the longitudinal data at each planned post-baseline visit of the double-blind period (W004, W012 and W026) will be performed as if patients had continued their randomised treatment. In this analysis, all data occurring after all intercurrent events will be considered as missing. The comparison associated with the analysis will be the contrasts between Bumetanide and placebo for the change from baseline to W026 of the CARS2 total score. The model will include the fixed, categorical effects of country, gender, treatment, visit, treatment by visit interaction, country by visit interaction and gender by visit interaction as well as the continuous, fixed covariates of baseline and visit-by-baseline interaction:

*MMRM: change at each visit = baseline country gender treatment visit treatment*visit baseline*visit country*visit gender*visit [4]*

Statistical elements: the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment arm means.
- Two-sided 95% CI of the estimate.

- W012 analysis:

The primary analysis will be repeated for the CARS2 total score expressed in term of change from baseline to W012.

- CARS2 responders analysis:

Definition:
CARS2 responders will be defined as patients having an improvement of at least 4.5 points in the CARS2 total raw score as compared to baseline (Jurek *et al.*, 2021).

Main Analysis:
CARS2 responders at W012 and at W026 will be analysed using a logistic regression model with the fixed categorical effect of treatment, country and gender.
Note: In case of a complete separation of data (zero cell count in a treatment group), only descriptive statistics will be provided. In case of quasi-completed separation of data, a logistic regression with penalized maximum likelihood will be used (see Appendix 5.1.4).

Missing data handling:
The 100 datasets imputed for the primary analysis and supplementary analysis at W012 (based on the changes from baseline in CARS2 total scores) will be used to derive, the CARS2 responders. Analysis step will be based on the logistic regression model described above and then the combination step will be applied.

Statistical elements: the following elements will be provided in a summary table:
- Estimate (standard error) of the odds ratio between treatment groups.
- Two-sided 95% CI of the estimate.

- Descriptive analysis:

The CARS2 total score and subscores will be summarized using descriptive statistics and graphs (expect for subscores):
- During the double-blind period: baseline, each planned post-baseline visit (W004, W012 and W026) as well as change from baseline to each planned post-baseline visit will be described in the RS by treatment groups (S95008 *vs* placebo).
- During the open-label period: OL baseline, each OL planned post-baseline visit (W038 and W052) as well as change from OL baseline to each OL planned post-baseline visit will be described in the RSO for the Placebo/S95008 arm.
- During the combined period: baseline, each planned post-baseline visit as well as change from baseline to each planned post-baseline visit will be described in the RSC for the S95008/S95008 arm.
- At the follow-up visit (WEND), the change from the last post-baseline visit under treatment during the open-label period to WEND will be described in the RSO by treatment groups (S95008/S95008 *vs* Placebo/S95008).

The proportions of CARS2 responders will also be summarized at each planned post-baseline visit for each period.
The proportions of CARS2 performed remotely at baseline, W012 and W026 will be described in the RS by treatment groups (S95008 *vs* placebo).

3.4.3. Secondary efficacy endpoints

3.4.3.1. Social Responsiveness Scale

Definition: The main analytical approach will be defined as the change in SRS-2 total raw score from baseline to W026.

Main Analysis: A General Linear Model with the fixed categorical effect of treatment, country and gender as well as the continuous, fixed covariate of baseline will be used for the SRS-2 change from baseline in the RS.

GENERAL LINEAR MODEL (ANCOVA): change = baseline country gender treatment

Missing data handling: For missing data handling, the same approach as for the primary endpoint will be performed. Data considered as missing for the analysis due to the strategy used to handle intercurrent event will be imputed in the same way.

Statistical elements: the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.

Supplementary analyses:

- W012 analysis:

The same analysis will be repeated for the SRS total score expressed in term of change from baseline to W012.

- Descriptive analysis:

The same approaches as for the primary endpoint will be performed for the SRS total score.

3.4.3.2. Clinical Global Impression Scale

Definition:

- The main analytical approach will be defined as the CGI-I score at W026.
- Responders CGI will be defined as patients with the CGI-I equal to "much improved" or "very much improved".

Main analysis: A Robust General Linear Model (RGLM) using a rank-based analysis (Wilcoxon scores) (see Appendix 5.1.1.2) with the fixed categorical effect of treatment, country and gender will be used for the CGI-I score at W026 in the RS.

RANK-BASED GLM: CGI-I score = country gender treatment

Missing data handling: for missing data handling, the same approach as for the primary endpoint will be performed (see Appendix 5.1.2).

Statistical elements: the following elements will be provided in a summary table:

- Estimate (standard error) of the difference between adjusted treatment arms medians.
- Two-sided 95% CI of the estimate.

Supplementary analyses:

- W012 analysis:

The RGLM will be repeated for the CGI-I at W012.

- CGI responders analysis:

CGI responders at W012 and W026 will be analysed using a logistic regression model with the fixed categorical effect of treatment, country and gender in the RS.
Note: In case of a complete separation of data (zero cell count in a treatment group), only descriptive statistics will be provided. In case of quasi-completed separation of data, a logistic regression with penalized maximum likelihood will be used (see Appendix 5.1.4).

Statistical elements: the following elements will be provided in a summary table:

- Estimate (standard error) of the odds ratio between treatment groups.
- Two-sided 95% CI of the estimate.

- Descriptive analysis:

The CGI-I, CGI-S and CGI responders will be summarized using descriptive statistics and graphs:

- During the double-blind period: number and percentage of patient at each planned visit (for CGI-I and CGI responders: W000+Day 17, W004, W008, W012, W016, W020 and W026 / for CGI-S: W012 and W026) and the value at each planned visit for the CGI-S will be described in the RS by treatment groups (S95008 VS placebo).

- During the open-label period: number and percentage of patient at each planned visit (W038 (only for CGI-I and CGI responders) and W052) and the value at each planned visit for the CGI-S will be described in the RSO for the Placebo/S95008 arm. Of note that only CGI-I compared to W026 will be considered.
- During the combined period: number and percentage of patient at each planned visit and the value at each planned visit for the CGI-S will be described in the RSC for the S95008/S95008 arm. Of note that only CGI-I compared to W000 will be considered.

Moreover, the CGI-I (compared to W052) will be described at the follow-up visit (WEND) during the open-label period in the RSO by treatment groups (S95008/S95008 *vs* Placebo/S95008).

The number and percentage of patients at each planned visit in the extension period for the CGI-I (of note that CGI-I compared to the inclusion in the extension period will be considered) will be described in the RS patients included in the extension period.

3.4.3.3. Vineland Adaptative Behaviour Scale II (VABS II)

Definition: The main analytical approach will be defined as the change in standard score of each domain from baseline to W026.

Main Analysis: A General Linear Model with the fixed categorical effect of treatment, country and gender as well as the continuous, fixed covariate of baseline will be used for the change in standard score of each domain in the RS.

GENERAL LINEAR MODEL (ANCOVA): change = baseline country gender treatment

Missing data handling: for missing data handling, the same approach as for the primary endpoint will be performed. Of note that there is only one post-baseline planned visit (W026) to consider.

Statistical elements: The following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment arm means.
- Two-sided 95% CI of the estimate.

Supplementary analysis:
VABS II standard domains scores, and Adaptative Behavior Composite standard score will be summarized using descriptive statistics and graphs:
- During the double-blind period: baseline, W026 as well as change from baseline to W026 will be described in the RS by treatment groups (S95008 *vs* placebo).
- During the open-label period: OL baseline, W052 as well as change from OL baseline to W052 will be described in the RSO for the Placebo/S95008 arm.
- During the combined period: baseline, each planned post-baseline visit as well as change from baseline to each planned post-baseline visit will be described in the RSC for the S95008/S95008 arm.

3.4.3.4. Paediatric Quality of Life Inventory

Definition: the four dimensions (Physical, Emotional, Social and School), the physical* and psychosocial health summary scores, and the total score will be calculated.

*Note: the physical health summary score is in fact the same as the Physical domain score. Only the latest will be presented, with a note in the output mentioning that it is the same.

The PedsQL questionnaire has been already validated in children with chronic conditions (Varni *et al*., 2003) and pediatric population with psychiatric disorders (Limbers *et al*., 2011). The psychometric properties will be checked separately.

Analysis:
A General Linear Model with the fixed categorical effect of treatment, country and gender as well as the continuous, fixed covariate of baseline will be used for the change from baseline to W026 in the four dimensions, the physical and psychosocial health summary scores and the total score in the RS.

GENERAL LINEAR MODEL (ANCOVA): change = baseline country gender treatment

Missing data handling: missing data will be imputed using the multiple imputation approach (MI).

Statistical elements: the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.

Supplementary analysis:
The four dimensions, the physical and psychosocial health summary scores and the total score will be summarized using descriptive statistics and graphs:
- During the double-blind period: baseline, each planned post-baseline visit (W004, W012 and W026) as well as change from baseline to each planned post-baseline visit will be described in the RS by treatment groups (S95008 *vs* placebo).
- During the open-label period: OL baseline, each OL planned post-baseline visit as well as change from OL baseline to each OL planned post-baseline visit (W030, W038 and W052) will be described in the RSO for the Placebo/S95008 arm.
- During the combined period: baseline, each planned post-baseline visit as well as change from baseline to each planned post-baseline visit will be described in the RSC for the S95008/S95008 arm.

3.4.3.5. Abbreviated World Health Organization Quality of Life - Bref

Definition: the four domain scores (Physical, Psychological, Social Relations and Environment) will be calculated.
The WHOQOL-Bref questionnaire has been validated in parents of children with autistic disorder (Latefa *et al*., 2014). The psychometric properties will be checked separately.

Analysis: A General Linear Model with the fixed categorical effect of treatment, country and gender as well as the continuous, fixed covariate of baseline will be used for the change from baseline to W026 in the four domains scores in the RS.

GENERAL LINEAR MODEL (ANCOVA): change = baseline country gender treatment

Missing data handling: missing data will be imputed using the multiple imputation approach (MI).

Statistical elements: the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.

Supplementary analysis:
The four domains scores will be summarized using descriptive statistics and graphs:
- During the double-blind period: baseline, each planned post-baseline visit (W004, W008, W012 and W026) as well as change from baseline to each planned post-baseline visit will be described in the RS by treatment groups (S95008 *vs* placebo).
- During the open-label period: OL baseline, each OL planned post-baseline visit (W030, W034, W038 and W052) as well as change from OL planned baseline to each OL post-baseline visit will be described in the RSO for the Placebo/S95008 arm.
- During the combined period: baseline, each planned post-baseline visit as well as change from baseline to each planned post-baseline visit will be described in the RSC for the S95008/S95008 arm.

3.4.4. Subgroups analysis

Not applicable.

3.5. Exploratory analysis

3.5.1. EuroQol – 5-Dimension – Three-level

Definition: The EuroQoL-5D-3L questionnaire (EQ-5D) is composed of 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), and a Visual Analog Scale (EQ-5D VAS).
The dimensional 5-level system will be converted into a single index utility score (EQ-5D index score).
Psychometric properties of the EQ-5D-3L Questionnaire were studying in caregivers of autistic children (Khanna *et al.*, 2013).
The psychometric properties will be checked separately.

Analysis:
A General Linear Model with the fixed categorical effect of treatment, country and gender as well as the continuous, fixed covariate of baseline will be used for the change from baseline to W026 in the EQ-5D index score (resp. EQ-5D VAS) in the RS.

GENERAL LINEAR MODEL: change = baseline country gender treatment

Missing data handling: missing data will be imputed using the multiple imputation approach (MI).

Statistical elements: the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.

Supplementary analysis:
The EQ-5D VAS and the EQ-5D index score will be summarized using descriptive statistics and graphs:
- During the double-blind period: baseline and change from baseline to each post-baseline visit will be described in the RS by treatment groups (S95008 *vs* placebo).
- During the open-label period: OL baseline and change from OL baseline to each OL post-baseline visit will be described in the RSO for the Placebo/S95008 arm for patients.
- During the combined period: baseline and change from baseline to each post-baseline visit will be described in the RSC for the S95008/S95008 arm.

3.6. Safety analysis

All safety analyses will be performed on the taken treatment at W000 as following:
- For the double-blind period in the SS by treatment group (S95008 *vs* placebo).
- For the open-label period: in the SSO on the Placebo/S95008 arm.
- For the combined period: in the SSC on the S95008/S95008 arm.

Six periods are of interest to present accurately the safety endpoints. The start and end date of these periods are summarized in the following table:

Table (3.6) 1 - Definition of periods of emergence for safety events

Period ID	ANALYSIS PERIOD	START	END
#1	Emergent during the double-blind period	First intake in double-blind	End of double-blind period (including W26)
#2	Treatment Emergent during the double-blind period	First intake in double-blind	Min(Last intake in double-blind + 2 days, First intake in open-label)
#3	Emergent during the open-label period	First intake in open-label (W26 not included)	End of open-label period (including W52 and follow-up)
#4	Treatment Emergent during the open-label period	First intake in open-label (W26 not included)	Last intake in open-label + 2 days (included)
#5	Emergent from the combined period	First intake in double-blind	End of open-label period (including W52 and follow-up)
#6	Treatment Emergent from the combined period	First intake in double-blind	Last intake in open-label + 2 days (included)

Listing of main safety data (including adverse events) and start/end date of each period will be provided for:
- Patients who had a delay superior to 30 days between the double-blind and the open-label period will be listed separately.

3.6.1. Adverse events

Definition:

- **Emergent Adverse Events** (EAE) during the double-blind *(resp. open-label, combined)* period are defined as all adverse events which occur in the period #1 *(resp. #3, #5)* (as per table above), or which started strictly before period #1 *(resp. #3, #5)* but which worsen (in terms of intensity) or become serious according to the investigator opinion during period #1 *(resp. #3, #5)*.
- **Treatment Emergent Adverse Events** (TEAE) during the double-blind (resp. open-label, combined) period are defined as all adverse events which occur in the period #2 (resp. #4, #6) (as per table above), or which started strictly before period #2 (resp. #4, #6) but which worsen (in terms of intensity) or become serious according to the investigator opinion during period #2 (resp. #4, #6).
- **AE starting during the run-in period** are defined as all adverse events which started strictly before period #1.

Analyses:

Number of events as well as number, percentage and annual incidence rate of patients reporting at least one event, presented by primary System Organ Class (SOC), and Preferred Term (PT), will be provided for:

- Serious EAE and deaths for period #1, #3 and #5.
- TEAE, TEAE leading to IMP withdrawal, TEAE requiring new treatment or increase of on-going treatment, TEAE requiring surgical or medical procedure, TEAE related to IMP, serious TEAE, severe TEAE, non-serious TEAE for periods #2, #4 and #6.

TEAE will be described according to the seriousness, the intensity, the relationship with the IMP, the action taken regarding the IMP, the requirement of added therapy and the outcome.

Moreover, TEAE, TEAE related to IMP and serious TEAE will be presented according to their time to onset in weeks calculated from the start of the period #2 (resp. #4 and #6).

AE starting during the run-in period will be also described by SOC and PT in the SS.
All adverse events reported after the last Bumetanide intake + 2 days (excluded) will be listed.
All COVID-19 infection reported during the study will be listed.

Number of events as well as number, percentage of patients reporting at least one event, presented by primary System Organ Class (SOC), and Preferred Term (PT), will be provided for all AE and all serious AE occurring after the first intake in the extension part in the patients included in the extension period.

3.6.2. Columbia Suicide Severity Rating Scale Children's version

Definition:

The C-SSRS-C data will be mapped to Columbia Classification Algorithm for Suicide Assessment (C-CASA) (Guidance for Industry, Suicidality, FDA CDER, September 2010) following C-SSRS Scoring and Data Analysis Guide (February 2013) (C-SSRS Guide). The following endpoints will be thus considered:

11 preferred C-SSRS categories:

- For Suicidal ideation:
  - Wish to be dead (category 1).
  - Non-specific active suicidal thoughts (category 2).
  - Active suicidal ideation with any methods (not plan) without intent to act (category 3).

  - Active suicidal ideation with some intent to act, without specific plan (category 4).
  - Active suicidal ideation with specific plan and intent (category 5).
- For Suicidal behaviour:
  - Preparatory acts of behaviour (category 6).
  - Aborted attempt (category 7).
  - Interrupted attempt (category 8).
  - Actual attempt (non-fatal) (category 9).
  - Completed suicide (category 10).
- Self-injurious behaviour without suicidal intent (category 11).

Endpoints derived from these C-SSRS categories:
- Suicidal ideation or behaviour: at least one "yes" answer present at any assessable visit *(i.e. categories 1 to 10 analysable)* during the considered period to any one of the 1 to 10 categories.
- Suicidal ideation: at least one "yes" answer present at any assessable visit during the considered period to any one of the 1 to 5 categories.
- Suicidal behaviour: at least one "yes" answer present at any assessable visit during the considered period to any one of the 6 to 10 categories.
- Suicidal ideation score: maximum suicidal ideation category (1-5 on the CSSRS) among any assessable visit of the considered period; 0 if no ideation is present.

Baseline:
For suicidal ideation and suicidal behaviour (for items and the four associated derived C-SSRS endpoints), baseline will be defined as the value at W000, lifetime.

Analyses:
The 11 preferred C-SSRS categories as well as the four associated derived C-SSRS endpoints will be described during the double-blind (resp. open-label, combined) treatment period (*i.e.* periods #2, #4 and #6). Moreover, a listing of all patients having at least one post-baseline category equal to "yes" will be performed.

3.6.3. Clinical laboratory evaluation

Definition:
- A laboratory value is considered as analysable if non-missing and not flagged in the ClinTrial database as "not analysable".
- The analyses will be performed on periods #2, #4 and #6.
- A value is considered as occurring during the considered period if the sampling date is between the start and last date of the considered period.
- Baseline for each period is defined in Section 3.1.2.
- For each considered period:
  - High emergent abnormal value under treatment according to the laboratory reference ranges is defined as value ≤ ULN (Upper Limit of Normal laboratory reference range) or missing at the baseline of the considered period and > ULN under treatment.
  - Low emergent abnormal value under treatment according to the laboratory reference ranges is defined as value ≥ LLN (Lower Limit of Normal laboratory reference range) or missing at baseline of the considered period and < LLN under treatment.

- High emergent abnormal value under treatment according to the cut-offs for PCSA values is defined as value ≤ ULS (Upper Limit used to define potentially clinically Significant abnormal values) or missing at baseline of the considered period and > ULS under treatment.
- Low emergent abnormal value under treatment according to the cut-offs for PCSA values is defined as value ≥ LLS (Lower Limit used to define potentially clinically Significant abnormal values) or missing at baseline of the considered period and < LLS under treatment.

Analyses:
For each biochemical, haematological (blood and urine sample) parameter and for electrolytes monitoring (Potassium, Sodium), the following analysis will be performed for each period:
- Number and percentage of patients with at least one high/low emergent abnormal value under treatment, according to the laboratory reference ranges and to the cut-offs for PCSA values.
- Laboratory parameters classified (number and percentage of patients in each class) according to these reference ranges and cut-offs and using shift tables from baseline in the considered period to the worst (high and/or low) values under treatment.

Moreover, listings of patients with out-of-range or PCSA analysable values emergent under treatment in each period and of non-analysable values excluded from analyses will be provided.

Additionally, for potassium parameter, number and percentage of patients with hypokalaemia will be described by severity (mild: 3mEq/L ≤ K^+< 3.5 mEq/L / moderate: 2.2mEq/L ≤ K^+< 3 mEq/L / severe: K^+< 2.2mEq/L) in each period.

3.6.4. Vital signs, clinical examination and other observations related to safety

3.6.4.1. Vital signs and clinical examination

Definition:
- A value is considered as analysable if collected under appropriate circumstances.
- The analyses will be performed separately for periods #2, #4 and #6.
- A value is considered as occurring during the considered period if the associated visit date is between the start and last date of the considered period.
- Baseline for each period is defined in Section 3.1.2.

Analyses
The following vital signs and clinical examination will be analysed:
- Weight (kg).
- BMI (kg/m²) (quantitative value and by class).
- HR (b.p.m).
- SBP (mmHg).
- DBP (mmHg).

Quantitative parameters will be described separately for each period, in terms of value at baseline, value at each planned post-baseline visit under treatment as well as in terms of change from baseline to each planned post baseline visit under treatment.

3.6.4.2. Electrocardiogram

Definition:

- ECG parameters will be the following: heart rate (b.p.m), RR interval (msec), PR interval (msec), QRS complex (msec), QT and QTcF interval (msec) (Fridericia's correction formula interval).
- A value is considered as analysable if non-missing value obtained from an interpretable ECG.
- An ECG is considered as interpretable if the quality of the trace is "Correct", with "Minor problems" or "Missing data".
- The analyses will be performed separately for periods #2, #4 and #6.
- A value is considered as occurring during the considered period if the assessment date is between the start and last date of the considered period.
- Baseline for each period is defined in Section 3.1.2.

Analyses:

Presence of clinically significant ECG abnormalities will be described as well as sinus rhythm at each planned visit under treatment.

ECG parameters will be described separately for each period, in terms of value at baseline, value at each planned post-baseline visit under treatment; as well as, for quantitative endpoints, in terms of change from baseline to each planned post baseline visit under treatment.

Moreover values and changes form baseline of the concerned period of corrected QT interval will be described in classes (*i.e.*, ≤ 450,]450 ; 480[, [480 ; 500] and > 500 msec for values, and ≤ 30,]30 ; 60] and > 60 msec for changes).

4. INTERIM ANALYSIS

Not applicable.

5. APPENDICES

5.1. Statistical methods details

5.1.1. General Linear Model

5.1.1.1. For continuous data

A model that will be used for this study is the General Linear Model studying treatment effect with baseline gender and country (as fixed effect) as covariates:

$$Y_{ijkt} = \gamma X_{ijkt} + a_i + b_j + d_k + \varepsilon_{ijkt}$$

Where:

- Y_{ijkt} is the response from the t^{th} patient, of gender k, in country j, that received treatment i.
- X_{ijkt} is the baseline value of the analysed variable for the t^{th} patient, of gender k, in country j, that received treatment i.
- γ denotes the common slope of the baseline covariate.
- a_i denotes the intercept of the i^{th} treatment.
- b_j denotes the fixed effect of country j.
- d_k denotes the fixed effect of the gender k.
- $\varepsilon_{ijkt} \sim$ iid $N(0, \sigma_\varepsilon^2)$ denotes the experimental unit error associated with the t^{th} patient of gender k, in country j, that received treatment i.

The model can also be written in matrix notation as:

$$Y = X\beta + \varepsilon$$

Where:

- Y is the vector of observations.
- X is the design matrix of the fixed effects and baseline factors.
- β is the unknown vector of the fixed effects and baseline factors.
- ε is the unobserved vector of independent and identically distributed random errors, such as $\varepsilon \sim N(0,R)$ where $R = \sigma_\varepsilon^2 I_n$.

Estimate:

The estimate $\hat{\beta}$ of β is given by:

$$\hat{\beta} = (X'X)^{-} X'Y$$

with $Var(\hat{\beta}) = (X'\hat{R}^{-1}X)^{-} = \hat{C}$

Confidence interval and p-value:

The Sum of Squares using a Least-Squares-means contrast is used for computing the Sum of Squares under H0: $L\beta = 0$ (where $L\beta$ denotes a linear estimable combination of the fixed effects).

The 95% confidence interval of $\hat{\beta}$ is given by:

$$L\hat{\beta} \pm t_{\nu,0.975} \sqrt{L\hat{C}L'}$$

Where:

- $\hat{\nu}$ denotes the approximate degrees of freedom by Kenward Roger.
- $t_{\hat{\nu},0.975}$ is the 0.975^{th} quantile of the t-distribution with $\hat{\nu}$ degrees of freedom.

The Least-Squares approach provides estimates of the linear parameters that are unbiased and have minimum variance among linear estimators.
The treatment effect is estimated by the difference of the Least-Squares-means of each studied treatment group.

Validation of hypothesis:
For the validation of the model used, the following points will be studied:
a) Normality and homoscedasticity of residuals
The assumptions of normality and homoscedasticity of residuals will be investigated on the chosen final model using some graphs and descriptive statistics.

b) Detection of outliers
The detection will be done using graphics.

5.1.1.2. Robust General Linear Model using rank-based analysis

The rank-based analysis is analogous to the least-squares analysis except that the Euclidean distance to minimize is replaced by a different measure of distance that is based on the dispersion function of Jaeckel (ref. Rfit: Rank-based Estimation for Linear Models):

$$\hat{\beta} = \text{Argmin}||Y\text{-}X\beta||\varphi$$

With $\|.\|\varphi$ the following measure of distance, R being the rank and φ the Wilcoxon score function:

$$\left|\left|u\right|\right|_{\varphi} = \sum_{i=1}^{n} \varphi\left(\frac{R(u_i)}{n+1}\right) u_i$$

As such analysis cannot be implemented directly from SAS, the function *rfit* of the package Rfit of R will be used. Of note that only the phase 2 analysis set (see Section 5.1.2) will be performed in R.

5.1.2. Pattern imputation based

We will implement a hybrid approach using control-based pattern imputation and standard multiple imputation depending on the treatment group and the reasons of withdrawal.

Standard multiple imputation (Rubin, 1987) replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. Each value is a Bayesian drawn from the conditional distribution of the missing observation given the observed data, made in such a way that the set of imputation properly represents the information about the missing value that is contained in the observed data for the chosen model.

The multiply imputed data sets are then analysed by using standard procedures for complete data and combining the results from these analyses, based on the change from baseline.
Multiple imputation inference involves 3 consecutive phases (phase 1: imputation step, phase 2: analysis step and phase 3: combination step).

1/ PHASE 1: Imputation step

To guarantee the reproduction of results, the seed will be fixed at 95008.
This part will be divided in 3 steps summarized hereafter:

	Dataset	
Starting point	*init* *(N incomplete raws)*	Intitial incomplete dataset
Step1:	*DATA_BASIS* *(N*m incomplete raws)*	Impute intermittent missing data (arbitrary) - MI approach based on MCMC method
Step2a:		Impute data for all placebo patients and S95008 patients with a WD due to other reasons - classical MI imputation, with treatment, baseline, covariates, and time points of interest in the regression model
Step2b:		Impute data for S95008 patients with a WD due to AE or lack of efficacy with a jump-to-reference approach, with baseline, covariates and time points of interest in the regression model.
Step 3:	*DATA_COMPLETE* *(N*m complete raws)*	Assemble back all data together

Step 1:

First, data collected post treatment withdrawal (if any) should be considered as missing. All intermittent missing data (*i.e.* with an arbitrary missing pattern) of baseline and variable of interest at the planned visits will be first imputed using a MI approach based on MCMC method (with single chain) by treatment groups, in case the initial dataset has not a strict monotone missing pattern.

A dataset is said to have a monotone missing pattern if missing values always occur at the end of data records; if we note $Y_i=(Y_{i,0},Y_{i,1},\ldots,Y_{i,t},\ldots Y_{i,T})$ the measured variable at each time t from baseline (t=0) to the end of the study (t=T) for an individual i, the missing pattern is monotone if, for all individuals and all times, if $Y_{i,t}$ is missing, then $Y_{i,t+1}$ is also missing and if $Y_{i,t}$ is not missing, then $Y_{i,t-1}$ is not missing also.

A total of 100 imputed partially completed datasets will be generated.

In case of patients with no data available at baseline and at any post baseline visit, a regression method will be used on the 100 imputed partially completed datasets to perform the imputation at the baseline visit only, using in this order: baseline covariates (*e.g.* country, gender), treatment in the model.

This dataset will serve as a basis for Step 2a and Step 2b (*DATA_BASIS*).

For the following step, the entry dataset will be sorted by imputation, by treatment and by patient.

Step 2:

In this sequence, remaining missing data of *DATA_BASIS* will be imputed. The treatment group and the reason for study or drug premature withdrawals have to be considered as the imputation rules will be different. For this imputation step, countries with less than 10 patients will be gathered to avoid strata with no observed values.

a) For all placebo patients and patients in S95008 arm with study or drug premature withdrawals due to other reasons:

To impute those missing data, all data will be considered (regardless of the treatment and of the reason of withdrawal) in order to benefit of all data available at each visit. However, at the end, only data (imputed or initially complete) from all placebo patients and data from S95008 patients, completers or with initially missing data due to other reasons will be considered (the other being imputed with jump-to-placebo method, as explained in step 2b).
A regression method will be used to perform the imputation, using in this order: baseline, baseline covariates (e.g. country, gender), treatment and variable of interest at the planned visits in the model. Of note, for imputing data of visit i, data up to visit i-1 are to be considered for the imputation, even if the data of visit i-1 already results of an imputation (this is taken into account in PROC MI with the order of the variables put in the VAR statement). This is applicable for variable of interest at each planned post-baseline visits.
Of note, for subgroup analysis, the subgroup should not be added as a covariate in the regression model.

b) For patients with study or drug premature withdrawals due to lack of efficacy or AE in S95008 arm

To impute those missing data, an input dataset will be prepared with data from all placebo patients (completers or with missing values regardless of the reason), and from S95008 patients, with missing data due to AE/lack of efficacy.

The jump-to-placebo approach will be set up for these patients' imputations, by using the %mistep macro from James Roger (https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#dia-missing-data , zip file MISTEP20180327).

A sequence of %mistep will be run, considering as the first model (for the first post-baseline visit): baseline baseline covariates (*e.g.* country, gender), treatment*. The subsequent model (for the second post-baseline visit) will include, in addition to variables included in the first model, the residuals of this first model (put at the end of the statement). The other subsequent models will be implemented in the same way.

**at each visit the treatment variable is derived such as the missing value in the S95008 will be imputed using the placebo arm as reference.*

This macro will be used because no SAS procedure are currently available to perform the jump to reference approach.

Step 3:
Finally we can assemble back all subjects with complete data: all patients from Step2a and remaining S95008 patients from Step2b (*DATA_COMPLETE*).

In case of imputed data at W026 (respectively W012 depending on the analysis) are outside of the endpoint ranges, the imputed data will be replaced by the boundary that was overpassed.

Endpoint	Min	Max
CARS2 total score	15	60
SRS2 total score	0	195
CGI-I	1	7
VABS II Adaptative Behavior Composite standard score	20	160
VABS II Communication domain standard score	20	160
VABS II Daily Living Skills Domain standard score	20	160
VABS II Socialisation Domain standard score	20	160
PedsQL total score	0	2300
PedsQl Physical functioning score	0	800
PedsQl Emotional functioning score	0	500
PedsQl Social functioning score	0	500
PedsQl School functioning score	0	500
PedsQL Psychosocial Health Summary score	0	1500
WHOQOL-Bref Physical Domain Score	0	100
WHOQOL-Bref Psychological Domain Score	0	100
WHOQOL-Bref Social Relations Domain Score	0	100
WHOQOL-Bref Environment Domain Score	0	100
EQ-5D-3L index score	0	1
EQ-5D-3L VAS	0	100

2/ PHASE 2: Analysis step

The planned model will be applied to each of the 100 imputed datasets obtained at the previous step (*DATA_COMPLETE*). For each imputed dataset i (i=1,2…,m), the estimate of the difference between groups and the associated standard error will be stored.

Of note that for CGI-I analysis, before this analysis step, the imputed data will be rounded to the nearest unit. Moreover, this analysis step will be done on the change from baseline (whereas imputations were done on visit's value).

3/ PHASE 3: Combination step

Statistical inferences will be generated by combining results from the 100 analyses using Rubin's formulae. The multiple imputation estimator of the difference between Bumetanide and placebo is the average of the individual 100 estimators. The variance of the estimator is the combination of the between- and within-imputation variability (Carpenter and Kenward, 2007).

	Mean	Variance
Estimate	Mean($\hat{O}_i$)	Variance($\hat{O}_i$)
Standard error	Mean($\hat{U}_i$)	Variance($\hat{U}_i$)

Let $\hat{O}_i$ and $\hat{U}_i$ be the point and variance estimates from the ith imputed data set, i=1,2…,m. Then the point estimates for O from multiple imputations is the average of the m complete-data estimates:

$$\hat{O} = \frac{1}{m}\sum_{i=1}^{m}\hat{O}_i$$

Let $\hat{U}$ be the within-imputation variance, which is the average of the m complete-data estimates:

$$\hat{U} = \frac{1}{m}\sum_{i=1}^{m}\hat{U}_i$$

and B be the between-imputation variance:

$$B = \frac{1}{m-1}\sum_{i=1}^{m}(\hat{O}_i - \hat{O})^2$$

Then the variance estimate associated with $\hat{0}$ is the total variance:

$$T = \hat{U} + (1 + \frac{1}{m})B$$

5.1.3. Mixed-effects model with repeated measurements

A model that will be used for this study is the mixed-effects model with repeated measurements (MMRM). The model will include the fixed, categorical effects of country, gender, treatment, visit, treatment by visit interaction, country by visit interaction and gender by visit interaction as well as the continuous, fixed covariates of baseline and visit-by-baseline interaction:
*MMRM: change at each visit = baseline country gender treatment visit treatment*visit baseline*visit country*visit gender*visit*

The analysis will fit an unstructured (UN) variance-covariance matrix. This model allows different variances of the response at each time-point; that is there is no constraint of common variance at all times. In case of non-convergence of the algorithm with an unstructured variance-covariance matrix, the following structures will be tested in the following order: Heterogeneous Toeplitz, Heterogeneous compound symmetry, Toeplitz then Compound symmetry. The first (co)variance structure yielding convergence will be used as the primary analysis.
The Kenward-Roger method will be used for the approximation of the denominator degrees of freedom.

Estimate:
Inference is made at the selected time-point by constructing estimates of the effects at that time: these are adjusted treatment effects or 'LS-means'. The effects are based on the observations at that actual time-point, combined with contributions from all the observations at other times, mediated by the estimated covariance matrix.

5.1.4. Logistic regression

To analyse the CARS2 and CGI responders, a logistic regression model with the fixed categorical effect of treatment, country and gender will be used.

In case of quasi-completed separation of data with the model mentioned above, the following models will be tested sequentially, until no facing the warning anymore:

- A logistic regression model with the fixed categorical effect of treatment, pooled countries and gender
- A penalized logistic regression model with the fixed categorical effect of treatment, country and gender
- A penalized logistic regression model with the fixed categorical effect of treatment, pooled countries and gender.

5.2. Software

Statistical analyses will be performed using SAS®/PC Software version 9.4 and R version 3.1.1 for the non-parametric approach with adjustment.

6. REFERENCES

Guideline

Guideline on Missing Data in Confirmatory Clinical Trials – Adopted by CHMP, June 2010, issued as EMA/CPMP/EWP/1776/99 Rev. 1.

ICH E14 – The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic drugs – Adopted by CHMP, May 2005, issued as CHMP/ICH/2/04.

ICH E3 - Structure and Content of Clinical Study Reports – Adopted by CPMP, December 1995, issued as CPMP/ICH/137/95/step 5.

ICH E3 Q&A - Structure and Content of Clinical Study Reports - Questions & Answers (R1) - July 2012, issued as EMA/CHMP/ICH/435606/2012.

ICH E9 - Statistical Principles for Clinical Trials - Adopted by CPMP, March 1998, issued as CPMP/ICH/363/96/step 5.

ICH E9(R1) - Addendum – Adopted by CPMP, 30 January 2020, issued as EMA/CHMP/ICH/436221/2017.

Points to consider

Points to Consider on Adjustment for Baseline Covariates – Adopted by CPMP, May 2003, issued as CPMP/EWP/2863/99.

Points to Consider on Multiplicity Issues in Clinical Trials – Adopted by CPMP, September 2002, issued as CPMP/EWP/908/99.

Bibliography

Carpenter James and Kenward Michael G. "Multiple imputation: current perspectives." Statistical methods in medical research 16.3 (2007): 199-218.

Carpenter, James, and Michael Kenward. Multiple imputation and its application. John Wiley & Sons, 2012.

Jurek, Balthazar, Gulati, Novakovic, Núñez, Oakley, O'Hagan. Response (minimum clinically relevant change) in ASD symptoms after an intervention according to CARS-2: consensus from an expert elicitation procedure. European Child & Adolescent Psychiatry, 2021.

Khanna, Jariwala, Bentley. Health Utility Assessment Using EQ-5D among Caregivers of Children with Autism. Value in health 16 (2013) 778-788.

Kloke, John D., and Joseph W. McKean. "Rfit: Rank-based estimation for linear models." The R Journal 4.2 (2012): 57-64.

Dardas, Ahmad Validation of the World Health Organization's Quality of Life Questionnaire with Parents of Children with Autistic Disorder. J Autism Dev Disord (2014) 44:2257–2263

Limbers, Ripperger-Suhler, Heffer, Varni. Patient-Reported Pediatric Quality of Life Inventory™ 4.0 Generic Core Scales in Pediatric Patients with Attention-Deficit/Hyperactivity Disorder and Comorbid Psychiatric Disorders: Feasibility, Reliability, and Validity. Value in health 14 (2001) 521-530.

Nilsson ME, Suryawanshi S, Gassmann-Mayer C, Dubrava S, McSorley P, Kiang K. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide; 2013.

Rubin, Multiple imputation for non-response in surveys. New York: John Wiley & Sons, 1987.

Varni JW, Burwinkle TM, Seid M, Skarr D: The PedsQL 4.0 as a pediatric population health measure: Feasibility, reliability, and validity. Ambul Pediatr 2003; 3:329-41.