Department	:	Data management and Biostatistics
Information Type	:	Statistical Analysis Plan (SAP)

Title	:	A Randomized, Double-Blind, Single Center, Phase 2, Efficacy and Safety Study of Autologous HB-adMSCs vs Placebo for the Treatment of Patients with Multiple Sclerosis
Product	:	HB-adMSCs – Hope Biosciences adipose derived mesenchymal stem cells
Effective date	:	20-Sep-2021

Description:

- The purpose of this SAP is to describe the planned analyses and output to be included in the Clinical Study Report for Protocol HBPCOVID02.
- This SAP is intended to describe the planned safety, efficacy and tolerability analyses required for the study.
- This SAP is to convey the content of the complete statistical analysis deliverables.

1. INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to describe the analyses to be included in the Clinical Study Report for Protocol HBMS01.

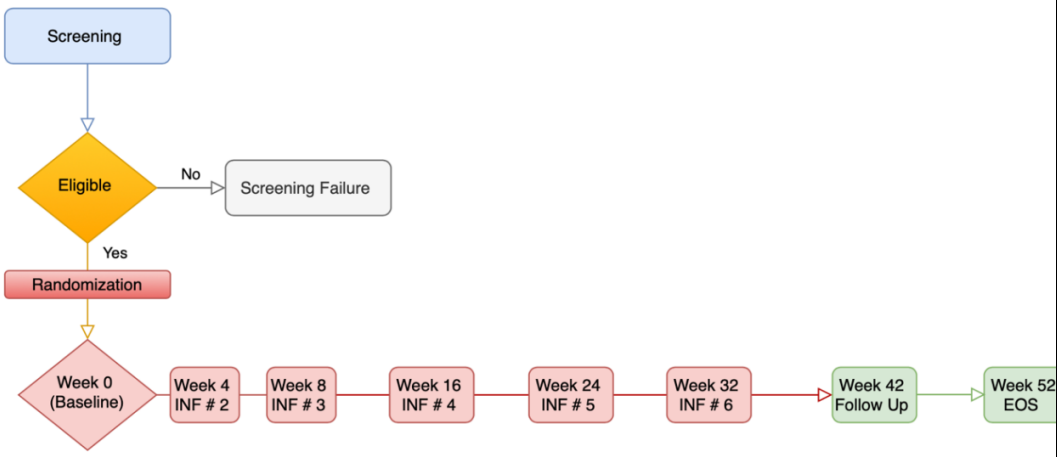
2. SUMMARY OF KEY PROTOCOL INFORMATION

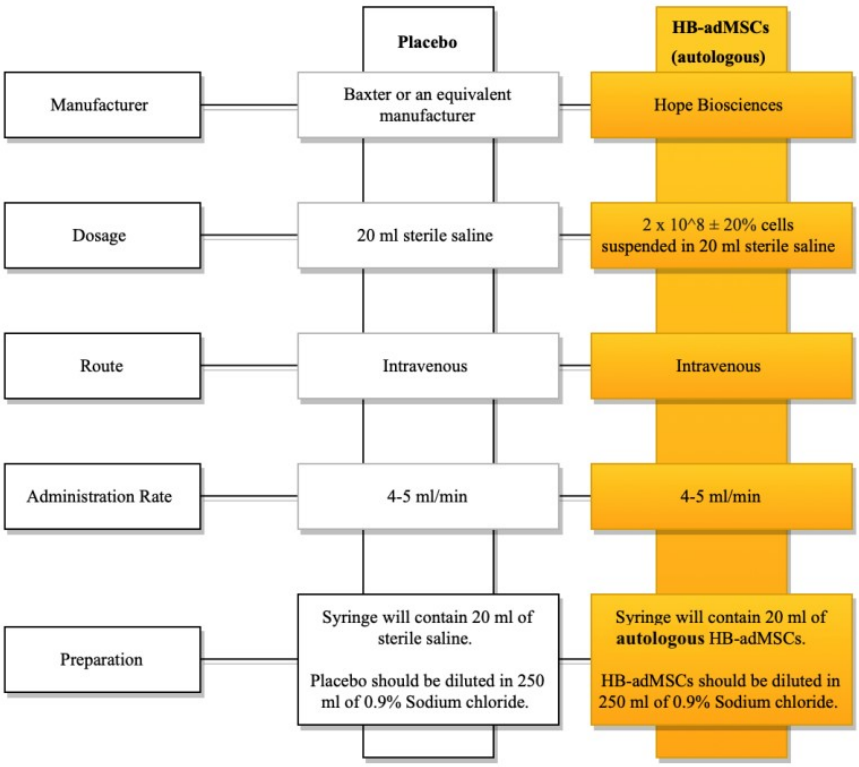
2.1. Study Objective(s), Estimand(s) and Endpoint(s)

Objectives	Endpoints
Primary Objective	**Primary Endpoint**
• To investigate the efficacy of intravenous infusions of HB-adMSCs vs Placebo in patients with Multiple Sclerosis as determined by improvements in Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument. (Time frame: Baseline to Week 52).	• Changes from Baseline at Weeks 52 in Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument scores following treatment with HB- adMSCs or Placebo. This 54-item questionnaire provides 12 subscales, two summary scores, and two extra single-item measures. The subscales are physical function, role limits-physical, emotional role restrictions, pain, emotional well-being, energy, health perceptions, social function, cognitive function, health distress, overall quality of life, and sexual function.
Secondary Objective	**Secondary Endpoint**
• To evaluate the efficacy of intravenous infusions of HB-adMSCs vs. Placebo in patients with Multiple Sclerosis as determined by changes in Expanded Disability Status Scale (EDSS) (Time frame: Baseline to Week 52).	• Change from baseline at Week 52 in Total Expanded Disability Status Scale (EDSS) scores following treatment with HB-adMSCs or Placebo
• To assess the efficacy of intravenous infusions of HB-adMSCs vs. Placebo in patients with Multiple Sclerosis as determined by changes in The Barthel Index (Time frame: Baseline to Week 52).	• Change from baseline at Week 52 in Barthel Index scores following treatment with HB- adMSCs or Placebo
• To determine the efficacy of intravenous infusions of HB-adMSCs vs. Placebo in patients with Multiple Sclerosis as determined by changes in 9-Hole Peg Test (Time frame: Baseline to Week 52).	• Change from baseline at Week 52 in 9-Hole Peg Test scores following treatment with HB- adMSCs or Placebo

Objectives	Endpoints
• To identify the safety of intravenous infusions of HB-adMSCs vs. Placebo in patients with Multiple Sclerosis as determined by changes in Patient Health Questionnaire (PHQ-9) (Time frame: Baseline to Week 52). • To assess the safety of intravenous infusions of HB-adMSCs vs. Placebo in patients with Multiple Sclerosis as determined by the incidence of adverse events or serious adverse events (Time frame: Baseline to Week 52).	• Change from baseline at Week 52 in Patient Health Questionnaire (PHQ-9) scores following treatment with HB-adMSCs or Placebo • Incidence of treatment-emergent Adverse Event (TEAEs) and serious Adverse Events (SAEs) • Incidence and risk of AEs of special interest (serious or nonserious), including thromboembolic events, peripheral events defined as, thromboembolism of the extremities, also infections and hypersensitivities. • Clinically significant changes in laboratory values, vital signs, weight, physical examination results and Multiple Sclerosis concomitant medications.

2.2. Study Design

Overview of Study Design and Key Features	
	
Design Features	This study will be a randomized, double-blind, single center, phase 2 study to assess efficacy and safety of multiple HB-adMSCs vs. Placebo for the treatment of Multiple Sclerosis. The trial includes, • A screening period of up to 4 weeks • A 32-week treatment period while on randomized study treatment ▪ Infusion 1 (Week 0), Infusion 2 (Week 4), Infusion 3 (Week 8), Infusion 4 (Week 16), Infusion 5 (Week 24) and Infusion 6 (Week 32) • A safety Follow-up period of 20 weeks after the last investigational product administration. ▪ Follow up at week 42 and end of study at Week 52 • 24 Subjects • 2 Treatment groups ▪ Group 1 → HB-adMSCs ▪ Group 2 → Placebo
Study Intervention	• Active Product: HB- adMSCs (Hope Biosciences adipose derived mesenchymal stem cells) ▪ Dose: 200 million ▪ Route: Intravenous ▪ Regimen: Weeks 0, 4, 8, 16, 24 and 32. • Placebo: Saline Solution 0.9% ▪ Dose: N/A ▪ Route: Intravenous ▪ Regimen: Weeks 0, 4, 8, 16, 24 and 32. • Duration of administration 1 hour • Study treatment details,

Overview of Study Design and Key Features	
	
Study Intervention Assignment	• Participants will be randomised 1:1 to receive HB-adMSCs active treatment or Placebo.
Interim Analysis	• No interim analysis will be performed in this study

3. STATISTICAL HYPOTHESES

The primary analysis will test whether HB-adMSCs is superior to Placebo according to the following statistical hypotheses:

Null hypothesis H0: The difference in change from baseline at Weeks 52 in Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument scores between treatment groups (HB-adMSCs – Placebo) is equal to zero.

Alternative hypothesis H1: The difference in change from baseline at Weeks 52 in Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument scores between treatment groups (HB-adMSCs – Placebo) is not equal to zero.

Secondary analysis will be tested for mean difference for change from Baseline to Week 52 for efficacy endpoints for both treatment groups,

Null hypothesis H0: $\Delta = 0$
Alternative hypothesis H1: $\Delta \neq 0$

3.1. Multiplicity Adjustment

The Bonferroni-Holm method for adjustment of multiplicity adjustment is performed for all secondary efficacy endpoint of interest.

Bonferroni-Holm method starts with ordering the p-values in increasing order and starts with testing the hypothesis with the lowest p-value on the 5/k% level (two-sided), where k is the number of hypotheses in the procedure. If significant, the testing proceeds to comparing the next lowest p-value with 5/(k-1)%.

4. ANALYSIS SETS

Population	Definition / Criteria	Analyses Evaluated
Safety analysis set	• All randomised subjects who received at least one dose of HB-adMSCs infusion or placebo. • If participants receive a treatment different to their randomized treatment, they will be analysed according to the treatment actually received.	• Safety • Study Population
Efficacy analysis set	• All randomized participants who received all 6 infusions of HB-adMSCs or placebo. • Participants will be analysed according to their randomized treatment.	• Efficacy
Screened Population	• This population consists of all subjects who signed an ICF to participate in the clinical trial. • This population will be used for summarizing screening failures and reasons for screening failures.	• Study Population

5. STATISTICAL ANALYSES

5.1. General Considerations

The final planned primary analyses will be performed after the completion of the following sequential steps:

1. All participants have completed (or withdrawn from) the study as defined in the protocol.
2. All required database cleaning activities have been completed and final database release and database lock has been declared by Data Management.

5.1.1. General Methodology

Unless otherwise stated, all hypotheses will be tested at a 2-sided significance level of 0.05 and 95% confidence interval. All continuous measurements will be summarised descriptively at each visit by treatment using observed data.

Summary of continuous variables will be presented using N, Mean, 95% confidence interval of mean, Standard Error of mean (SE), Standard Deviation (SD), Median and Range (Minimum and Maximum). The categorical variables will be presented using number and percentage based on N.

For measurements over time mean values will be plotted to explore the trajectory over time. Observed data will be used as the basis for plotting data along with bars as +/- SE, not otherwise specified.

As a primary analysis, a standard Analysis of Covariance (ANCOVA) will be applied for primary and secondary endpoint to test the significance of the effects of the treatment at Week 52. The model includes treatment as fixed factor and the corresponding baseline value as a covariate.

A parametric Repeated Measures Analysis (RMA) Model will be applied as a secondary analysis to test the significance of the effects of the treatment at Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52 including baseline as covariate.

Presentation of results from a statistical analysis model will include the estimated mean treatment effects (Least Square Means (LSMeans)). For all endpoints analysed statistically, estimated mean treatment differences will be presented together with two-sided 95% confidence intervals and p-values,

HB-adMSCs - Placebo

Pairwise t-test will be performed on the efficacy endpoints to test the difference between baseline and week 52 (EOS) as secondary analysis. Data for all the efficacy outcomes will be checked for normality (Shapiro-Wilk Test). When there is a larger deviation of data distribution from normality, appropriate non-parametric test will be used.

Additionally, RMA and ANCOVA in a Bayesian framework using a non-informative prior will be conducted to estimate the posterior distribution and to estimate the probablity that the true treatment difference of change from baseline to Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52.

Study population analyses including analyses of subject disposition, demographic and baseline characteristics, medical history, prior and concomitant medications.

Disposition summary includes, subject screened, randomized and disposition at end of study – Week 52 along with reasons for withdrawals. Subjects in different analysis populations also will be presented.

The screen failure table includes total number of screened subjects and reasons. The percentage in the screen failure table will be calculated based on total number of screened subjects as denominator.

5.1.2. Baseline Definitions

For all endpoints, the baseline value will be the latest pre-treatment assessment visit with a non-missing value. i.e., If an assessment has been made both at screening visit (Visit 1) and Week 0 infusion 1 visit (Visit 2, Week 0), the value from the Week 0 visit is used as the baseline value. If the value measured at the Week 0 visit is missing and

the assessment also has been made at screening, then the screening value is used as the baseline value.

Unless otherwise stated, if baseline data is missing no derivation will be performed and baseline will be set to missing.

Change from baseline calculated as: Post-baseline value – Baseline

5.2. Primary Endpoint(s) Analyses

The primary objective of this study is to compare HB- adMSCs to Placebo on Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument scores. Efficacy analysis set will be used for this analysis. The details of the planned displays are in programming specification document.

5.2.1. Definition of endpoint(s)

Objectives	**Endpoints**
Primary Objective	**Primary Endpoint**
• To investigate the efficacy of intravenous infusions of HB-adMSCs vs Placebo in patients with Multiple Sclerosis as determined by improvements in Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument. (Time frame: Baseline to Week 52).	• Changes from Baseline at Weeks 52 in Multiple Sclerosis Quality of Life (MSQOL)-54 Instrument scores following treatment with HB-adMSCs or Placebo. This 54-item questionnaire provides 12 subscales, two summary scores, and two extra single-item measures. The subscales are physical function, role limits-physical, emotional role restrictions, pain, emotional well-being, energy, health perceptions, social function, cognitive function, health distress, overall quality of life, and sexual function.

Data from the MSQOL-54 will be scored according to the scoring algorithm suggested by the authors of the questionnaires (Vickrey 1995). Scale scores will be generated by averaging the items within scales and transforming the mean scores linearly from 0 to 100 possible scores, with higher scores indicating a better QOL. Physical and mental health composite scores will be generated by weighting the sum of selected scales to generate a simplified two-factor solution and the composite scores of the MSQOL-54.

Derivation of MSQOL-54 Composite Scores

Two composite scores are defined for MSQOL-54 using the subscale scores defined in the table below. The MSQOL-54 Physical Health Composite score and the MSQOL-54 Mental Health Composite Scores are calculated as a weighted-sum using the subscale scores and the weights defined in table below. If a subscale score is missing then it will not contribute to the Composite Score calculation (in effect, it is treated as a score of 0).

Composite Score	Subscale	Weight for Composite
Physical Health	Physical Function	0.17
	Health Perceptions	0.17
	Energy	0.12
	Role Limitations - Physical	0.12
	Pain	0.11
	Sexual Function	0.08
	Social Function	0.12
	Health Distress	0.11
Mental Health	Health Distress	0.14
	Overall Quality of Life	0.18
	Emotional Well-being	0.29
	Role Limitations - Emotional	0.24
	Cognitive Function	0.15

5.2.2. Main analytical approach

Primary analysis
Endpoint/Variables
• Change from baseline at Weeks 52 in Multiple Sclerosis Quality of Life (MSQOL)-54 Composite scores following treatment with HB-adMSCs or Placebo.
ANCOVA Model Specification

• To compare the HB-adMSCs to Placebo on Change from baseline to Week 52, a standard Analysis of Covariance (ANCOVA) model will be fitted
• Terms fitted in the mixed effect model will include:
  • Primary endpoint efficacy measurements available at - Week 52 will be response variable in a linear mixed model using a variance-covariance residual matrix.
  • The model will include:
  • Fixed factors: treatment
  • Fixed strtatification factors: Multiple's sclerosis severity (Mild disability (EDSS < 4) ; Moderate disability (EDSS > 4 and < 6.5)), Age (< 55 ; > 55) and Gender (Female ; Male)

- Covariate: baseline

Model Checking and Diagnostics

For the ANCOVA Model assumptions will be checked, and appropriate adjustments may be applied based on the data.

Distribution assumptions underlying the model used for analysis will be examined by obtaining a normal probability plot of the residual and a plot of the residuals versus the fitted values (i.e., checking the normality assumption and constant variance assumption of the model respectively) to gain confidence that the model assumptions are reasonable.

Non-parametric analyses, Mann-Whitney U (or Wilcoxon Rank-Sum) Test will be conducted if the normality assumptions do not hold.

Results presentation

Presentation of results from a statistical analysis model will include the estimated mean treatment effects (Least Square Means (LSMeans)) and estimated mean treatment differences will be presented together with two-sided 95% confidence intervals standard error for mean difference and p-value.

Secondary analysis

Endpoint/Variables

Change from baseline in Multiple Sclerosis Quality of Life (MSQOL)-54 Composite scores following treatment with HB-adMSCs or Placebo at Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52.

Repeated Measures Model Specification

- To compare the HB-adMSCs to Placebo on Change from baseline, a Repeated measures analysis (RMA) model will be fitted
- Terms fitted in the mixed effect model will include:
  - All primary endpoint efficacy measurements available at post-bseline at scheduled measurements will be response variable in a linear mixed model using an unstructured residual covariance matrix.
- The model will include:
  - Fixed factors: treatment and visit
  - Fixed strtatification factors: Multiple's sclerosis severity (Mild disability (EDSS < 4) ; Moderate disability (EDSS > 4 and < 6.5)), Age (< 55 ; > 55) and Gender (Female ; Male)
  - Covariate: baseline

- Furthermore, the model will include:
  - Interaction terms between treatment and visit

Model Checking and Diagnostics

For the Repeated Measures Analysis (RMA) Model assumptions will be checked, and appropriate adjustments may be applied based on the data.

Distribution assumptions underlying the model used for analysis will be examined by obtaining a normal probability plot of the residual and a plot of the residuals versus the fitted values (i.e., checking the normality assumption and constant variance assumption of the model respectively) to gain confidence that the model assumptions are reasonable.

Non-parametric analyses, Mann-Whitney U (or Wilcoxon Rank-Sum) Test will be conducted if the normality assumptions do not hold.

Results presentation

Presentation of results from a statistical analysis model will include the estimated mean treatment effects (Least Square Means (LSMeans)) and estimated mean treatment differences will be presented together with two-sided 95% confidence intervals standard error for mean difference and p-value.

Additional analysis

Summary analysis method

To compare difference between baseline to Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52, a paired t-test will be used for primary endpoint change from baseline Multiple Sclerosis Quality of Life (MSQOL)-54 Composite scores to check the statistical significance.

To further evaluate if the improvements are clinically relevant, effect size of the treatment difference at Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52 will also be calculated for the efficacy endpoint using Cohen's d calculation. Cohen's d is a standardized measure of effect size (ES) that provides information on the amount of change in the outcome meaure relative to the variation within the measure. Cohen's d is calculated as the difference between the HB-adMSCs mean score and Placebo mean scores divided by the standard deviations of the scores. An absolute effect size of <0.2 will be considered trivial, ≥ 0.2 as small, ≥0.5 as medium and >0.8 as large [1].

Cohen's d $= \frac{\mathrm{M1-M2}}{(sd1+sd2)/2}$, where M1 and M2 are the Mean scores for HB-adMSCs and Placebo respectively, and SD1 and SD2 are the corresponding standard deviations.

Results presentation

Secondary endpoints will be summarized using n, mean, confidence interval of mean, SD, median, minimum and maximum. And p-value for paired comparison will be displayed for all post-baseline visits.

Normality checking

Distribution assumptions used for the statistical analysis will be examined by obtaining a normal plot and Shapiro-Wilk Test.

Non-parametric analyses, Wilcoxon signed-rank test will be conducted if the normality assumption does not hold.

Supportive Statistical Analysis

Bayesian Method 1 Endpoint/Variable

Change from baseline at Weeks 52 in Multiple Sclerosis Quality of Life (MSQOL)-54 Composite scores following treatment with HB-adMSCs or Placebo.

Bayesian model specification

Construct a linear model for each visit in order to model within-subject observation covariance structures by multivariate normal (MVN) distribution in the MCM procedure.

For subject i on treatment j the model can be written as:

$CHGijk=\beta0 + \beta1\ treatment + \beta2\ baseline + \beta3strataEDSS + \beta4strataAge + \beta5strataGender + \varepsilon ij\ \ \varepsilon ij \sim Normal\ (0, \sigma^2)$

Where CHGij refers to change from baseline in subject i = 1, 2, ..., n indicator variable, which is intrinsically formed within PROC MCMC from the design matrix.

Initial values of the MCMC chains will be selected at random. Seed for random number generation will be specified as "seed = 12345" with three MCMC chains of 8000 total iterations each (with 4000 of these discarded as burn-in iterations). The simulation size and number of burn-in interations can be updated during the convergence check.

The choice of prior distribution for level two variances will follow Gelman's recommendations from the literature. Models will use weakly informative priors to maximize the influence of the present data on posterior probabilities (PP). The non-informative priors for the model parameters ~ Normal (0, σ^2= 10) and 1/σ2 ~ Gamma (0.001, rate = 0.001) will be used.

Models were evaluated via posterior probability guidelines in the literature, suggesting that PP = 75% to 90% indicates "moderate evidence," PP = 91% to 96% indicates "strong evidence," and PP ≥ 97% indicates "very strong to "extreme evidence."

Consistent with prior research, a PP ≥75% (equivalent to a Bayes factor = 0.33 or 3.00) that an effect exists will be taken as a minimum threshold of evidence in favor of the alternative hypothesis. This probability was chosen to emphasize the value in identifying a signal for the effects of treatment group, time, and their interaction.

The median, standard deviation, and 95% credible intervals (CrI) of the posterior distribution will be used to provide a point estimate and corresponding range of uncertainty for the magnitude of each predictor effect, including treatment group.

Model Checking and Diagnostics

Convergence diagnostics via scale reduction factors (Rhat), effective sample size, and posterior predictive distributions will be examined to ensure satisfaction of Bayesian modeiling assumptions, including Rhat < 1.01, sufficiently large effective sample size, and graphical inspection that the observed distribution of each outcome fell within the range of distributions produced by 1000 replications drawn from the posterior predictive distributions of the outcome.

Adequate values for the number of MCM samples/thinning/number of burn-in samples should be chosen to ensure that the ratio Monte Carlo Standard Errors (MCSE) and standard deviation of the posterior distribution for all the parameters in the model as small as possible, typically close to 0.01.

In addition, if possible, the number of tuning units and maximum number of tuning iterations may be increased to find a better proposal distribution for the model parameters, which in turn may reduce the MCSE/SD ratio.

The Geweke diagnostics test checks whether the mean estimates have converged by comparing means from the early and latter part of the Markov chain using a z score t-test. Large absolute values of the z-score statistic indicate rejection of the null hypothesis of no difference between the mean estimates obtained from the early and latter parts of the chain.

The convergence diagnostics for all parameters in the Bayesian analysis will be visually checked by the trace plots.

If the trace plots show apparent trend or the autocorrection plots show significant positive or negative correlation, number of iterations will be increased or reparameterization might be explored.

Results presentation

Presentation of results from the Bayesian statistical models will include the estimated conditional/marginal mean treatment effects and standard deviation. For all endpoints analysed statistically, estimated mean treatment differences will be presented together with 95% credible intervals and posterior probabilities.

Bayesian Method 2 Endpoint/Variable
Change from baseline in Multiple Sclerosis Quality of Life (MSQOL)-54 Composite scores following treatment with HB-adMSCs or Placebo at Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52.
Bayesian model specification
Construct a linear model for each visit in order to model within-subject observation covariance structures by multivariate normal (MVN) distribution in the MCM procedure. For subject i on treatment j at Visit k, the model can be written as: $CHGijk=\beta0 + \beta1\ treatment + \beta2\ baseline + \beta3\ Visitk + \beta3strataEDSS + \beta4strataAge + \beta5strataGender + \varepsilon ijk$ $\varepsilon ijk \sim Normal\ (0, \sigma^2)$ Where CHGijk refers to change from baseline in subject i = 1, 2, ..., n and Visitk for k= 1, 2, 3, 4, 5, 6 represents Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52, respectively are indicator variables, which are intrinsically formed within PROC MCMC from the design matrix. Initial values of the MCMC chains will be selected at random. Seed for random number generation will be specified as "seed = 12345" with three MCMC chains of 8000 total iterations each (with 4000 of these discarded as burn-in iterations). The simulation size and number of burn-in interations can be updated during the convergence check. The choice of prior distribution for level two variances will follow Gelman's recommendations from the literature. Models will use weakly informative priors to maximize the influence of the present data on posterior probabilities (PP). The non-informative priors for the model parameters ~ Normal ($0, \sigma^2= 10$) and $1/\sigma2 \sim$ Gamma (0.001, rate = 0.001) will be used. Models were evaluated via posterior probability guidelines in the literature, suggesting that PP = 75% to 90% indicates "moderate evidence," PP = 91% to 96% indicates "strong evidence," and PP ≥ 97% indicates "very strong to "extreme evidence." Consistent with prior research, a PP ≥75% (equivalent to a Bayes factor = 0.33 or 3.00) that an effect exists will be taken as a minimum threshold of evidence in favor of the alternative hypothesis. This probability was chosen to emphasize the value in identifying a signal for the effects of treatment group, time, and their interaction. The median, standard deviation, and 95% credible intervals (CrI) of the posterior distribution will be used to provide a point estimate and corresponding range of uncertainty for the magnitude of each predictor effect, including treatment group.
Model Checking and Diagnostics
Convergence diagnostics via scale reduction factors (Rhat), effective sample size, and posterior predictive distributions will be examined to ensure satisfaction of Bayesian

modeiling assumptions, including Rhat < 1.01, sufficiently large effective sample size, and graphical inspection that the observed distribution of each outcome fell within the range of distributions produced by 1000 replications drawn from the posterior predictive distributions of the outcome. Adequate values for the number of MCM samples/thinning/number of burn-in samples should be chosen to ensure that the ratio Monte Carlo Standard Errors (MCSE) and standard deviation of the posterior distribution for all the parameters in the model as small as possible, typically close to 0.01. In addition, if possible, the number of tuning units and maximum number of tuning iterations may be increased to find a better proposal distribution for the model parameters, which in turn may reduce the MCSE/SD ratio. The Geweke diagnostics test checks whether the mean estimates have converged by comparing means from the early and latter part of the Markov chain using a z score t-test. Large absolute values of the z-score statistic indicate rejection of the null hypothesis of no difference between the mean estimates obtained from the early and latter parts of the chain. The convergence diagnostics for all parameters in the Bayesian analysis will be visually checked by the trace plots. If the trace plots show apparent trend or the autocorrection plots show significant positive or negative correlation, number of iterations will be increased or reparameterization might be explored.
Results presentation
Presentation of results from the Bayesian statistical models will include the estimated conditional/marginal mean treatment effects and standard deviation. For all endpoints analysed statistically, estimated mean treatment differences will be presented together with 95% credible intervals and posterior probabilities.

5.3. Secondary Endpoint(s) Analyses

5.3.1. Efficacy Endpoints / Variables

- Change from baseline at Week 52 in Total Expanded Disability Status Scale (EDSS) scores following treatment with HB- adMSCs or Placebo

- Change from baseline at Week 52 in Barthel Index scores following treatment with HB- adMSCs or Placebo

- Change from baseline at Week 52 in 9-Hole Peg Test scores following treatment with HB- adMSCs or Placebo

- Change from baseline at Week 52 in Patient Health Questionnaire (PHQ-9) scores following treatment with HB-adMSCs or Placebo

Primary analysis

Primary analysis of Repeated Measures Analysis (RMA) Model will be performed for above secondary efficacy endpoints above as given in the Section 5.2.2.

Secondary analysis

Secondary analysis of ANCOVA model will be perfomed to test the significance of treatment difference for secondary efficacy endpoints above as given in the Section 5.2.2.

Additional analysis

Summary analysis will be perfomed to compare difference between baseline to Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52, a paired t-test performed for secondary efficacy endpoints above as given in the Section 5.2.2.

To further evaluate if the improvements are clinically relevant, effect size of the treatment difference at Week 4, Week 8, Week 16, Week 24, Week 32, Week 42 and Week 52, Cohen's d will also be calculated for the secondary efficacy endpoints as given in the Section 5.2.2.

5.3.2. Safety Analyses

The safety analyses will be based on the Safety analysis set, unless otherwise specified.

5.3.2.1. Extent of Exposure

- Number of days of exposure to study drug will be calculated based on the formula:

Duration of Exposure in Weeks = (Treatment stop date – Treatment start date)/7

- Duration will be summarized by treatment group. Each subject will contribute duration of exposure to the treatment taken.
- Participants who were randomized but not report a treatment start date will be categorised as having zero days of exposure.

A listing and summary table of exposure will be created.

The details of the planned displays are in programming specification document.

5.3.2.2. Adverse Events

AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA).

A treatment emergent adverse event (TEAE) is defined as an event that has onset date on or after the first day of exposure to infusion treatment and on or before the first safety follow up at week 42/early withdrawal date. Here the first day of exposure is defined as the first day of exposure to infusion treatment.

Treatment emergent Adverse events TEAEs are summarized descriptively, whereas non-TEAEs are presented in listings. TEAE data will be displayed in terms of the number of subjects with at least one event (N), percentage of subjects with at least one event (%) and the number of events (E).

Summaries of TEAEs and of serious AEs will be presented as an overview including all AEs, serious AEs, AEs by severity, AEs by relation to treatment, action to AEs and treatment advised, and outcome of AEs.

Furthermore, summary tables based on system organ class and preferred terms are made for:

- All TEAEs
- Serious AEs
- AEs leading to withdrawal of study
- Incidence and risk of AEs of special interest (serious or nonserious), including thromboembolic events, peripheral events defined as, thromboembolism of the extremities, also infections and hypersensitivities.

Individual adverse events will be listed.

The details of the planned displays are in programming specification document.

5.3.2.3. Clinical Laboratory data

Laboratory evaluations will include the analyses of Biochemistry laboratory tests, Hematology laboratory tests and Urinalysis. The details of the planned displays are in programming specification document.

All laboratory parameters, including numerical urine analysis parameters will be summarized descriptively. Categorical urine analysis results will be summarized using count and percentage based on subjects.

Results of urine pregnancy test will be listed in individual subject data listings only.

Individual laboratory evaluations will be listed. In addition, a listing containing individual subject laboratory values outside the normal reference ranges will be provided.

Data recorded at unscheduled assessments will not be included in tables and figures but will be listed.

5.3.3. Additional Safety Assessments

The analyses of non-laboratory safety test results will include physical examination and vital signs.

Physical Examination and Vital signs will be summarized using count and percentage based on subjects. The vital signs based on visit and change from baseline will be summarized using descriptive statistics.

Individual Vital signs, Physical Examination evaluations will be listed.

The details of the planned displays are in programming specification document.

5.4. Changes to Protocol Defined Analyses

Analysis is planned as per protocol. No deviation from the planned protocol specified analysis.

6. SAMPLE SIZE DETERMINATION

Sample size consists of up to 24 subjects who have been diagnosed with Relapsing Remitting Multiple Sclerosis.

7. SUPPORTING DOCUMENTATION

7.1. Appendix 1 Study Population Analyses

Unless otherwise specified, the study population analyses will be based on the "Safety" population. Screen failures will be summarized or listed based on the "Screened" population. A summary of the number of participants in each of the participant level analysis set will be provided.

7.1.1. Subject Disposition

A summary of the number and percentage of subjects who completed the study as well as those who withdrawn from the study will be provided by treatment. Reason of study withdrawn will be summarized by treatment.
A summary of the study intervention status will be provided. This display will show the number and percentage of subjects who have completed the Week 52, as well as primary reasons for withdrawn.

The study analysis set will be summarized in the subject disposition table.

The details of the planned displays are in programming specification document.

7.1.2. Demographic and Baseline Characteristics

The demographic characteristics including, age, sex, ethnicity, race, height at baseline, weight at baseline, BMI at baseline will be summarized with descriptive statistics. In addition, the following categories will be summarized: 18-64, 65-84 and >=85 based on the randomized analysis set.

Listings of demographic characteristics will also be produced.

The details of the planned displays are in programming specification document.

7.1.3. Protocol Deviations

Important protocol deviations will be summarized.

Protocol deviations will be tracked by the study team throughout the conduct of the study.

The details of the planned displays are in programming specification document.

7.1.4. Prior and Concomitant Medications

Prior and concomitant medications will be coded using WHO Drug Dictionary. Concomitant medications will be summarized as number and percentage of subjects. For classifying study phase for concomitant medications, use the following definition.

Study Phase	Definition
Prior	If medication end date is not missing and is before date of first dose of study medication.
Concomitant	Any medication that is not a prior

Please refer to Section 7.2.6 for handling of missing and partial dates for concomitant medication.

The details of the planned displays are in programming specification document.

7.1.5. Study Intervention Exposure and Compliance

A summary of Overall cumulative exposure to HB-adMSCs will be produced.

The details of the planned displays are in programming specification document.

7.2. Appendix 2 Data Derivation Rule

7.2.1. Criteria for Potential Clinical Importance

The potential clinical importance criteria are not defined this trial. A laboratory value that is outside the reference range is considered either high abnormal or low abnormal will be displayed based on lab normal range data.

7.2.2. Study Period

Adverse events will be classified according to the time of occurrence relative to the study intervention period.

Treatment emergent	Definition
Y	If event start date is not missing and is before date of first dose of study medication.
N	Any event started on or after date of first dose of study medication or event date is missing or partial

7.2.3. Study Day and Reference Dates

Study Day

• Study Day 1 is defined as the day the first dose was taken.
• Study day >1 is calculated as the number of days from the date of the Study Day 1:
  • Ref Date = Missing → Study Day = Missing
  • Ref Date < Date of Study Day 1 → Study Day = Ref Date – Date of Study Day 1
  • Ref Data ≥ Date of Study Day 1 → Study Day = Ref Date – (Date of Study Day 1) + 1

7.2.4. Assessment Window

For data summaries by visit, scheduled visits with nominal visit description will be displayed. Unscheduled visits will not be displayed or slotted into a visit window. While in the baseline derivation or post-baseline worst scenarios are derived, unscheduled visits are considered. All unscheduled visits will be displayed in listings, as appropriate.

7.2.5. Multiple measurements at One Analysis Time Point

For lab tests on a study day, if more than one assessment is taken on the same day, the test from the latest non-missing lab measurements will be used for the analysis. All lab measurements will be displayed in the listings, as appropriate.

7.2.6. Handling of Missing and Partial Dates

<table>
<tr><th>Element</th><th>Reporting Detail</th></tr>
<tr><td>General</td><td>• Partial dates will be displayed as captured in participant listing displays.</td></tr>
<tr><td>Adverse Events</td><td>• Partial dates for AE recorded in the CRF will be imputed using the following conventions:
<table>
<tr><td>Missing start day</td><td>If study intervention start date is missing (i.e, subject did not start the study medication), then set start date = 1st of month.

Else if study intervention start date is not missing:

• If month and year of start date = month and year of study intervention start date, then
o If stop date contains a full date and stop date is earlier than study intervention start date, then set start = 1st of month.
o Else set start date = study intervention start date.

• Else set start date = 1st of month</td></tr>
</table></td></tr>
</table>

Element	Reporting Detail		

		Missing start day and month	If study intervention start date is missing (ie., subject did not start study medication), then set start date = January 1.
			
			• Else if study intervention start date is not missing:
			  o If year of start date = year of study intervention start date, then
			  ▪ If stop date contains a full date and stop date is earlier than study intervention start date, then set start date = January 1.
			  o Else set start date = study intervention start date.
			Else set start date = January 1.

		Missing end day	A '28/29/30/31' will be used for the day (dependent on the month and year)
		Missing end day and month	No imputation
		Completely missing start/end date	No imputation

<table>
<tr><th>Element</th><th>Reporting Detail</th></tr>
<tr><td>Concomitant Medications</td><td>

Partial dates for any concomitant medications recorded in the CRF will be imputed using the following convention:
<table>
<tr><td>Missing start day</td><td>If study intervention start date is missing (i.e, subject did not start the study medication), then set start date = 1st of month.

Else if study intervention start date is not missing:

If month and year of start date = month and year of study intervention start date, then

If stop date contains a full date and stop date is earlier than study intervention start date, then set start = 1st of month.
Else set start date = study intervention start date.

Else set start date = 1st of month

</td></tr>
</table>

</td></tr>
</table>

Element	Reporting Detail		
		Missing start day and month	If study intervention start date is missing (ie., subject did not start study medication), then set start date = January 1. • Else if study intervention start date is not missing: o If year of start date = year of study intervention start date, then ▪ If stop date contains a full date and stop date is earlier than study intervention start date, then set start date = January 1. o Else set start date = study intervention start date. Else set start date = January 1.
		Missing end day	A '28/29/30/31' will be used for the day (dependent on the month and year)
		Missing end day and month	A '31' will be used for the day and 'Dec' will be ised for the month
		Completely missing start/end date	No imputation

8. REFERENCES

Hope Biosciences protocol number HBMS01 A Randomized, Double-Blind, Single Center, Phase 2, Efficacy and Safety Study of Autologous HB-adMSCs vs Placebo for the Treatment of Patients with Multiple Sclerosis

Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. Qual Life Res. 1995 Jun;4(3):187-206. doi: 10.1007/BF02260859. PMID: 7613530.