Statistical Analysis Plan

Trial ID: TAUASD

Taurine supplementation in children with autism spectrum disorders

——An Exploratory Trial

Authors:

Wennan He	**Trial Statistician (SAP writer)**
Weili Yan	**Senior Statistician Responsible**
Hao Zhou	**Principle Investigator**
Yun Chen	**Project Clinical Director**
Qun Deng, Zhongbi Peng	**Project Assistant**

Contents

Scope of the statistical analysis plan

- This statistical analysis plan(SAP) covers a more technical and detailed elaboration of statistical analysis and detailed procedures for executing the statistical analysis in TAUASD trial, which is consistent with the principal features of the statistical methods described in the protocol.
- A change log describing any amendment to the statistical analyses plan is provided on section 3.
- This is a supportive document and replication with the protocol is minimized[1].
- The SAP will be finalized before database lock (i.e. completion of last visit of last subject).

List of abbreviations

ASD	*Autism Spectrum Disorders*
ASRS	*Autism Spectrum Rating Scales*
ATEC	*Autism Treatment Evaluation Checklist*
CARS	*Childhood Autism Rating Scale*
CSHQ	*Children's Sleep Habits Questionnaire*
FAS	*Full analysis set*
fNIRS	*Resting-state functional near-infrared spectroscopy*
GLMM	*Generalized linear mixed model*
6-GSI	*Six Gastrointestinal Severity Index*
IMFeD	*Identification and Management of Feeding Difficulties*
ITT	*Intention-to-treat*
MAR	*Missing at random*
MI	*Multiple imputation*
M0	*Baseline*
M1	*the 1st month*
M2	*the 2nd month*
M3	*the 3rd month*
M6	*the 6th month*
M9	*the 9th month*
M12	*the 12th month*
PPAS	*Per-protocol analysis set*
PSI	*The Parenting Stress Index*
SAEs	*Serious adverse events*
SAP	*Statistical analysis plan*
SAS	*Safety analysis set*
SC	*Social/communication*
SRS	*Social Responsiveness Scale*
TEAEs	*Emergent adverse event*
UB	*Unusual behaviors*
WISC	*Wechsler Intelligence Scale*

1. Introduction

1.1 Trial information

1.1.1 Background

Autism Spectrum Disorders (ASD) is the most common neurodevelopment dysfunction and disabling disorder in childhood, with a lack of specific clinical treatment methods. Evidence-based study reveals that nutrients supplement can improve the core symptom of ASD. Our previous study demonstrated the children with ASD were found to have lower taurine levels in serum and urine samples. However, whether the taurine supplementation can improve the core symptom of ASD is unclear, and no ASD trial has ever focused on this intervention.

1.1.2 Objective(s)

1.1.1.1 Primary objective

To compare the effect of 3-month taurine supplementation (age-specific doses, 3 times/day) plus behavioral rehabilitation therapy versus placebo plus behavioral rehabilitation therapy on better rehabilitation of ASD core syndrome, reflected by differential age-specific ASD questionnaires evaluation from baseline at month 0 to month 3.

1.1.1.2 Secondary objectives

To compare the effect of taurine supplementation (age-specific doses, 3 times/day) plus behavioral rehabilitation therapy versus placebo plus behavioral rehabilitation therapy on better rehabilitation of ASD core syndrome, reflected by differential age-specific ASD questionnaires evaluation from baseline at 0 month to 6 month, 9month and 12 month-follow-up.

To compare the safety and adherence of taurine supplementation (age-specific doses, 3 times/day) plus behavioral rehabilitation therapy versus placebo plus behavioral rehabilitation therapy.

To compare the effect of taurine supplementation (age-specific doses, 3 times/day) plus behavioral rehabilitation therapy versus placebo plus behavioral rehabilitation therapy on:

- Brain function
- Intestinal flora
- Oxidative stress markers
- Bile acid metabolites

- Taurine concentrations

1.1.3 Endpoints

1.1.1.3 Primary endpoints

The primary endpoints addressing the primary objectives:

Change from baseline at month 0(M0) to month 3(M3) in:

- The Gesell score
- The Wechsler intelligence scale (WISC) score
- The Childhood Autism Rating Scale (CARS) score

Change from baseline at M0 to month 1(M1), month 2(M2), month 3(M3) in:

- The Autism Treatment Evaluation Checklist (ATEC) score
- The Autism Spectrum Rating Scales (ASRS)score
- The Social Responsiveness Scale (SRS) score
- The Identification and Management of Feeding Difficulties (IMFeD) score
- The Six Gastrointestinal Severity Index(6-GSI) score
- The Children's Sleep Habits Questionnaire (CSHQ) score
- The Parenting Stress Index (PSI)

1.1.1.4 Secondary endpoints

The secondary endpoints addressing the secondary objectives:

Change from baseline at M0 to month 6(M6) ,month 9(M9) and month 12(M12) in:

- The Gesell score
- The Wechsler intelligence scale (WISC) score
- The Childhood Autism Rating Scale (CARS) score
- The Autism Treatment Evaluation Checklist (ATEC) score
- The Autism Spectrum Rating Scales (ASRS)score
- The Social Responsiveness Scale (SRS) score
- The Identification and Management of Feeding Difficulties (IMFeD) score
- The Six Gastrointestinal Severity Index(6-GSI) score
- The Children's Sleep Habits Questionnaire (CSHQ) score
- The Parenting Stress Index (PSI)

1.1.1.5 Supportive endpoints

The supportive endpoints addressing the secondary objectives:

Effect endpoints:

Change from baseline at M0 to M3 on:

- Hemoglobin concentration in cortical metabolism
- Intestinal flora
- Serum metabolites level, including:
  - Oxidative stress markers
  - Bile acid metabolites
  - Taurine concentrations

Safety endpoints:

- Number of treatment emergent adverse events (TEAEs) from baseline at M0 to M3
- Number of serious adverse events (SAEs) from baseline at M0 to M3.
- Number of adverse events (AEs) from baseline at M0 to M3.

1.1.4 Type of trial

This is a 3-month intervention with a total of 12-month follow-up, randomized, double-blind, placebo-controlled, two-armed, parallel group, single-center, exploratory trial comparing taurine supplementation (age-specific doses, 3 times/day) with placebo on better rehabilitation of ASD core syndrome in ASD children aged 2-18 years.

2. Statistical considerations

2.1 Sample size determination

Sample size calculation is precluded from this trial due to exploratory design, as well as a lack of literature and data. 60 subjects (30 per group) are expected to meet the regular outpatient workload in the study site, which will be expected to reach the size within a year.

2.2 Definition of analysis sets

- The full analysis set (FAS):
  includes all randomized subjects according to the intention-to-treat (ITT) principle, regardless of actual treatment subjects receive. If any subject fails to complete at least one evaluation after randomization, modified ITT will be used as FAS instead.
- The per-protocol analysis set (PPAS):
  includes all randomized subjects who receive and complete the randomized intervention during the whole intervention period (i.e. M0 to M3).
- The safety analysis set (SAS):

includes all randomized subjects exposed to at least one dose of randomized treatment. Any observation excluded from the analysis will be documented before database lock with the reason for exclusion provided.

Two observation periods are defined for each subject:

- In-trial: The in-trial period is defined as the uninterrupted time interval from date of randomization to date of last contact with trial site.
- On-treatment: The on-treatment period will therefore be from the date of first trial product administration to date of last trial product administration excluding potential off-treatment time intervals triggered by at least 7-day missed doses.

The in-trial and on-treatment periods define the patient years of observation and patient years of intervention, respectively, as the total time duration in the periods.

2.3 Statistical analysis

Effect endpoints will be analyzed using the FAS and PPAS; safety endpoints will be analyzed using the SAS.

Results from effect endpoints will generally be accompanied by two-sided 95% confidence intervals(CI) and corresponding p-values. Significant difference will be claimed if two-sided 95% CI lie away from null. Effect size will be evaluated by corresponding Cohen's D values, which are calculated as mean difference between groups divided by standard deviation for each endpoint in corresponding analysis set (details in appendix Table 1).

2.3.1 Baseline characteristics

Baseline characteristics of subjects are defined as baseline values of the last available and eligible observation at or before randomization covering all subjects, including:

- Continuous:
  - Age (years)
  - Birth weight (kg)
  - Body Mass Index
  - Serum taurine concentration (relative concentration standardized by quality-control samples)
- Categorical:
  - Sex (1,Boy;2,Girl)
  - Nationality (1,Han; 2,Other)
  - Residential area (1,Urban;2, Rural)
  - Household monthly income (1,<=8500;2,>8500 and <20000;3,>=20000 and <50000; 4,>=50000)
  - Number of siblings (0,Single child; 1,One; 2,Two or more)

- Nourish style (0,Parental; 1,Grandparental; 2,Mixed)
- Paternal occupation (1,Civil servant;2 Farmer;3,Techican;4, Freelancer)
- Maternal occupation (1,Civil servant;2 Farmer;3,Techican;4, Freelancer)
- Allergy history (0,No; 1, Yes)
- Birth defects (0,No; 1, Yes)
- Abnormal birth history (0,No; 1, Yes)
- Serious medical history (0,No; 1, Yes)
- Receiving intervention previously (0,No; 1, Yes)
- Baseline physical examination (0,Normal; 1,Abnormal)
- ASD severity (0, Normal;1,Marginal;2,Mild;3,Medium;4,Severe;5,Extremely severe)

Baseline characteristics of subjects will be presented as mean and standard deviation for continuous variables with normal or approximately normal distribution; otherwise, median values and interquartile range (IQR) will instead be calculated. Categorical variables will be described as proportions(number of subjects within category ÷ all subjects×100%).

2.3.2 Primary endpoints

The primary endpoints are defined as change in scores of ASD-related scales from baseline at M0 to M3.

Descriptive statistics of primary endpoints

Mean and standard deviation of primary endpoints at each visit, and changes of mean and standard deviation from baseline to each visit will be presented.

Primary analysis for primary endpoints

Using unimputed data of primary endpoints in FAS, exploratory mixed-effects models, preferably generalized linear mixed model (GLMM) for repeated measures analysis, in the condition for endpoints of >2 visits. This analysis investigates if there are group-specific differences over time. The maximum likelihood method will be used to estimate the mean difference between the two groups and their 95% CI, with Gaussian distribution and identity link function; baseline score, group (group), time point (visit) and group × visit interaction as fixed effects. The time point(visit) is treated as categorical variable in the models. Random slope with unstructured covariance will be preferred as priority; if convergence is not able to be achieved, random intercept will be performed as an alternative.

Otherwise, ANCOVA will be performed to estimate the mean difference between the two groups and their 95% CI, with group (group) as independent variable, and baseline score, time point (visit) and group × visit interaction as covariates for outcomes with only 2 visits available. The available visits are defined as >10 observations at a single visit. The reason for this alternative method is explained in the Change Log.

For imputed data of primary endpoints, multiple imputation will be considered for handling missing outcomes. More details are described in 2.3.4.

Secondary analysis for primary endpoints

- **Per-protocol analysis**

Using unimputed data of primary endpoints in PPAS, exploratory mixed-effects models, preferably generalized linear mixed model (GLMM) for repeated measures analysis, in the condition for endpoints of >2 visits. Otherwise, ANCOVA will be performed if only 2 visits are available for the scores of ASD-related scales. The statistical method is the same as primary analysis for primary endpoints.

- **Covariate-adjusted analysis**

Covariate-adjusted analysis for primary endpoints is performed in FAS and PPAS using the same statistical method as primary analysis.

Covariates are baseline values (missing imputed) including:

- Age (years)
- Sex (1,Boy;2,Girl)
- Serum taurine concentration (relative concentration standardized by quality-control samples)

- The following covariates will be adjusted only if imbalanced baseline characteristics between group, and further filtered to avoid strong correlation or colinearity in case of linear violation.
- Birth weight (kg)
- Body Mass Index
- Nationality (1,Han; 2,Other)
- Residential area (1,Urban;2, Rural)
- Household monthly income (1,<=8500;2,>8500 and <20000;3,>=20000 and <50000; 4,>=50000)
- Number of siblings (0,Single child; 1,One; 2,Two or more)
- Nourish style (0,Parental; 1,Grandparental; 2,Mixed)
- Paternal occupation (1,Civil servant;2 Farmer;3,Techican;4, Freelancer)
- Maternal occupation (1,Civil servant;2 Farmer;3,Techican;4, Freelancer)
- Allergy history (0,No; 1, Yes)
- Birth defects (0,No; 1, Yes)
- Abnormal birth history (0,No; 1, Yes)
- Serious medical history (0,No; 1, Yes)
- Receiving intervention previously (0,No; 1, Yes)
- Baseline physical examination (0,Normal; 1,Abnormal)

Analysis based on imputation missing baseline covariates data

- Continuous: If no assessments are available, the mean of baseline values within the same sex and age(±1year) is used as the baseline value.
- Categorical: If no assessments are available, a new category coded as 999 is used as the baseline value.

2.3.3 Secondary and supportive endpoints

2.3.3.1 Secondary endpoints

Descriptive statistics of secondary endpoints

Mean and standard deviation of secondary endpoints at each visit, and changes of mean and standard deviation from baseline to each visit will be presented.

Analysis for Secondary endpoints of the same scale in M6,M9 and M12:

Using unimputed data of secondary endpoints with the same exploratory mixed-effects models as primary endpoints (i.e. the same scale in different time points will be included simultaneously), preferably generalized linear mixed model (GLMM) for repeated measures analysis. The statistical method is the same as primary analysis, per-protocol analysis and covariate-adjusted analysis for primary endpoints.

2.3.3.2 Supportive endpoints

Analysis for Effect endpoints:

- Change in hemoglobin concentration in cortical metabolism between M0 and M3:
  Using Functional near-infrared spectroscopy (fNIRS) analysis, by Wuhan Yiruide Medical Treatment Equipment New Technology Co.,Ltd., with detailed procedure described in a separate file provided by the company.
- Change in intestinal flora between M0 and M3:
  Using fecal metagenomics analysis[2], by Yunsheng Biotechnology Co(Shanghai, a company engaged in technology development and technical services in the field of biotechnology.), with detailed procedure described in a separate file provided by the company.
- Change in Serum metabolites level between M0 and M3:
  Using non-targeted metabolomics analysis, by Yunsheng Biotechnology Co(Shanghai, a company engaged in technology development and technical services in the field of biotechnology.), with detailed procedure described in a separate file provided by the company.

Analysis for Safety endpoints:

Number of TEAEs, SAEs and AEs will be calculated as counts with proportion (%) in each group and in total from baseline at M0 to M3.

2.3.4 Other Sensitivity analysis

Complete data analysis

- **Handling of missing data for primary and secondary endpoints**
  Missing at random (MAR) missingness mechanism is assumed for incomplete endpoints. For

missing data <50% of the endpoint, multiple imputation (MI) using regression (arbitrary missing pattern with 10 imputations) on missing data, dependent on baseline characteristics (age, sex), baseline scores is likely to be performed. Imputation will not be considered if missing data ⩾50% of the endpoint.

Therapy-adjusted analysis for primary endpoints

Based on covariate-adjusted analysis for primary endpoints, further adjusted for behavioral rehabilitation therapy (total class hours between M1,M2 and M3) to see the extent the effect size could be explained by the randomized intervention.

Analysis for sub-scales within the same primary endpoint

The statistical method is the same as descriptive statistics and primary analysis for primary endpoints. The analysis investigates if there are group-specific differences over time.

Analysis for sub-scales of certain ASD-related scale scores include:

- Gesell
  - Gesell DA adaptability
  - Gesell DA Fine movements
  - Gesell DA language
  - Gesell DA personal social
  - Gesell DA big movements
  - Gesell DQ adaptability
  - Gesell DQ Fine movements
  - Gesell DQ personal social
  - Gesell DQ big movements
- ATEC
  - Language
  - Social activity
  - Perception
  - Behavior
- SRS
  - Social perception
  - Social cognition
  - Social communication
  - Social motivation
  - SD behavior
- ASRS
  - Social/Communication
  - Unusual Behaviors

Dose-response analysis for primary endpoints

Using on-treatment period(see 2.2) for each subject, to perform dose-response analysis to see the effect of different on-treatment period (days without missed doses) on primary endpoints. The statistical method is the same as primary analysis. The analysis investigates if there are group-

specific differences over time.

Subgroup analysis

Subgroup analysis is not planned in this trial due to limited sample size.

2.3.5 Special statistical consideration

ASRS score calculation

Due to differential scale scoring between ASRS (2-5 years) and ASRS (6-18 years), the total score is unlikely to be weighed and used as a single composite endpoint. Instead, the standardized T score of ASRS in total , and two subscales (Social/Communication, SC and Unusual Behaviors, UB) will be calculated and combined between ASRS (2-5 years) and ASRS (6-18 years). The formulas and items for T score calculation is provided in Appendix Table 2.

2.3.6 Statistical software

STATA® (version 15.1, Stata Corp, Texas, USA) will perform all data analyses and generate most data displays. R4.1.2 (Vienna, Austria) may also be used for some data analyses and generating statistical graphs.

3. Change log

Version	Date	Reasons for change
0.0	18 Aug 2023	Clinical Trial registration
1.0	7 Mar 2024	Clinical assessments at M1 and M2 were cancelled because clinicians found it difficult for subjects to keep intense clinical visits;
2.0	10 Jan 2025	Secondary endpoints updated
2.5	17 Feb 2025	Supportive endpoint analysis and Other sensitivity analysis updated.
3.0	28 Feb 2025	As drop-outs in M6,M9,M12 for Gesell, CARS and WISC are more than expected (<10 at each visit), ANCOVA will be preferred as an alternative.

4. Appendix

Table 1 The formulas for calculating Cohen's D values

Condition	Statistical display	Formula

Population estimated mean group difference in change over time	$\beta_{T\text{-}T0}$ (95%CI)	$\left\lvert \frac{\beta_{T-T0}}{(upperlimit - lowerlimit) \times \frac{2 \times 1.96}{\sqrt{N_{T-T0}}}} \right\rvert$

Abbreviations: $_A$:placebo group; $_B$, taurine group; T, time visit; T0,baseline.

Table 2 The formulas for calculating ASRS T-score

Condition	Age range	Raw score of items	Formula[1]
SC	2-5 years	Σ(i1 i3 i4 i5 i7 i13 i14 i15 i16 i17 i18 i19 i21 i22 i24 i25 i28 i29 i30 i32 i33 i35 i36 i37 i38 i40 i43 i44 i49 i50 i51 i52 i54 i55 i57 i61 i62 i63 i67)	50+10×((raw score-50.06)/10.12)
	6-18 years	Σ(i3 i4 i8 i9 i12 i23 i28 i31 i32 i33 i39 i42 i43 i45 i55 i56 i61 i69 i7)	50+10×((raw score -24.3)/11.95)
UB	2-5 years	Σ(i2 i8 i9 i10 i11 i12 i20 i26 i27 i39 i41 i42 i45 i46 i47 i48 i53 i56 i60 i64 i65 i69 i70)	50+10×((raw score -50)/10.51)
	6-18 years	Σ(i2 i13 i17 i20 i21 i22 i24 i25 i26 i27 i29 i38 i40 i46 i48 i49 i50 i51 i54 i62 i63 i65 i67 i68)	50+10×((raw score -27.74)/10.79)
Total	2-5 years	-	50+10*((T score$_{SC+UB}$ -50.03)/10.35)
	6-18 years	-	50+10*((T score$_{SC+UB}$ -47.89)/7.85)

[1]the reference values for calculation are based on Chinese reference population[3, 4].
Abbreviations: SC, Social/Communication; UB, Unusual Behaviors

5. References

1. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P, Berlin J, Senn S, Day S, Barbachano Y and Loder E. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *Jama*. 2017;318:2337-2343.
2. Tringe SG and Rubin EM. Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet*. 2005;6:805-14.
3. Zhou H, Li C, Luo X, Wu L, Huang Y, Zhang L, Zou X, Xu X, Jiang YH, Yan W and Wang Y. Cross-Cultural Revision and Psychometric Properties of the Chinese Version of the Autism Spectrum Rating Scale (2-5 Years). *Front Neurol*. 2018;9:460.
4. Zhou B, Zhou H, Wu L, Zou X, Luo X, Fombonne E, Wang Y, Yan W and Xu X. Assessing the Accuracy of the Modified Chinese Autism Spectrum Rating Scale and Social Responsiveness Scale for Screening Autism Spectrum Disorder in Chinese Children. *Neurosci Bull*. 2017;33:168-174.