Myopia Treatment Study 2
(MTS2)

A Randomized Placebo-Controlled Trial of Spectacles with Highly Aspherical Lenslets or 0.05% Atropine to Slow Progression of Myopia in Children

Statistical Analysis Plan

15 July 2025

Version History

This SAP was written with reference to protocol version 1.0. If the protocol is subsequently updated, then this SAP will be reviewed to ensure consistency with the new protocol. The SAP will not be revised unless the protocol changes require modification of the analyses.

Version	Protocol Version	Author	Approver	Effective Date	Study Stage
1.0	1.0	Rui Wu	Wesley Beaulieu	15 Jul 2025	The first participant has not yet been enrolled.

Version	Revision Description
1.0	Original Version

Approvals

Role	Digital Signature or Handwritten Signature/Date
Author (Statistician)	Rui Wu Digitally signed by Rui Wu DN: cn=Rui Wu ou=North Wing Reason: I am the author of this document Location: Date: 2025-07-15 09:55-04:00
Approver (Senior Statistician)	Wesley Beaulieu I agree to the terms defined by the placement of my signature in this document 2025-07-15 08:46-04:00

1. Study Overview

The primary objective of MTS2 is to compare the effectiveness of 0.05% atropine eye drops versus placebo and of HAL lenses vs. single vision lenses for slowing myopia progression (change in axial length) over a two-year treatment period in children aged 5 to less than 12 years with spherical equivalent refractive error (SER) myopia of 0.75 D to 6.00 D and at least 0.75 D myopia in both principal meridians of each eye at the time of enrollment. On-treatment follow-up visits are at 6, 12, 18, and 24 months (primary). Randomized treatment stops at 24 months, and participants return at 30 months to assess for rebound effects. Complete eligibility criteria and study procedures are described in the study protocol.

Participants will be randomized 1:1:1:1 to placebo eye drops + single vision lenses, 0.05% atropine eye drops + single vision lenses, HAL lenses + placebo eye drops, or HAL lenses + 0.05% atropine eye drops. Randomization is stratified by baseline age (5 to <9 years, 9 to <12 years). Analyses involving the combination group (HAL lenses + 0.05% atropine eye drops) are considered exploratory. Sample size has been set at 87 per group (261 total); details are in the study protocol.

2. Consistency with the Protocol

The author of this document has confirmed the analyses described here are consistent with the version of the protocol indicated on the version history page except for the following:

- Flat corneal radius, anterior chamber depth, and lens thickness will not be collected and therefore the analyses of these outcomes have been removed.
- Monocular amplitude of accommodation will be analyzed at 18 and 30 months in addition to 6, 12, and 24 months.
- The Barnard unconditional exact test will be used to compare binary safety outcomes instead of the Fisher exact test.
- The adaptive false discovery rate procedure will be used to account for multiplicity in exploratory analyses.

A subsequent protocol amendment will reflect these changes. Should there be any further discrepancy between the associated protocol and this SAP, the content of the SAP shall prevail.

3. Statistical Hypotheses

A test of superiority will be used to evaluate two hypotheses for the change in axial length from baseline at the 24-month visit (primary outcome):

- Between spectacles with single vision lenses + nightly placebo eye drops (hereafter **PLACEBO** group) and spectacles with HAL lenses + nightly placebo eye drop (hereafter **HAL** group)
- Between **PLACEBO** group and spectacles with single vision lenses + nightly atropine 0.05% eye drop (hereafter **ATROPINE** group)

Since two treatments with different mechanisms of action are being compared to a shared control group, no adjustment for multiplicity is necessary (Section 16).

3.1. ATROPINE Versus PLACEBO

The 0.05% atropine versus placebo eyedrops hypothesis is evaluated in the cohort using single-vision lenses.

- Null Hypothesis (H_0): There is *no difference* in the primary outcome between the PLACEBO group and the ATROPINE group.
- Alternative Hypothesis (H_a): There *is a nonzero difference* in the primary outcome between the PLACEBO group and the ATROPINE group.

3.2. HAL Versus PLACEBO

The HAL vs. single vision lenses hypothesis is evaluated in the cohort using placebo eye drops.

- Null Hypothesis (H_0): There is *no difference* in the primary outcome between the PLACEBO group and the HAL group.
- Alternative Hypothesis (H_a): There *is a nonzero difference* in the primary outcome between the PLACEBO group and the HAL group.

4. Outcome Measures

4.1. Primary Efficacy Outcome

- Change in axial length from baseline at 24 months

4.2. Secondary Efficacy Outcomes

- Change in spherical equivalent refractive error (SER) from baseline at 24 months
- Change in axial length from baseline at 30 months
- Change in SER from baseline at 30 months
- Change in axial length from baseline at 18 months
- Change in SER from baseline at 18 months
- Change in axial length from baseline at 12 months
- Change in SER from baseline at 12 months
- Change in axial length from baseline at 6 months
- Change in SER from baseline at 6 months

4.3. Exploratory Outcomes

- Change in monocular amplitude of accommodation from baseline at 6, 12, 18, 24, and 30 months
- Change in pupil size from baseline at 6, 12, 18, 24, and 30 months

- Change in axial length over 24 months (area under the curve)
- Change in axial length from 12 to 24 months
- Change in axial length from 24 to 30 months
- Change in SER over 24 months (area under the curve)
- Change in SER from 12 to 24 months
- Change in SER from 24 to 30 months
- Child and parent Treatment Impact Questionnaire scores at 6 months and 24 months.

5. Analysis Cohorts

- Intention-To-Treat (ITT) Analysis Cohort: all randomized participants, irrespective of treatment received, will be analyzed according to treatment assignment.
- Safety Analysis Cohort: participants who receive at least one dose of the randomly assigned study medication (placebo or 0.05% atropine) or wear the randomly assigned spectacles (HAL or SVL) for any amount of time.

The primary analysis will follow the ITT principle. It will include all randomized participants. The data from the ITT cohort will be analyzed according to the group to which the participants were assigned through randomization, regardless of treatment received.

6. Visit Windows

For primary, secondary, exploratory, and subgroup analyses, visits must be completed within the specified visit windows for data to be included. The table below defines the visit windows.

Table 1. Analysis Windows

Visit	Target Day	Target Window	Analysis Window
6 months	Randomization + 183 days	± 2 weeks 169 to 197 days	± 3 months 91 to 273 days
12 months	Randomization + 365 days	± 2 weeks 335 to 379 days	± 3 months 273 to 454 days
18 months	Randomization + 548 days	± 2 weeks 534 to 562 days	± 3 months 454 to 638 days
24 months	Randomization + 731 days	± 2 weeks 717 to 745 days	± 3 months 638 to 821 days
30 months	Randomization + 913 days	± 2 weeks 899 to 927 days	± 3 months 821 to 1003

7. Primary Efficacy Outcome

The average of three separate axial length measurements visits will be calculated for each eye at baseline and all follow-up visits. If fewer than three measurements are available for an eye at a

timepoint, the mean of available measurements will be used to calculate the mean axial length for each eye. The mean of the right and left eyes will be used for analysis. The change in mean axial length from baseline to the 24-month visit will be used as the primary outcome.

The primary analysis will be a treatment group comparison of change in axial length from baseline at 24-month visit, using a longitudinal discrete-time mixed effects model using axial length at randomization, 6, 12, 18, and 24 months as the dependent variable and adjusting for age to account for confounding due to potential imbalances between groups and to increase statistical power.[1,2] Denominator degrees of freedom will be estimated using the Kenward-Roger method (DDFM=KR2 in SAS/STAT version 15.2).[3] Non-independence due to repeated measures on the same participant will be accounted for using an unstructured covariance matrix. The treatment group difference (active treatment – placebo) for change in mean axial length from baseline to 24 months, 95% confidence interval, and P value for the null hypothesis of no difference will be calculated based on the model estimates at 24 months. Within-group summary statistics (mean and standard deviation) will be calculated from observed data.

The model assumptions for the mixed model will be assessed qualitatively without formal statistical testing. The linearity assumption of the continuous baseline covariates (SER and age) will be evaluated using scatterplots (dependent variable versus independent variables and residuals versus fitted values). If the assumption of linearity is seriously violated, then a transformation or median split will be considered. Normality and homoscedasticity will be assessed using plots of residuals (QQ plot, histogram, scatterplot of residuals versus fitted values). With equal sample sizes in each treatment group, the assumption of equal variance is not critical for the treatment group comparison. If assumptions are seriously violated, then alternative, robust approaches will be considered (e.g., M-estimation, generalized linear model using the t distribution, and/or heteroscedasticity consistent standard errors). Linearity and normality are not expected to be violated given prior experience in a recent study.[4]

There will be no explicit imputation of outcome data for exams not completed or completed outside of the analysis window, as the mixed model will produce an unbiased estimate of treatment effect via direct maximum likelihood if the missing outcome data are missing at random (MAR). Intercurrent events will be handled using the treatment policy strategy in which observed data are used regardless of whether the intercurrent event occurs (e.g., death, withdrawal. loss to follow-up, cessation of treatment, receipt of non-randomized treatment, etc.).[5]

7.1. Sensitivity Analysis

Sensitivity analyses will be performed to assess the robustness of the primary outcome.

7.1.1. Complete Case Analysis

A sensitivity analysis will be conducted to compare the mean change in axial length from baseline to 24 months between the treatment groups using an analysis of covariance (ANCOVA) model, adjusting for age and baseline axial length. This analysis will be limited to participants completing the 24-month visit and will not include imputation for missing data.

7.1.2. Tipping Point Analysis

A multiple imputation with a shift parameter that adjusts the imputed values will be performed with treatment group, age, and axial length at randomization, 6, 12, 18, and 24 weeks in the imputation model. The imputed data will be estimated from observations in the same treatment group (i.e., stratified by treatment group)[6]. A shift will be applied to each group and the estimated treatment effect will be displayed as a function of the two shifts. The tipping points where the significance changes direction (from significant to non-significant or *vice versa*) will be reported and clinical judgement will decide if the tipping points are plausible with a clear justification.

7.1.3. Confounding

Imbalances between groups in important covariates are not expected to be of sufficient magnitude to produce confounding. The primary analysis described above includes a pre-specified list of covariates identified in prior work as associated with the outcome. As a sensitivity analysis, any baseline demographic or clinical characteristics observed to be imbalanced between treatment groups will be added as covariates to the analyses of the primary outcome. The determination of a meaningful baseline imbalance will be based on clinical judgement and not a p-value. All variables obtained on a continuous scale will be entered into the model as continuous variables, unless it is determined that a variable does not have a linear relationship with the outcome. In such a case, categorization and/or transformation will be explored.

8. Secondary Efficacy Outcomes

Secondary outcome will use the ITT analysis cohort and test the null hypothesis between treatment groups.

8.1. Change in Spherical Equivalent Error from Baseline at 24 Months

The mean SER of each eye at baseline and all follow-up visits, measured by the masked examiner using cycloplegic autorefraction, will be calculated as the average of the three separate readings from autorefraction. If fewer than three readings are available, the average of available readings will be used. The mean of the right and left eyes will be used for analysis. The other aspects of the analysis are the same as outlined in the primary analyses (Section 5).

8.2. Change in Axial Length from Baseline at 30 Months

The same method described for the primary outcome (Section 5) will be used, but data from 6, 12, 18, 24, and 30 months will be included in the model.

8.3. Change in Spherical Equivalent Refractive Error from Baseline at 30 Months

The same method as described in Section 8.1 will be used but with data from 6, 12, 18, 24, and 30 months included in the model.

8.4. Changes in Axial Length and Spherical Equivalent Refractive Error at 6, 12, and 18 Months

The same methods described in Section 5 and Section 8.1 will be used.

9. Exploratory Outcomes

Exploratory outcomes will use the ITT cohort and test the null hypothesis of no difference between treatment groups.

9.1. Change in Monocular Amplitude of Accommodation at 6, 12, 18, 24, and 30 Months

Change in the monocular amplitude of accommodation at 6, 12, and 24 months will be analyzed using a discrete-time longitudinal mixed effects model adjusted for age similar to the primary outcome (Section 5).

9.2. Change in Pupil Size at 6, 12, 18, 24, and 30 Months

The change in pupil size at 6, 12, and 24 months will be analyzed using a discrete-time longitudinal mixed effects model adjusted for age similar to the primary outcome (Section 7).

9.3. Change in Axial Length Over 24 Months (Area Under the Curve)

The change in axial length over 24 months (area under the curve) will be calculated and compared between treatment groups using the same discrete-time longitudinal mixed effects model used for the analyses of the primary outcome (Section 5). The area under the curve can be interpreted as a weighted average of the change in axial length at each visit with weights proportional to the time between visits. AUC will be calculated by linear combination of model estimates using the trapezoidal rule and the following formula:

$$AUC = \sum_{i=1}^{n} \left(\frac{X_i + X_{i+1}}{2} \times m \right)$$

Where X_i is the axial length measured at the i^{th} visit, m is the number of months between visits i and $i+1$, and n is the number of outcome visits included in the analysis. In this analysis there are $n = 5$ visits total: 0, 6, 12, 18, and 24 months. For presentation, AUC will be divided by the number of months between baseline and the n^{th} visit (i.e., 24) so that the value shown will have units of millimeters rather than millimeter·months.

9.4. Change in Axial Length from 12 to 24 Months

The change in axial length from 12 to 24 months will be calculated and compared between treatment groups using the same discrete-time longitudinal mixed effects model used for the analyses of the primary outcome (Section 7).

9.5. Change in Axial Length from 24 to 30 Months

The change in axial length from 24 to 30 months will be calculated and compared between treatment groups using the same discrete-time longitudinal mixed effects model used for the analyses of the secondary outcome of change axial length from 24 to 30 months (Section 8.2).

9.6. Change in Spherical Equivalent Refractive Error Over 24 Months

The change in SER over 24 months will be analyzed using the same methods as axial length (Section 9.3).

9.7. Change in Spherical Equivalent Refractive Error from 12 to 24 Months

The change in SER from 12 to 24 months will be calculated and compared between treatment groups using the same discrete-time longitudinal mixed effects model used for the analyses of the secondary outcome in SER (Section 8.1).

9.8. Change in Spherical Equivalent Refractive Error from 24 to 30 Months

The change in SER from 24 to 30 months will be calculated and compared between treatment groups using the same discrete-time longitudinal mixed effects model used for the analyses of the secondary outcome in SER (Section 8.3).

9.9. Treatment Impact Questionnaire

The Treatment Impact Questionnaire (TIQ) will be used as a quantitative measure to evaluate opinions regarding the burdens and impact of the randomized treatment at 6 months and 24 months (as questions for the child – the Child TIQ and the parent themselves – the Parent TIQ).

The Child-TIQ and Parent-TIQ will undergo separate factor analysis to determine the number of domains for each TIQ. Each domain will be refined through the evaluation of misfitting items and will then be Rasch scored.

The Rasch scores will be compared between the two treatment groups using a t test to generate a mean difference and 95% CI. If assumptions of the t test are seriously violated (normality and homoscedasticity) then the Wilcoxon Rank-Sum test with Hodges-Lehmann estimator and 95% confidence interval will be used.

Note that because the TIQ is not administered at baseline (because treatment has not been started), there will be no adjustment for baseline score in any analysis.

10.Safety Analyses

Safety analyses will be performed among participants who receive at least one dose of their randomly assigned study medication (placebo or 0.05% atropine) or wear their randomly assigned spectacles (HAL or SVL) for any amount of time. Adverse events will be coded and tabulated based on the Medical Dictionary of Regulatory Activities (MedDRA) by treatment

group. The severity, frequency, and relationship to study treatment will also be tabulated. There will be no formal statistical comparison of adverse events.

The number and proportion of participants experiencing the following outcomes at any time post randomization will be tabulated for each group and compared using the Barnard Unconditional Exact Test; risk differences and 95% CIs will be estimated using the exact Mid-P method of Agresti and Min.[7]:

- Loss of ≥ 2 logMAR lines of binocular near visual acuity
- Loss of ≥ 2 logMAR lines of monocular distance visual acuity

These analyses will include all randomized participants without regard to visit completion because loss to follow-up is expected to be low based on prior experience in a recent RCT.[4]

11.Intervention Adherence

Adherence to study eyedrops (atropine and placebo eyedrop) and spectacles (SVL and HAL lenses) based on calendars brought to each follow-up visit will be tabulated in each treatment group.

12.Protocol Adherence and Retention

Protocol deviations and visit completion rates (excluding participant deaths) will be tabulated for each treatment group.

13.Baseline Descriptive Statistics

The baseline characteristics will be tabulated according to treatment group. At a minimum, the following will be included:

- Age
- Sex
- Race
- Ethnicity
- Iris color
- Number of biological parents with myopia
- Distance visual acuity in habitual refractive correction
- Axial length
- SER

14.Planned Interim Analyses

There are no formal planned interim analyses for this study. The Data and Safety Monitoring Committee will review safety and efficacy data approximately every 6 months and can recommend stopping the trial if deemed necessary.

15.Subgroup Analyses

Subgroup analyses, i.e., assessments of effect modification (interaction), will be conducted for the primary outcome, change in axial length at 24 months, and the key secondary outcome, change in SER at 24 months. These analyses will be considered exploratory. Subgroup analyses will be interpreted with caution, particularly if the corresponding overall analysis does not demonstrate a significant treatment group difference. The general approach for these analyses will be to add an interaction term for the subgroup factor by treatment into the analysis models described in Section 7 and Section 8.1. The P value for an interaction effect will be shown only if there are a minimum of 10 observations per level and treatment group. Statistical power for detecting interactions is expected to be low. Within-subgroup means, standard deviations, adjusted treatment differences, and 95% confidence intervals will be calculated for the following factors:

- Sex
- Race/Ethnicity
- Iris color
    - Brown vs not brown
- Age
    - 5 to <9 vs 9 to <12
- Axial Length
- SER

For continuous factors, the interaction P value will be calculated using the continuous version and within-subgroup means, standard deviations, differences, and 95% confidence intervals will be calculated based on a median split.

16.Multiple Comparison/Multiplicity

For the primary outcome of axial length, two tests of superiority will be conducted: ATROPINE vs PLACEBO and HAL vs PLACEBO. The tests will be performed independently, and each will be conducted with an alpha level of 0.05.

Although two pairwise comparisons are being evaluated, there will be no formal adjustment to the familywise error rate. Because the primary objective of this trial is to compare each of two active treatments (atropine eye drops and HAL lenses), which likely have different mechanisms of action, with a shared PLACEBO control group (not with one another), a multiplicity adjustment is not needed.[8-10] The risk of a false positive finding with this approach is lower than if the two hypotheses were evaluated in two studies with different control groups.[8] The same logic applies to secondary, exploratory, safety, and subgroup analyses.

For the secondary outcomes (Section 8), the familywise error rate will be controlled with a hierarchical (i.e., fixed sequence) approach. If the null hypothesis for the primary outcome (axial length) is rejected (for either HAL vs PLACEBO or ATROPINE vs PLACBO), then the first secondary outcome (change in SER at 24 months) will be compared without further adjustment to the type 1 error rate.[11] If the primary outcome null hypothesis is not rejected, then the

comparison of the change in SER at 24 months will be considered exploratory; a 95% confidence interval (without adjustment for multiplicity) will be presented, and a *p*-value will not be presented. Subsequent secondary outcomes will be tested in the order listed in Section 2.

For exploratory outcomes, the adaptive two-stage step-up procedure of Benjamini, Krieger, and Yekutieli[12] will be used to control the false discovery rate at 5% by adjusting both the 95% confidence intervals and P values for the analysis of multiple outcomes. The categories/families for FDR adjustment will be as follows:

- Pupil size and accommodation
- SER and axial length
- Treatment Impact Questionnaire

There will be no formal adjustment for safety analyses because type 2 errors (false negatives) are of greater concern than type 1 errors (false positives).

The adaptive two-stage step up procedure[12] will be used to control the false discovery rate at 5% to adjust for multiple subgroup analyses. Both interaction P values and within-group 95% confidence intervals will be adjusted; interaction P values and within-group 95% CIs will be considered separate families of tests. P values for interactions will only be presented if the overall analysis indicates a significant effect.

17.Missing Data

In general, the procedure for handling missing data is outlined in each section. Where not otherwise specified, missing data will be excluded, and only complete cases will be analyzed.

18.Additional Tabulations and Analyses

- A flow chart accounting for all participants for all visits and phone calls will be developed.
- Visit and phone contact completion rates for each follow-up visit will be tabulated.
- Proportion of participants with a change in myopia of ≥ 0.50 D, ≥ 0.75 D, and ≥ 1.0 D from baseline to 12, 24, and 30 months.
- Proportion of participants with a change in axial length ≥ 0.25 mm, ≥ 0.375 mm, and ≥ 0.50 mm from baseline to 12, 24, and 30 months.

19.Exploratory Analyses in COMBINED Atropine + HAL Lenses Group

Exploratory comparisons between the COMBINED group and the ATROPINE, HAL, and PLACEBO groups will parallel the analyses conducted for the ATROPINE vs PLACEBO and HAL vs PLACEBO comparisons described above. There comparisons will inform design of future studies.

19.1. Analysis of Main Effects

An exploratory analysis of change in axial length from baseline to 24 months (Section 7) will be repeated by pooling across cells in the factorial design (i.e., testing main effects) if the estimated interaction effect is less than the hypothesized mean difference of 0.17 mm based on the 95% confidence interval (e.g., the 95% CI excludes +/- 0.17 mm). The interaction effect will be estimated from a longitudinal mixed effects linear model with independent variables for baseline age, atropine (yes vs no), HAL (yes vs no), and the interaction of atropine vs HAL. The interaction effect will be tested based on the estimated marginal means as follows:

$$-0.17\ mm < \bar{x}_{PLACEBO+SVL} - \bar{x}_{PLACEBO+HAL} - \bar{x}_{ATROPINE+SVL} + \bar{x}_{ATROPINE+HAL} < 0.17\ mm$$

If the interaction effect is less than the hypothesized mean difference and there is a significant main effect for either atropine or HAL (P<.05) in the above analysis of change in axial length, then a similar exploratory analysis will be conducted for change in SER from baseline to 24 months (Section 8.1) but without testing for interaction. The study is expected to have greater power to detect differences in axial length than SER; additionally, SER and axial length are expected to be highly corrected. Therefore, the test for interaction will be more reliable in the axial length analysis.

20.Example Analysis Code

The code in the following sections is an example of how analyses may be performed in SAS/STAT version 15.2 or a comparable statistical package.

20.1. Primary Analysis

```
proc sort data=visitPts;
      by PtID month;
run;

proc mixed data=visitPts plots=(VCIRYPANEL);
      where month IN (0, 6, 12, 18, 24);
      class PtID TrtGroup month;
      model meanSER = TrtGroup|month ageRand|month / s ddfm=kr2;
      repeated month / type=un subject=PtID r rcorr;
      slice TrtGroup*month / diff cl sliceby=month plots=none;
run; title;
```

20.2. Area Under the Curve Analysis

```
proc sort data=visitPts;
      by PtID month;
run;

proc mixed data=visitPts plots=(VCIRYPANEL);
      where month IN (0, 6, 12, 18, 24);
```

```
        class PtID TrtGroup month Gender;
        model meanSER = TrtGroup|month ageRand|month / s ddfm=kr2;
        repeated month / type=un subject=PtID r rcorr;
        lsmeans TrtGroup*month;
        lsmestimate TrtGroup*month "24-m AUC Atropine"
              3 6 6 6 3 0 0 0 0 0 / divisor = 24 cl;
        lsmestimate TrtGroup*month "24-m AUC Placebo"
              0 0 0 0 0 3 6 6 6 3 / divisor = 24 cl;
        lsmestimate TrtGroup*month "24-m AUC Difference (Atropine - Placebo)"
              3 6 6 6 3 -3 -6 -6 -6 -3 / divisor = 24 cl;
run; title;
```

20.3. Subgroup Analysis

```
proc sort data=visitPts;
      by PtID month;
run;

proc mixed data=visitPts plots=(VCIRYPANEL);
      where month IN (0, 6, 12, 18, 24);
      class PtID TrtGroup month Gender;
      model meanSER = TrtGroup|month ageRand|month Gender|TrtGroup|month / s
ddfm=kr2;
      repeated month / type=un subject=PtID r rcorr;
      slice TrtGroup*month*Gender / diff cl sliceby=month*Gender plots=none;
      lsmestimate trtGroup*month*Gender "Difference of Differences (Interaction)"
            0 0 0 0 0 0 0 0 1 -1  0 0 0 0 0 0 0 0 -1 1;
run; title;
```

21. References

1. Weise KK, Repka MX, Zhu Y, et al. Baseline factors associated with myopia progression and axial elongation over 30 months in children 5 to 12 years of age. *Optom Vis Sci*. Oct 1 2024;101(10):619-626. doi:10.1097/OPX.0000000000002187
2. Food and Drug Administration. Adjusting for covariates in randomized clinical trials for drugs and biological products: guidance for industry. US Department of Health and Human Services; 2023.
3. Kenward MG, Roger JH. An Improved Approximation to the Precision of Fixed Effects from Restricted Maximum Likelihood. *Computational Statistics and Data Analysis*. 2009;(53):2583-2595.
4. Repka MX, Weise KK, Chandler DL, et al. Low-dose 0.01% atropine eye drops vs placebo for myopia control: A randomized clinical trial. *JAMA Ophthalmol*. Aug 1 2023;141(8):756-765. doi:10.1001/jamaophthalmol.2023.2855
5. Food and Drug Administration. E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials. 2021.
6. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? *Stat Methods Med Res*. Sep 2018;27(9):2610-2626. doi:10.1177/0962280216683570
7. Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. *Stat Methods Med Res*. Apr 2015;24(2):224-54. doi:10.1177/0962280211415469
8. Howard DR, Brown JM, Todd S, Gregory WM. Recommendations on multiple testing adjustment in multi-arm trials with a shared control group. *Stat Methods Med Res*. May 2018;27(5):1513-1530. doi:10.1177/0962280216664759
9. Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of Multi-Arm Parallel-Group Randomized Trials: Extension of the CONSORT 2010 Statement. *Jama*. Apr 23 2019;321(16):1610-1620. doi:10.1001/jama.2019.3087
10. Parker RA, Weir CJ. Non-adjustment for multiple testing in multi-arm trials of distinct treatments: Rationale and justification. *Clin Trials*. Oct 2020;17(5):562-566. doi:10.1177/1740774520941419
11. Dmitrienko A, D'Agostino RB, Sr. Multiplicity Considerations in Clinical Trials. *N Engl J Med*. May 31 2018;378(22):2115-2122. doi:10.1056/NEJMra1709701
12. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*. 2006;93(3):491-507. doi:10.1093/biomet/93.3.491