FDA Presentation
sNDA 21641(013):Rasagiline Delayed Start Trials in Parkinson's Disease

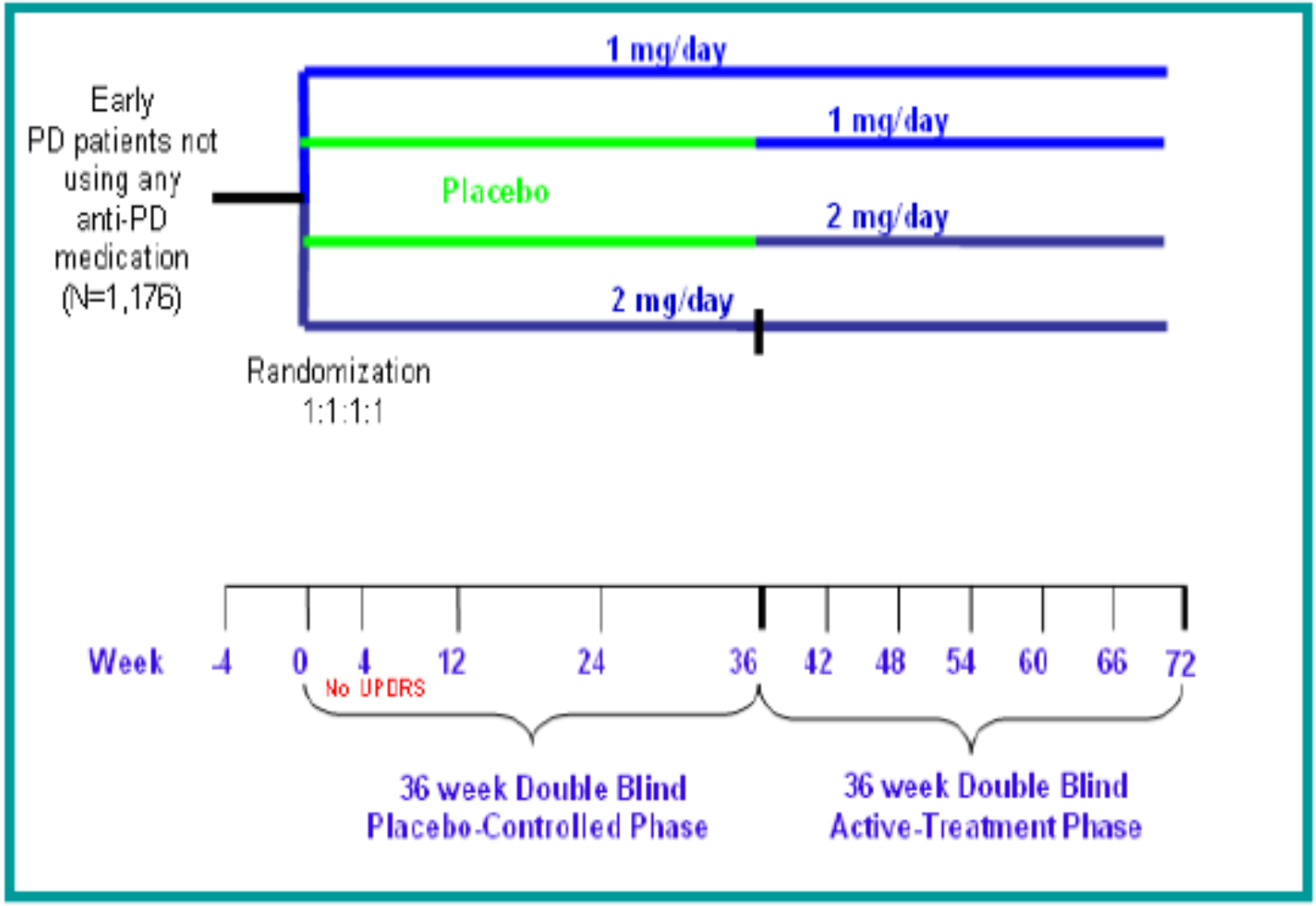
Peripheral and Central Nervous System Drugs Advisory Committee Meeting, October 17, 2011

Tristan Massie, Ph.D.
Mathematical Statistician
Division of Biostatistics I (DB-I)
Office of Biostatistics (OB)
Office of Translational Science (OTS), CDER, FDA

ADAGIO Issues

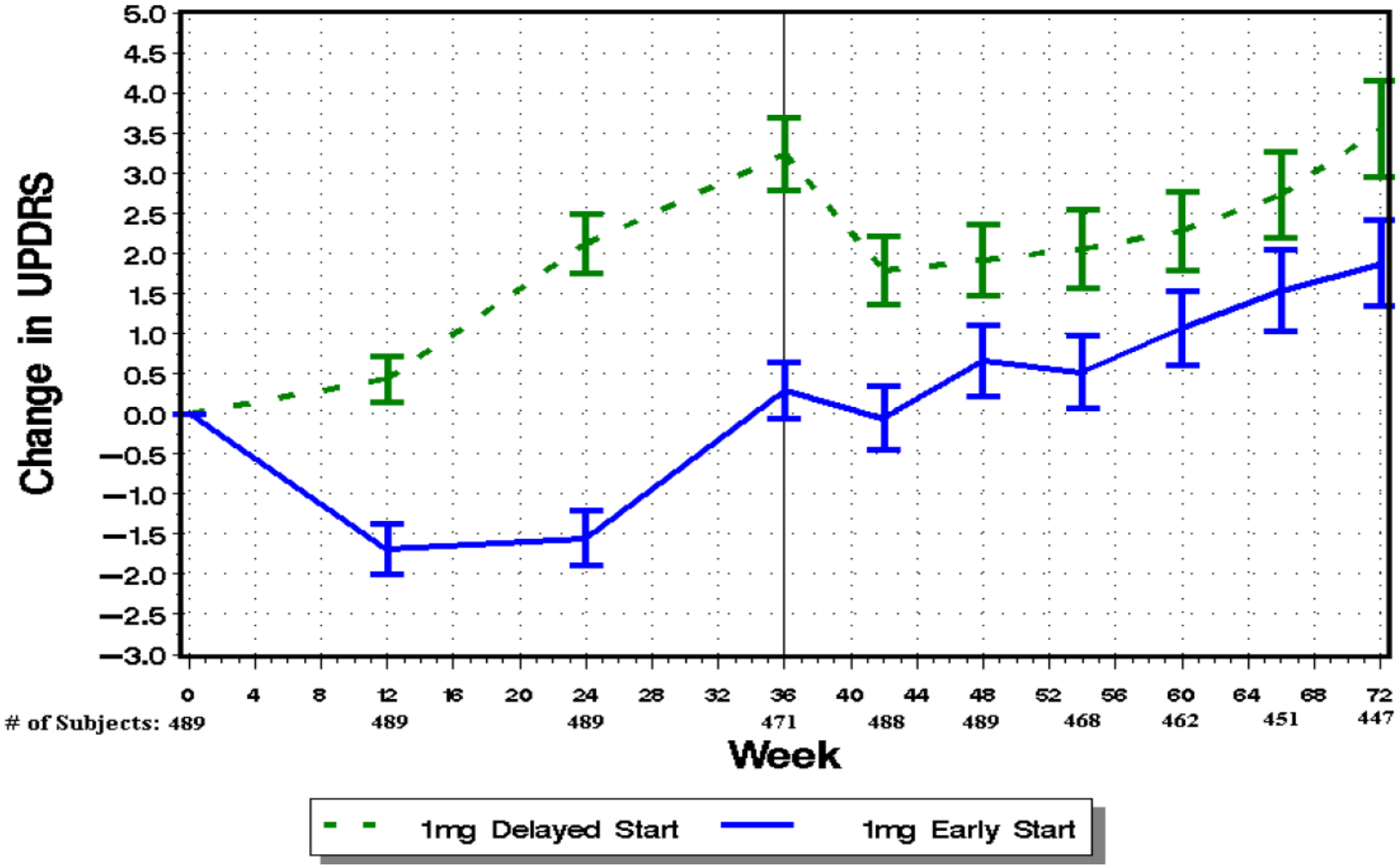
- 2 mg dose **clearly** not positive at end, Week 72
- Baseline UPDRS score > 25.5 quartile subgroup analysis
  - Unplanned, inconclusive
- Failure of higher dose raises concerns that 1 mg separate dataset result may be a false positive
  - Several other issues with 1 mg results

ADAGIO Study Design



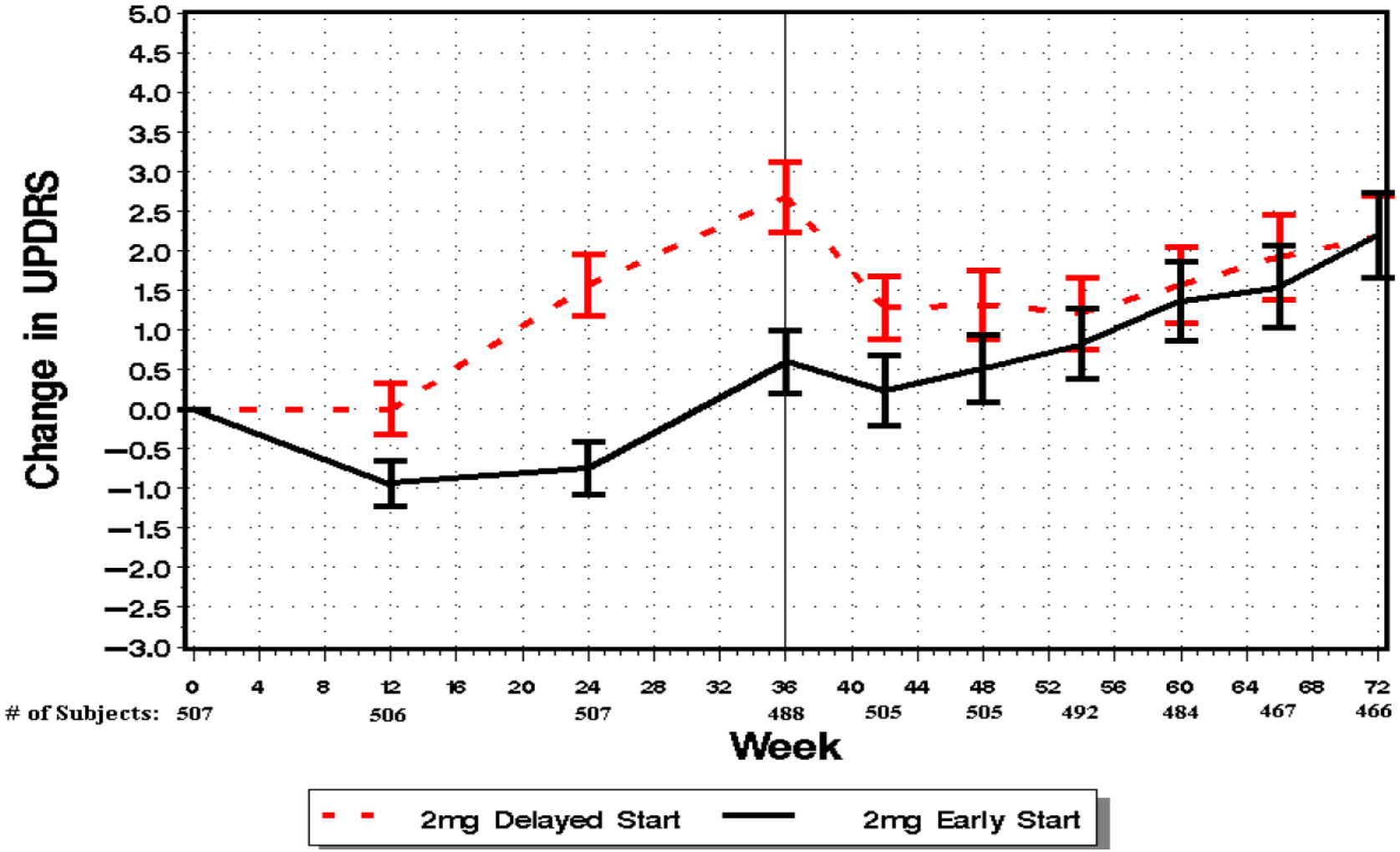
From NDA Study Report ADAGIO Study 500, page 58

ADAGIO 1 mg



From Teva Backgrounder, Figure 8, Page 68

ADAGIO 2 mg



From Teva Backgrounder, Figure 9, Page 69

Hypothesis 1:

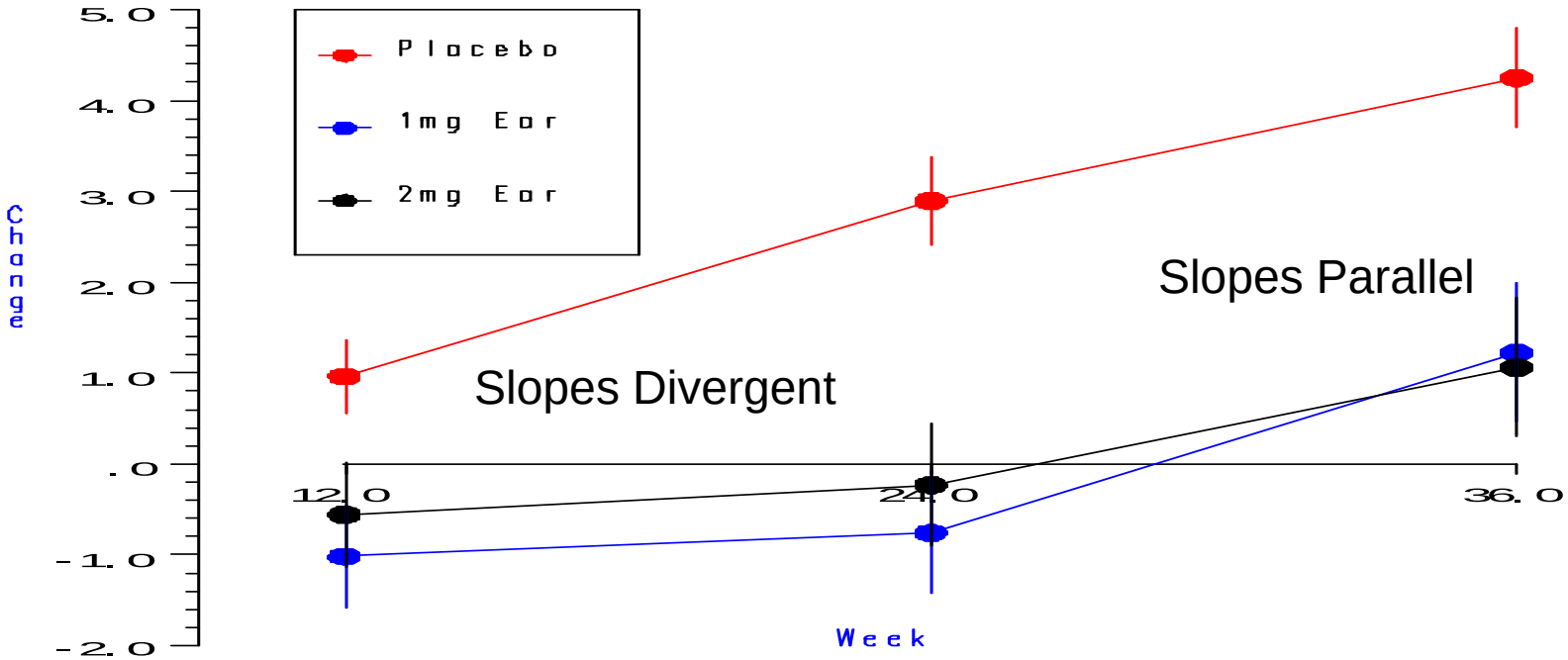
- Slope difference of:

  1 mg Early Rasagiline vs. Placebo (pooled delayed groups)

  2 mg Early Rasagiline vs. Placebo (pooled delayed groups)

Over Placebo Controlled Phase Weeks 12 through 36

ADAGIO Hypothesis 1: Slope Difference in Placebo Phase



- Slope Difference: p=0.0133 for 1 mg , p=0.0001 for 2 mg significant
- BUT failed pre-specified non-linearity (constant slope over time) test
- Therefore, Hypothesis 1 slope difference test inconclusive
- Two available time segments suggest conflicting conclusions

Adjustment for Multiple Dose Testing

- Two doses create two chances to win
- Need for adjustment to significance level to control overall study false positive rate to be comparable with single dose study

Sponsor's Chosen Method = **Hochberg**:

If both doses' p-values $\leq$ 0.050 conclude both doses statistically significant

Otherwise, if larger of two p > 0.050 then conclude smaller p statistically significant if it's p<0.025

Hypothesis 2:

- "Superiority" at the end of Active Phase, Week 72, of Early Start (ES) over Delayed Start (DS)
  - Week 72 Mean Change from Baseline in Total UPDRS
- This test was only to be performed if Hypothesis 1 was statistically significant
  - Placebo Controlled Phase Tests and Active Phase Tests were hierarchically ordered to control overall trial false positive rate
- Prespecified Dataset with all 4 Groups for Primary Analysis
  - Joint 4 group model to compare: 1 mg Early vs. 1 mg Delayed
    2 mg Early vs. 2 mg Delayed

ACTE Dataset: Primary Analysis Dataset for Hypothesis 2 in Active Phase

- Active Efficacy Data Analysis Set (ACTE) consists of all subjects entering the active-treatment phase of the trial with at least 24 weeks of treatment during the placebo-controlled phase of ADAGIO

  <u>AND</u>

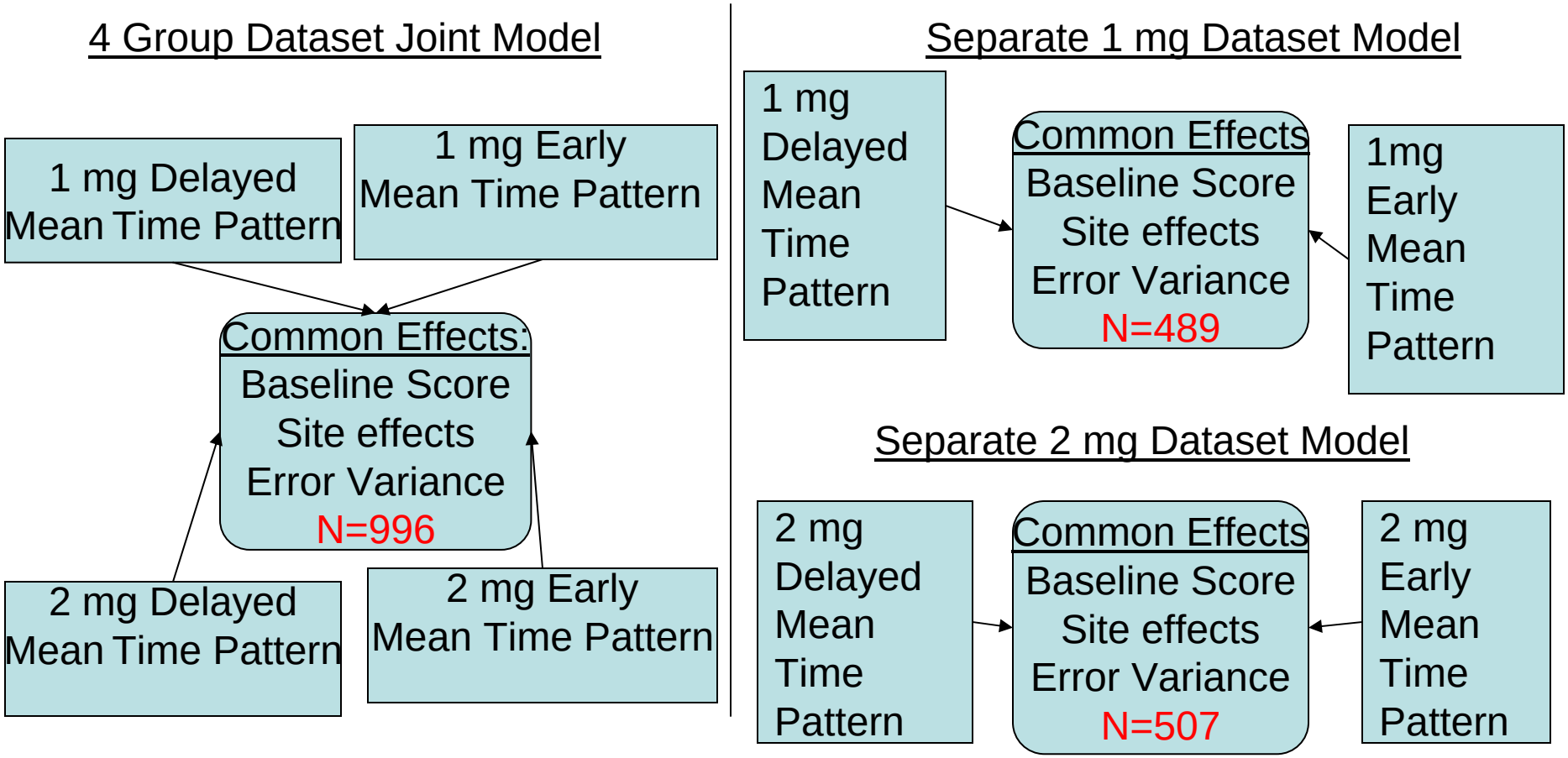
  at least one Total UPDRS measurement during the active-treatment phase from Week 48 or later

- If in Phase 1 the investigator determined patient needed additional anti-Parkinson's treatment, the patient could transition early to the active phase
  - These patients excluded from analysis if transition before week 24
  - This presents a challenge for obtaining an unbiased analysis with this design

Concerns About Post Hoc Revision of Primary Analysis Not Prespecified in Statistical Analysis Plan (SAP)

- Final SAP planned to analyze Hypothesis # 2 (and # 3) using combined datasets (1 mg ES, 1 mg DS, 2 mg ES, 2 mg DS) in statistical model for primary analysis
- No prespecified SAP provision for testing for interactions or if assumptions met
- No prespecified SAP alternative analyses if interaction(s) found or assumptions not met
- Sponsor conducted post hoc revised primary analysis for Hypothesis # 2 from using combined dataset in statistical model to using separate datasets in the statistical model because sponsor found interactions for treatment by baseline Total UPDRS and treatment by site
- There are concerns about this not specified, post-hoc revision of the primary analysis
  - Uncertainty if primary analysis should be revised and if so, how?

Why Four Group "Combined" Dataset Result and Separate Dose Dataset Results Can Differ



Group Unique Effect – Pattern of Treatment Group Mean over Time

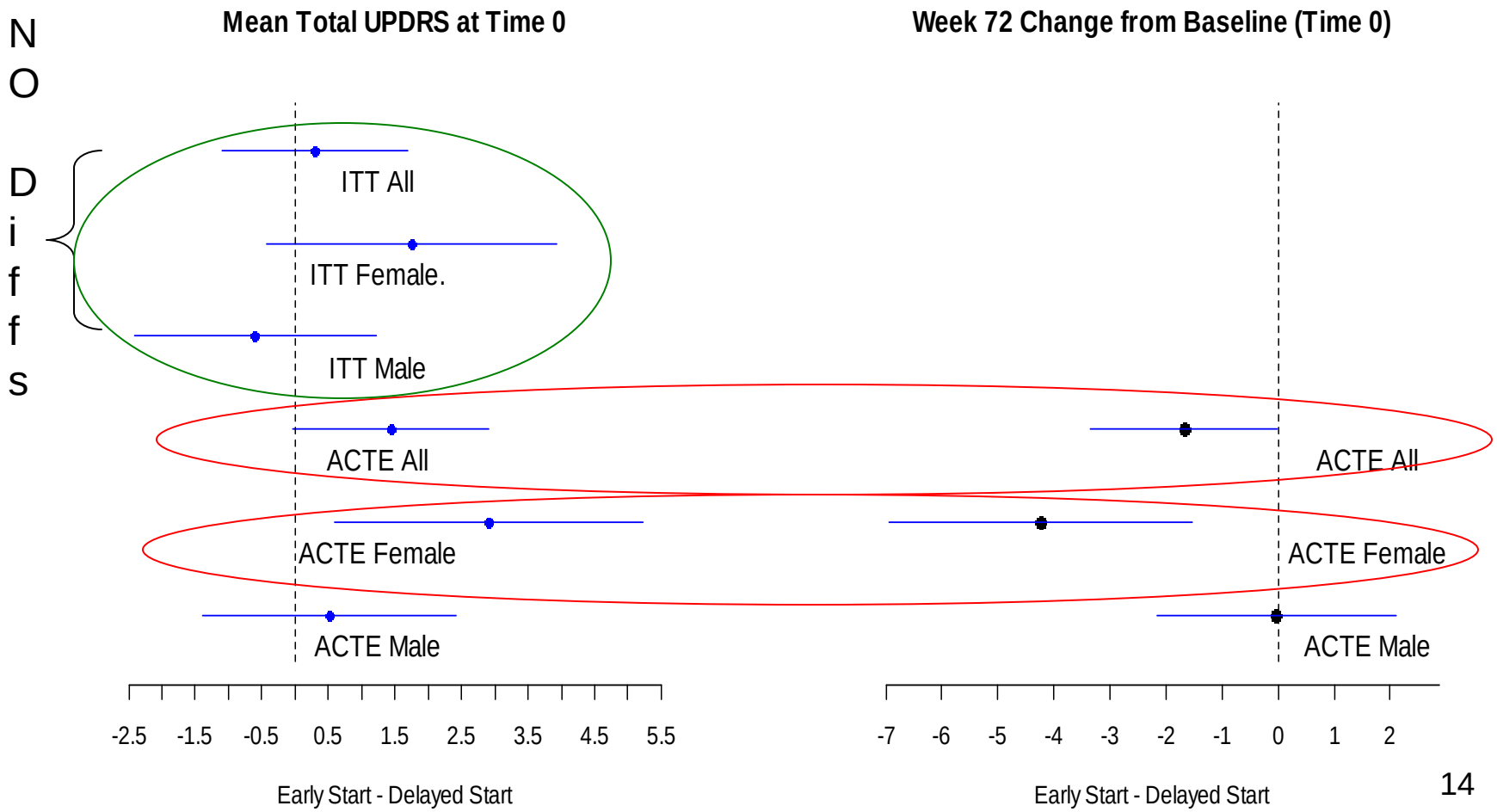
General Effects - Common to All Groups: Sites, Baseline Score, Variability

Post Hoc Hypothesis 2 Analysis Dataset Change

- **Prespecified analysis using Joint 4-group statistical model**
  - 1 mg Early vs. Delayed: -1.42, p=0.0506 (need ≤ 0.0250)
  - 2 mg Early vs. Delayed: +0.18, p=0.8014
- **Post-hoc analysis Separate 2-group model for each dose**
  - 1 mg Early vs. Delayed: -1.68, p=0.0250
  - 2 mg Early vs. Delayed: +0.36, p=0.6028

- Sponsor's Justification for change: Significant Interaction effects
  - Treatment by baseline Total UPDRS score interaction between doses (1 mg vs 2 mg)
  - Treatment by site interaction between doses (1 mg vs. 2 mg)
  - However, significantly different baseline score interaction and treatment by site interaction within dose (Early vs. Delayed) persists. Therefore, switching primary analysis is questionable.

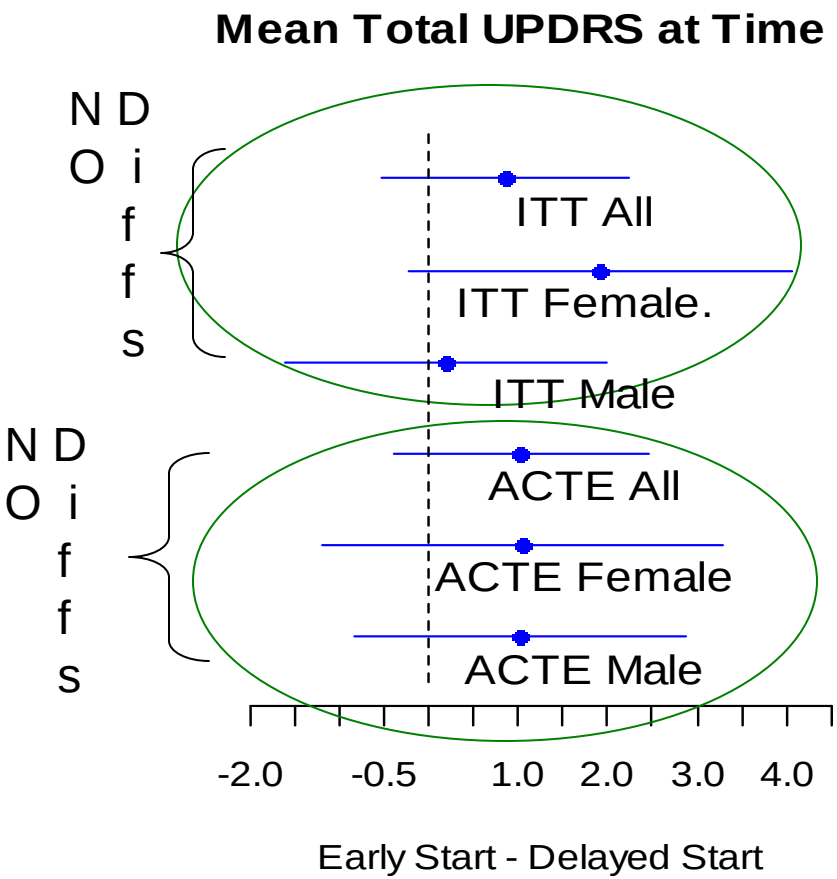
Differences in Total UPDRS for 1 mg at Time 0 and Week 72 in ADAGIO

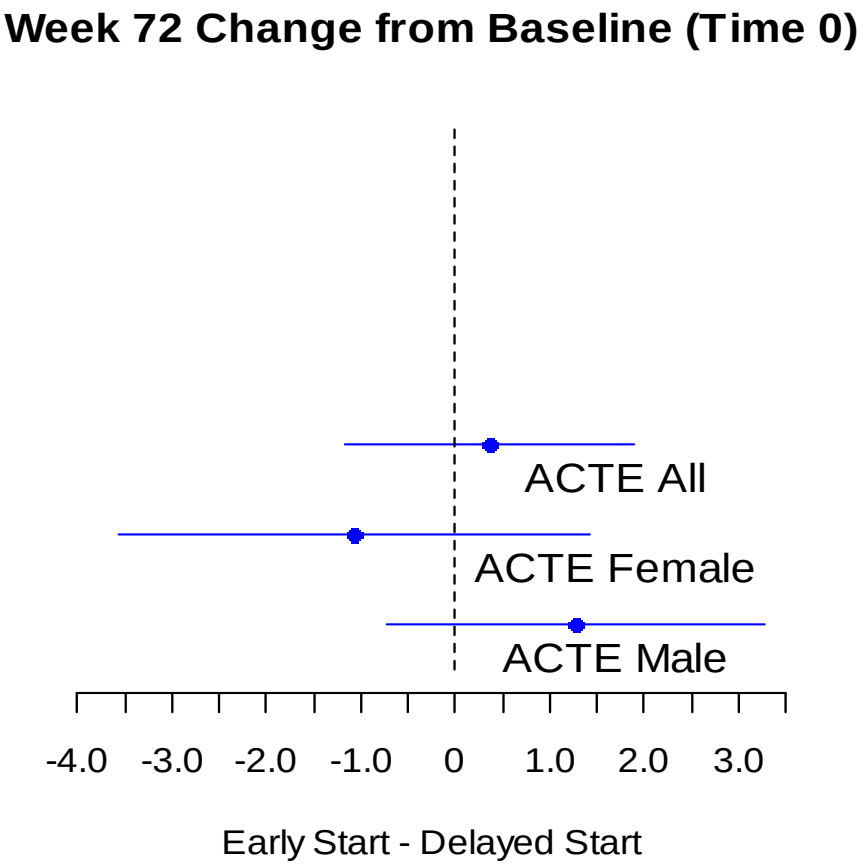


ADAGIO Female Subgroup (1 mg) with Large UPDRS Treatment Difference (Week 72) Also Has Baseline UPDRS Imbalance

- Dropouts lead to NON-RANDOMIZED treatment groups within ACTE population
  - Unable to know if baseline imbalance causes treatment difference BUT baseline imbalance raises concerns for other potential imbalances (measured or unmeasured)
  - Baseline score adjustment in model not perfect and cannot correct/adjust for other imbalances
  - For Hypotheses 2 and 3, 16% of ITT patients totally excluded from primary analyses
    - Equivalent to 16% of ITT population having no post-baseline data in ordinary single phase designed study
    - Assumption should be data missing completely AT RANDOM but we know NOT RANDOM; some patients RESCUED by early transition to Phase 2
    - Hypotheses 2 and 3 results may not be valid

Differences in Total UPDRS for 2 mg at Time 0 and Week 72 in ADAGIO





Sponsor's Original Hypothesis 2:
Weeks 48 through 72 Average Group Mean Difference

Dose	Difference Early-Delayed	Std. Err.	P-value
1 mg	-1.408	0.557	0.0115
2 mg	-0.272	0.544	0.6164

– Sponsor suggests 2 mg failure on Week 72 outcome due to underpowering

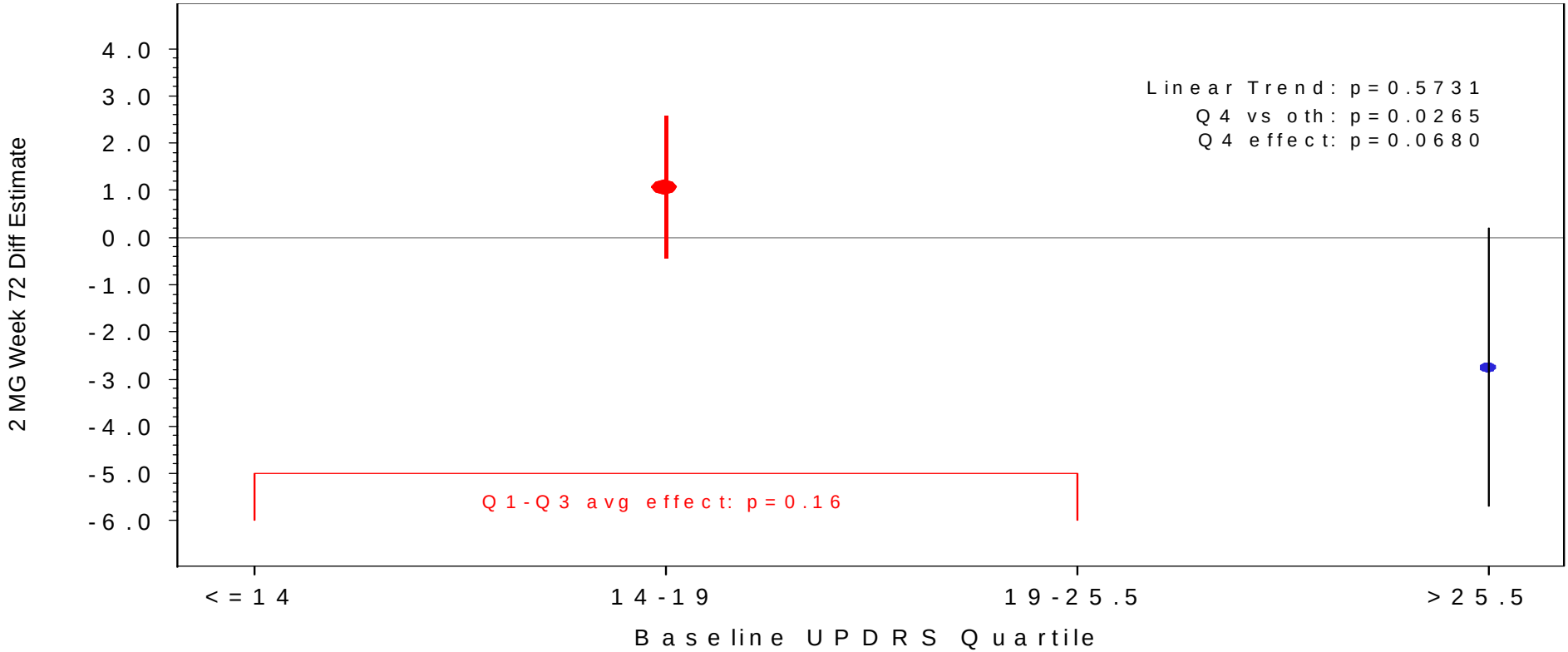
– Size of 2 mg effect in sponsor's original analysis not consistent with a slight underpowering issue

Sponsor's Explanation for Failure of ADAGIO 2 mg Efficacy

- In NDA, Sponsor presented many post-hoc analyses by baseline UPDRS Total quartiles (ITT) attempting to explain failure of high dose to show any benefit at end of study (72 weeks).

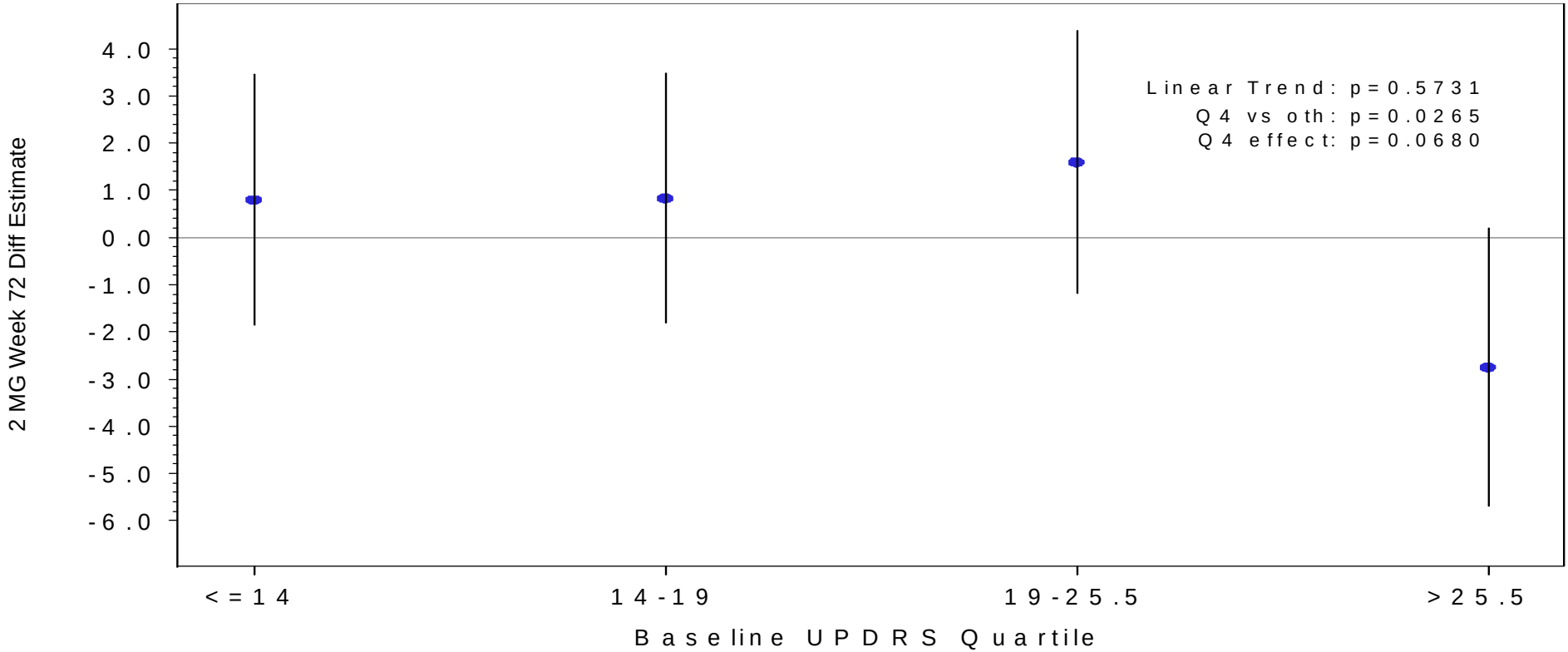
- Sponsor hypothesized lack of effect/difference for 2 mg may have been due to "floor"/"threshold" effects in which patients with more severe PD (i.e., higher baseline Total UPDRS) show greater responses.

ADAGIO: 2 mg Week 72 Difference for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ITT Population)



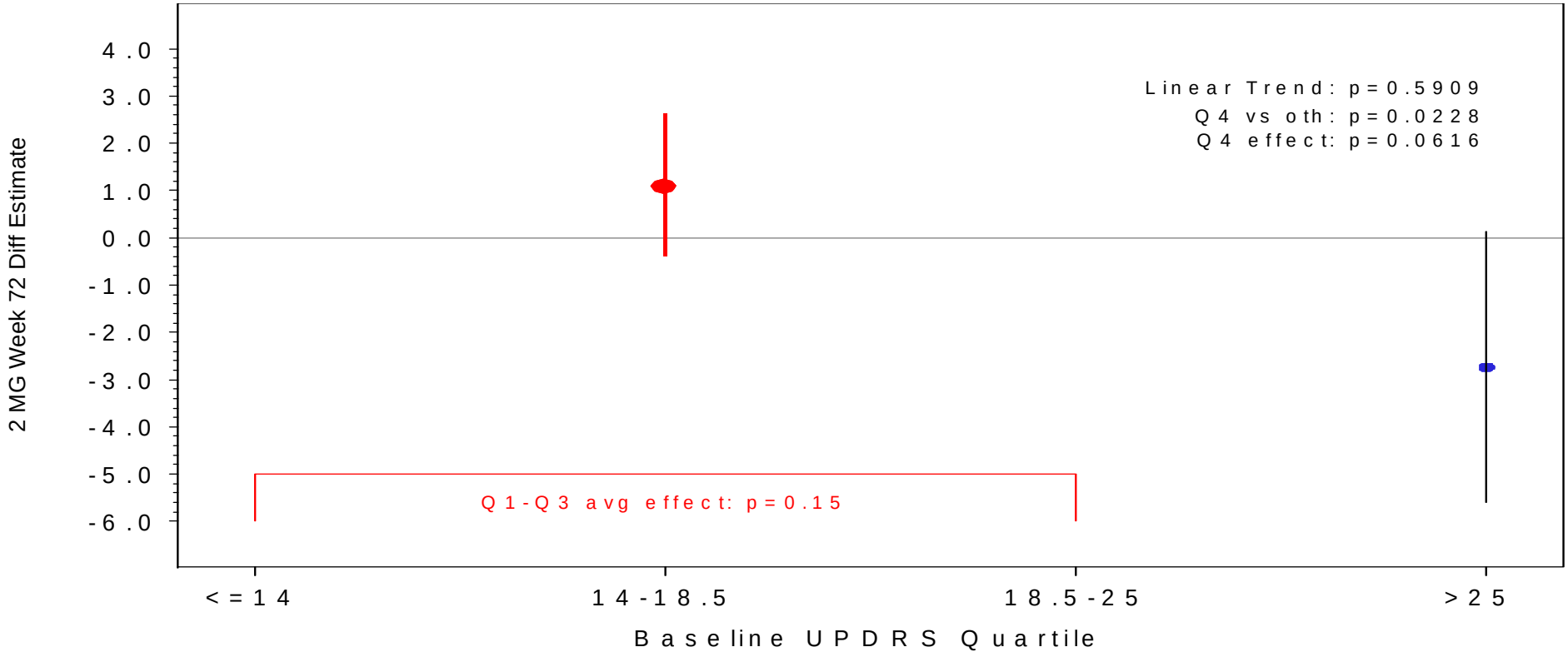
1. Quartiles are arbitrary split points (why 75%? why not use 66% or 80%?)
2. Not true quartiles for ACTE population because of dropouts

ADAGIO: 2 mg Week 72 Difference for ACTE Population by Baseline UPDRS Quartiles (Quartiles derived from ITT Population)

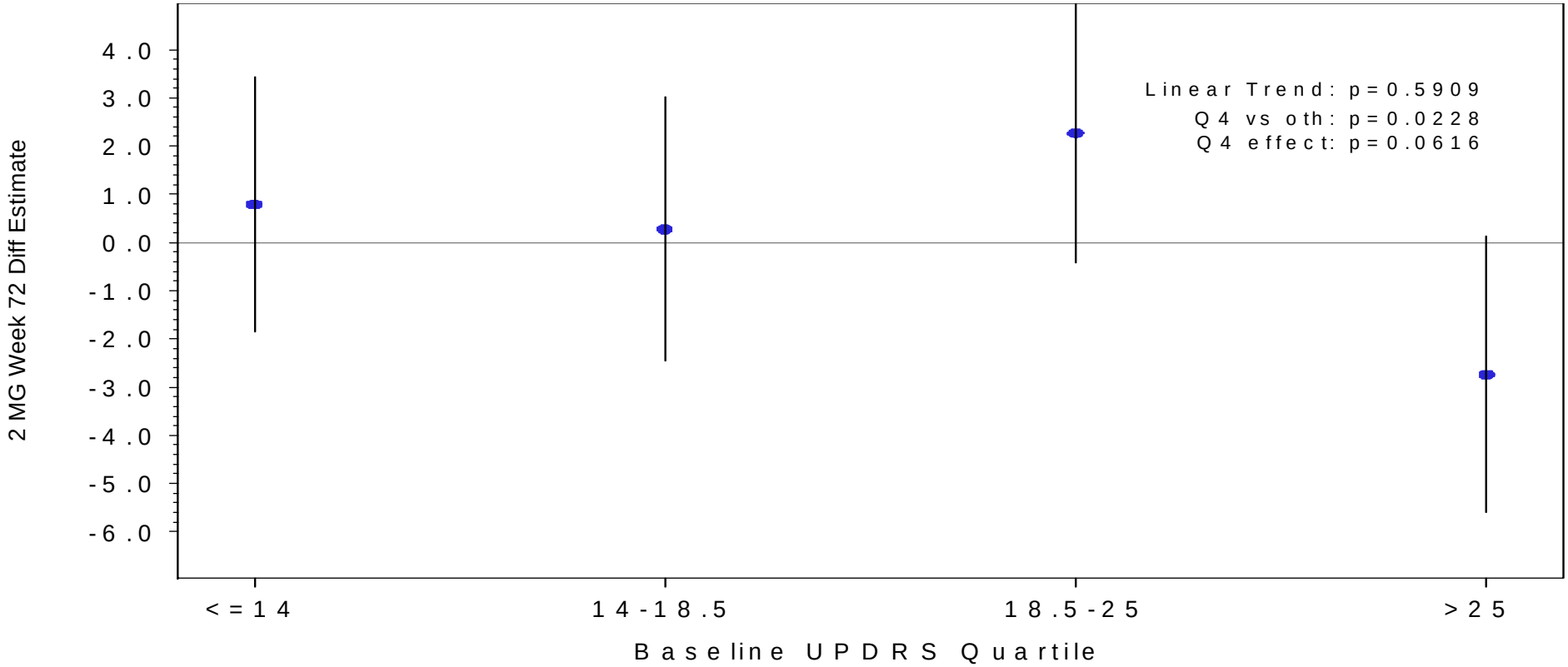


1. Quartiles are arbitrary split points (why 75%? why not use 66% or 80%?)
2. Not true quartiles for ACTE population because of dropouts

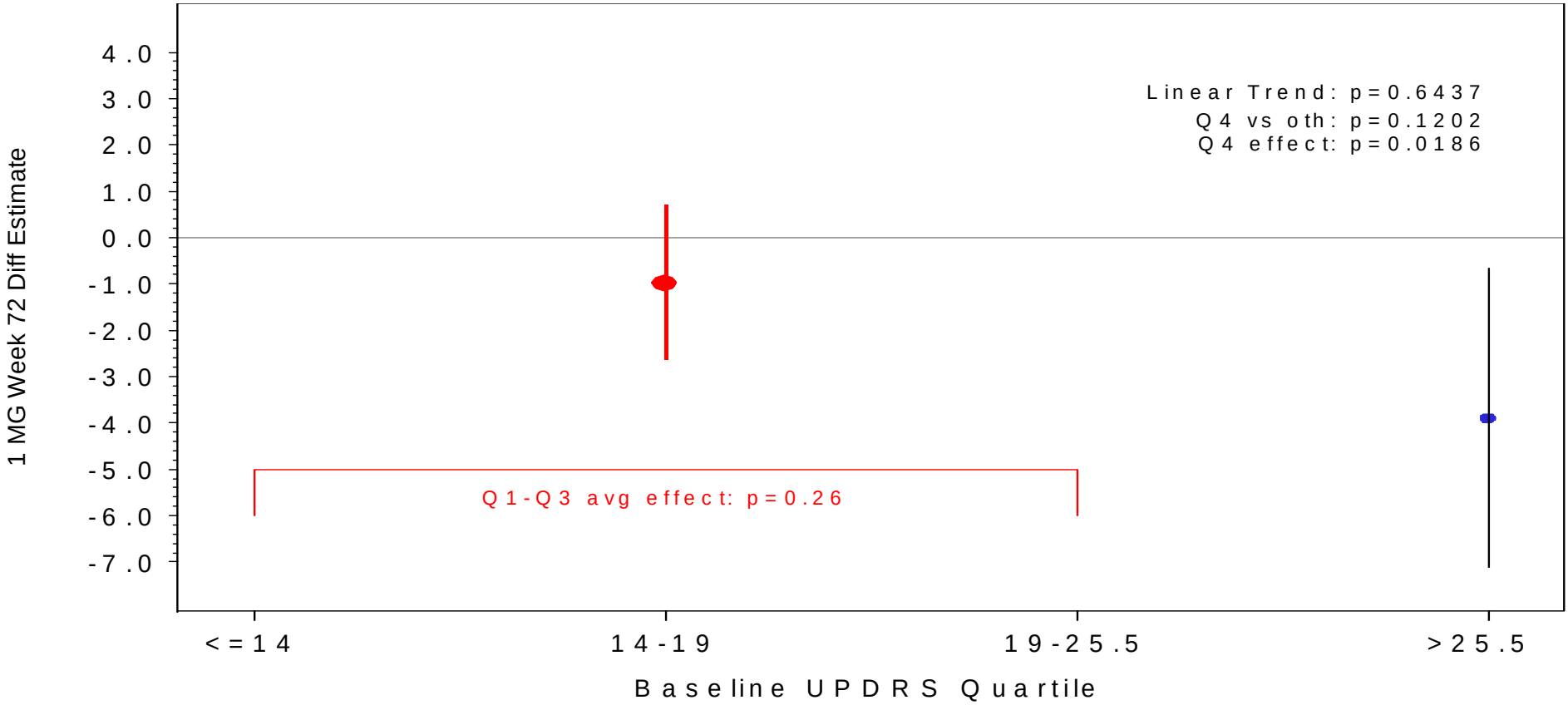
ADAGIO: 2 mg Week 72 Difference for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ACTE Population)



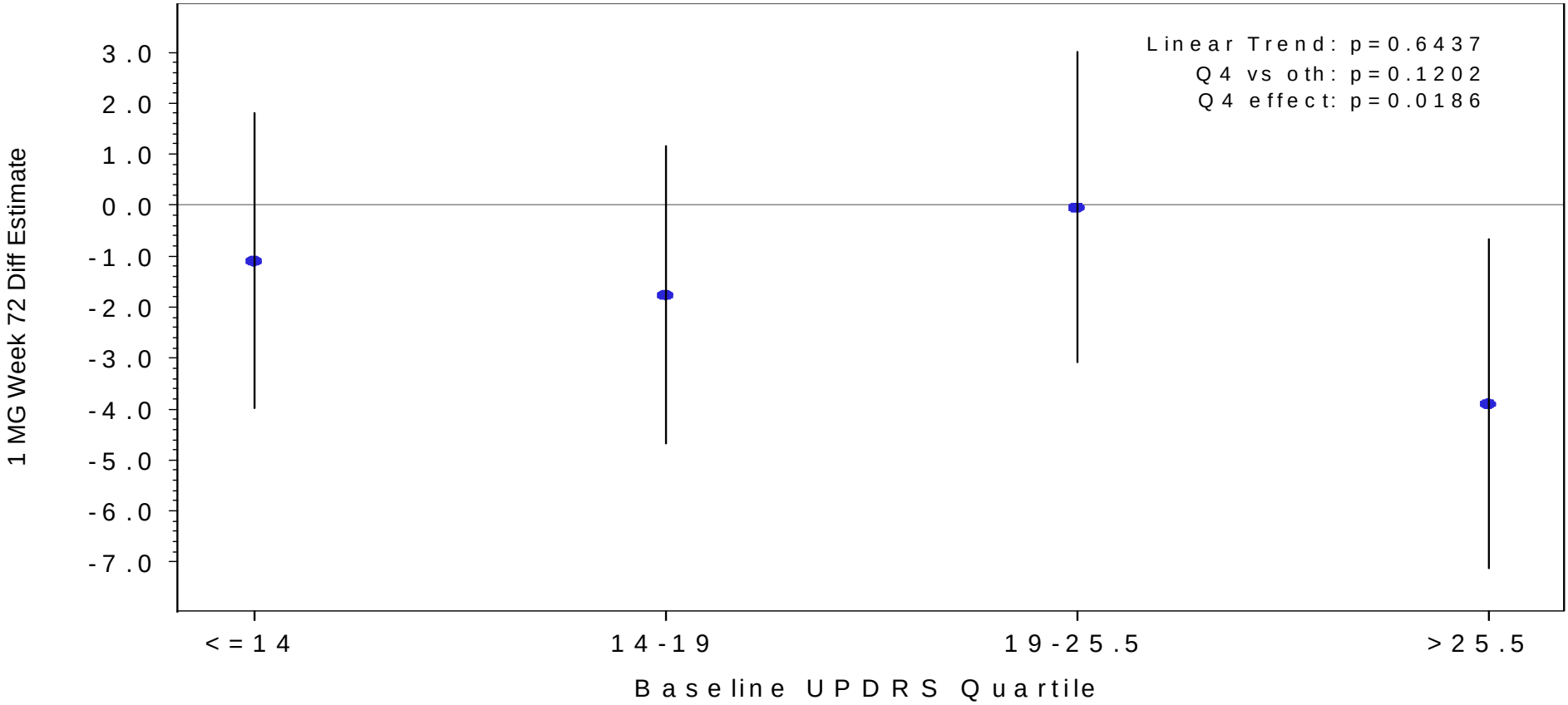
ADAGIO: 2 mg Week 72 Difference for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ACTE Population)



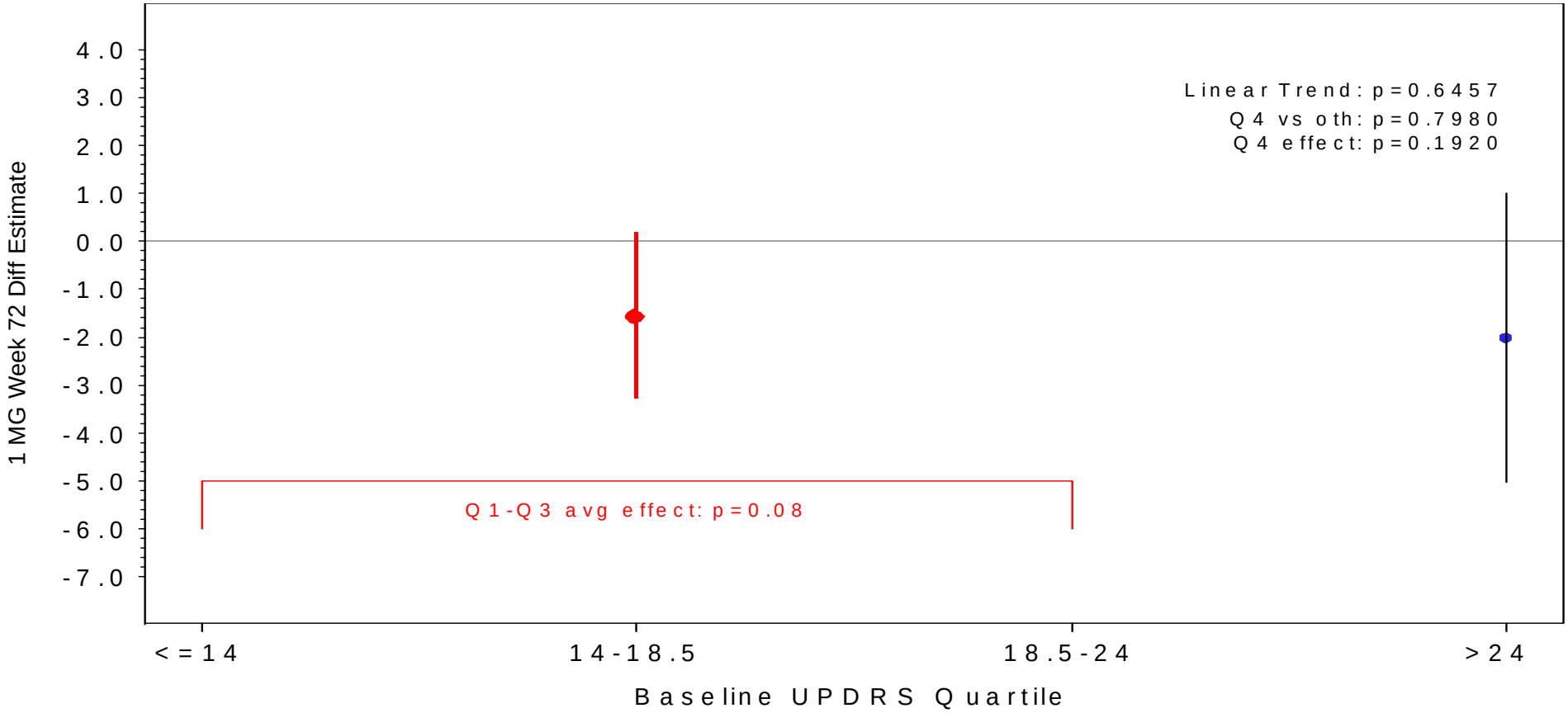
ADAGIO: 1 mg Week 72 Difference (ES – DS) for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ITT Population)



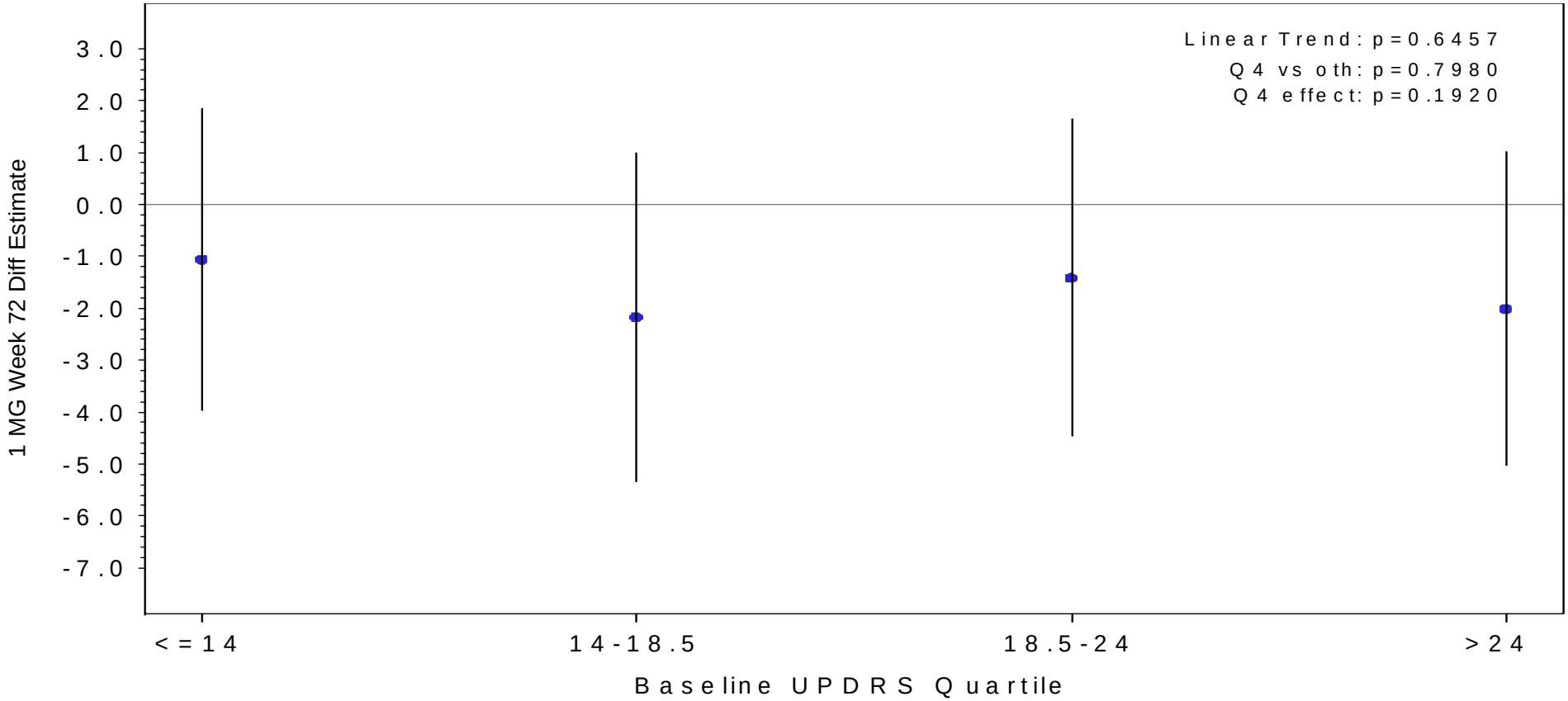
ADAGIO: 1 mg Week 72 Difference (ES – DS) for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ITT Population)



ADAGIO: 1 mg Week 72 Difference (ES – DS) for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ACTE Population)



ADAGIO: 1 mg Week 72 Difference (ES – DS) for ACTE Population by Baseline UPDRS Quartiles (Quartiles Derived from ACTE Population)



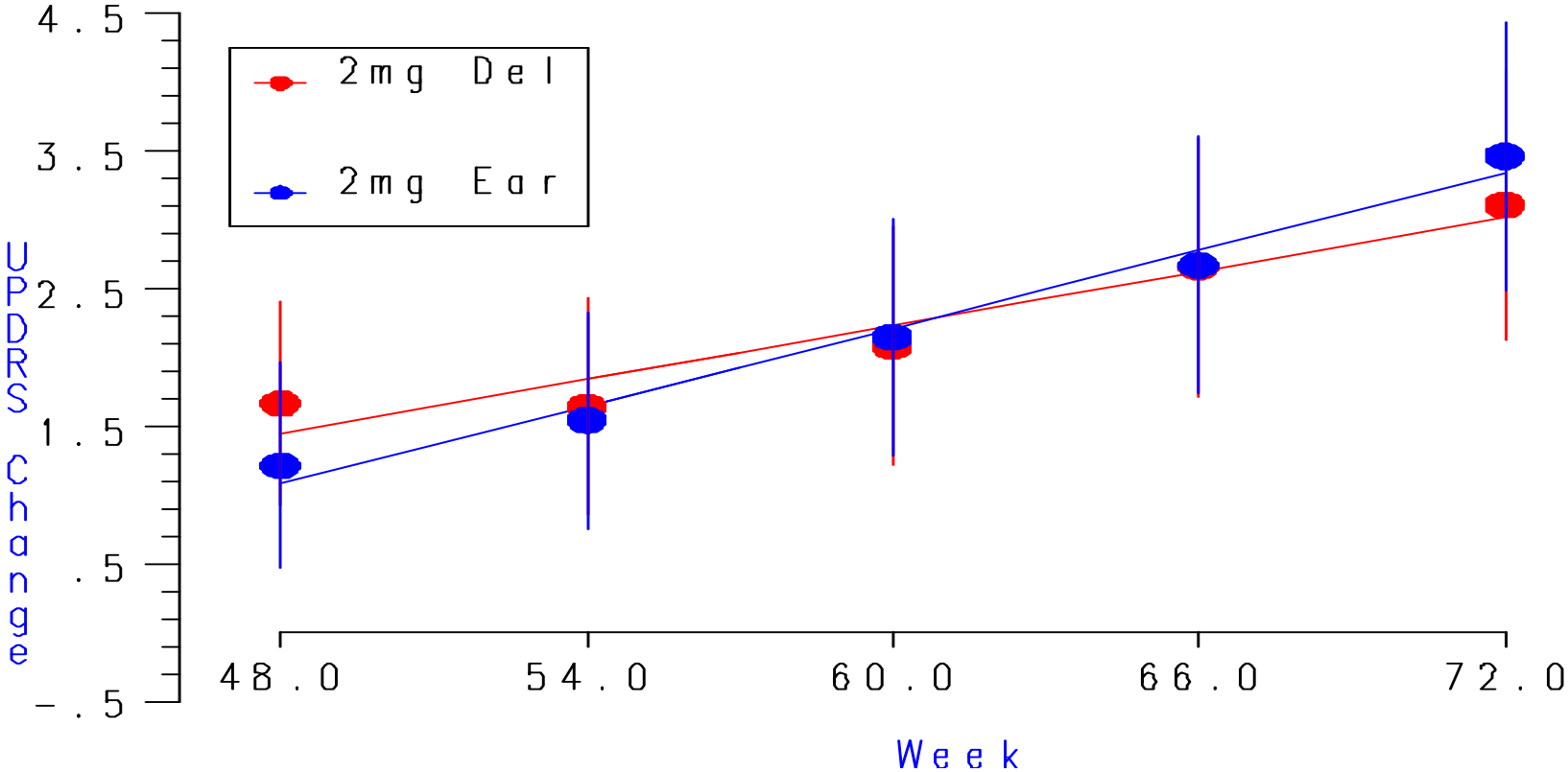
Concerns About Inconsistent Findings for Sponsor's Post-Hoc Hypothesis/Explanation Why 2 mg Dose Failed

- ITT derived and ACTE derived baseline quartiles give different picture for 1 mg
  - Possible floor effect vs. no floor effect
- 4th ITT quartile best, but 3rd worst
  - No evidence of linear trend so have to assume 4th quartile cut point is special
  - 4th quartile has higher proportion of early switchers relative to rest of delayed group
    - This may confound effect of 4th quartile on Week 72

Hypothesis 3:

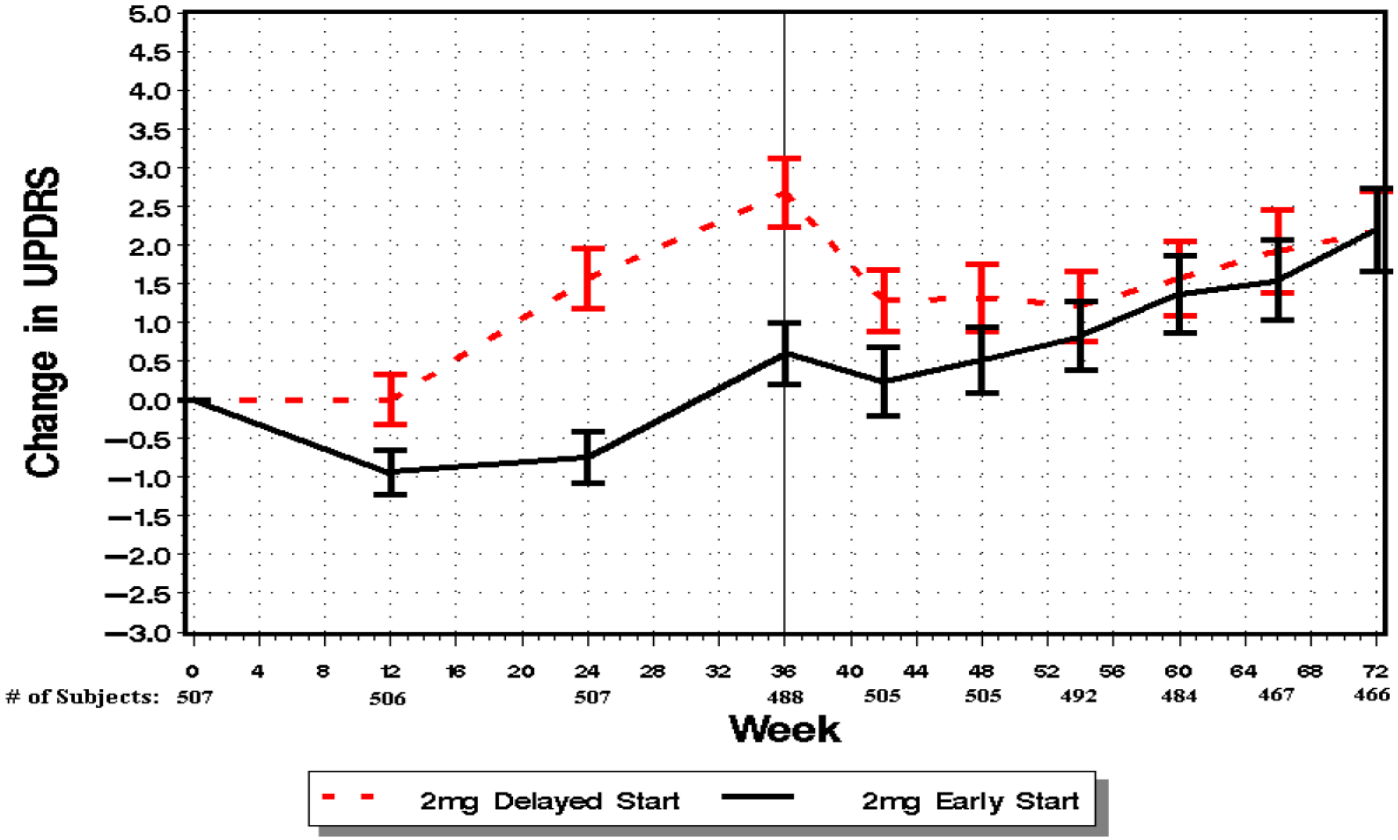
- Non-inferiority of slopes of Early Start vs. Delayed Start in the Active Phase <u>After Week 42.</u>
  - Test for “parallelism” (non-converging) of slopes over Active Phase weeks 48 through 72 for Early Start vs. Delayed Start .
    - Upper 90% confidence limit for slope difference (Early minus Delayed) should not exceed non-inferiority margin of 0.15 points / week
- This test was <u>only</u> to be performed if Hypothesis 1 and 2 tests were both statistically significant.
- Test of parallelism is irrelevant if there is no difference at Week 72.

ADAGIO Hypothesis 3: 0.15 Margin for Slope Difference = Parallelism?



- 2 mg early group slope numerically bigger: diff=0.029, 90% CI (-.005,.062)
- Early line crosses over Delayed line **but** upper confidence limit << 0.15 margin

Concerns that Margin (0.15) for Non-Inferiority of Slopes as Assessment for Slope Parallelism in Second, Active Treatment Phase is Inappropriate/Excessive



* Numbers represent the number of subjects per week. The larger number of subjects in Week 42 compared to Week 36 is due to the early transfer of subjects from Phase I to Phase II.

From Teva Backgrounder, Figure 9, Page 69

Summary of Concerns that Margin (0.15) for Non-Inferiority of Slopes as Test for Parallelism of Slopes in Second, Active Treatment Phase is Inappropriate/Excessive

- Visual inspection of active treatment phase 2 shows slopes for 2 mg Early Start group and 2 mg Delayed Start group ARE NOT PARALLEL.

- BUT statistical analysis using 0.15 non-inferiority margin for slope difference indicates 2 mg Early Start and 2 mg Delayed Start group slopes are statistically “parallel” despite fact that Early Start group actually crosses Delayed Start group and is worse at week 72.

Hypothesis 3 Linearity Assumption

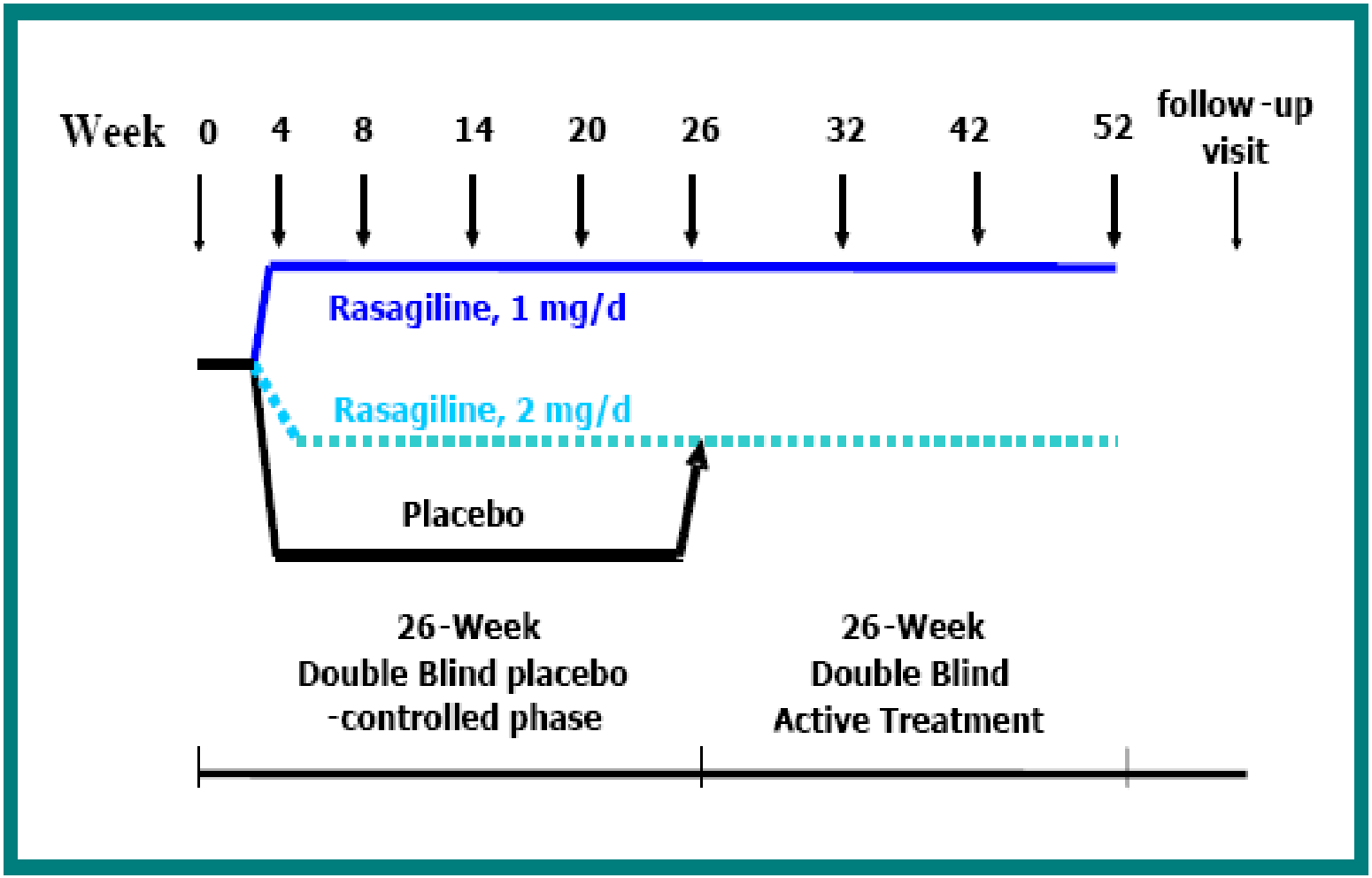
- Hypothesis 3 parallelism test assumes linearity (constant slope) of UPDRS over weeks 48 through 72

- Sponsor nonlinearity test involved all four groups: p=0.0893 so they conclude linearity
  - But only 1 mg eligible for Hypothesis 3

- Nonlinearity test applied to just 1 mg groups: p=0.0435

- Sample mean plot over time shows arguable linearity but plot is a simplification of actual correlated patient level data

TEMPO Study Design

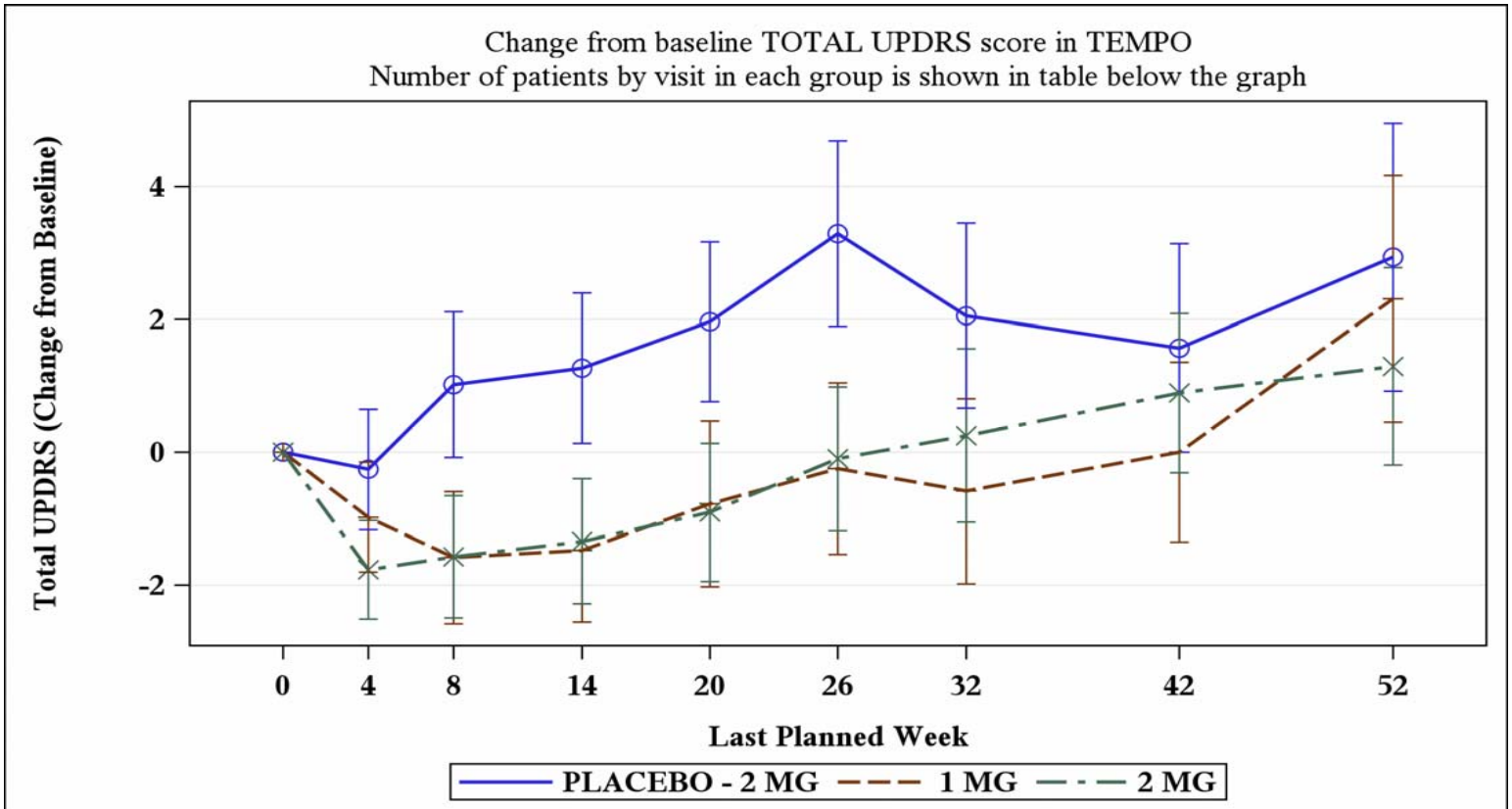


From Integrated Summary of Efficacy, page 15

TEMPO Trial

- Primary Analysis for Placebo-Controlled Phase at Week 26 (end of Phase 1)
  - Before Delayed Start Active Treatment Phase 2
- End of Active Phase Analysis
  - Active Phase designed primarily for safety and to explore efficacy
  - Statistical Analysis Plan for Week 52 analysis not submitted for FDA review and comment prior to unblinding
  - No single primary efficacy endpoint nor single analysis population specified

TEMPO (Observed UPDRS Data)



PLACEBO - 2 MG	126	125	126	125	114	109	117	95	87
1 MG	120	120	118	116	109	108	113	96	88
2 MG	118	118	116	114	112	105	106	91	80

Rationale for Conducting a Second Trial

- Preliminary review of TEMPO suggested a possible disease modifying effect though not definitive
- FDA informed sponsor of need for two positive, pivotal trials (i.e., replication of results) showing disease modification (or supportive data plus robust study demonstrating disease modification. FDA told sponsor (at 12/22/04 meeting):

"The Number of Required Studies"

- ***Ordinarily 2 trials are required to support efficacy.***
- ***The TEMPO study post hoc analysis may not be sufficient for review because it is not the primary analysis. If the next study is robustly positive, then the TEMPO study may provide supporting evidence."***

TEMPO Data Analyzed by ADAGIO Analysis Plan

- TEMPO more Phase 1 assessments, shorter overall duration
- This is an underpowered, exploratory analysis

2 mg Early vs. 2 mg Delayed Results

Hypothesis	Estimate	95% C.I.	P-value
1: Week 14-26 Slope Difference	-0.083	(-0.197,0.030)	0.1475
2: Week 52 Difference	-1.934	(-4.078, 0.209)	0.0768
3: WK 42-52 Slope Difference	-0.100	(-0.234, 0.033)[90%]	n/a

1 mg Early vs. 2 mg Delayed Results (No 1 mg Delayed Group)

Hypothesis	Estimate	95% C.I.	P-value
1: Week 14-26 Slope Difference	-0.085	(-0.197, 0.026)	0.1342
2: Week 52 Difference	-0.696	(-2.753, 1.360)	0.5055
3: WK 42-52 Slope Difference	0.073	(-0.057, 0.203)[90%]	n/a

TEMPO Active Phase Results According to Original Analysis Plan

– Statistically significant for 2 mg using Last Observation Carried Forward (LOCF) Imputation (protocol specified)

- LOCF problematic for assessing progression
- May bias in favor of earlier treated group

<u>2 mg Week 52 Group Mean difference (ES-DS)</u>

-1.2 Completers [65% of ITT has Week 52 UPDRS]

-2.2 LOCF

- Exploratory Repeated Measures Model:

2 mg Week 52 difference: -2.0, p=0.0501

Limitations of TEMPO Long Term, Follow-Up

- Open label
- Confounded by dropouts and concomitant medications
  - Variable discontinuation from follow-up by Group
  - Allowed confounding treatment with additional Parkinson's treatments
- Data were collected in non-randomized treatment groups

Conclusions

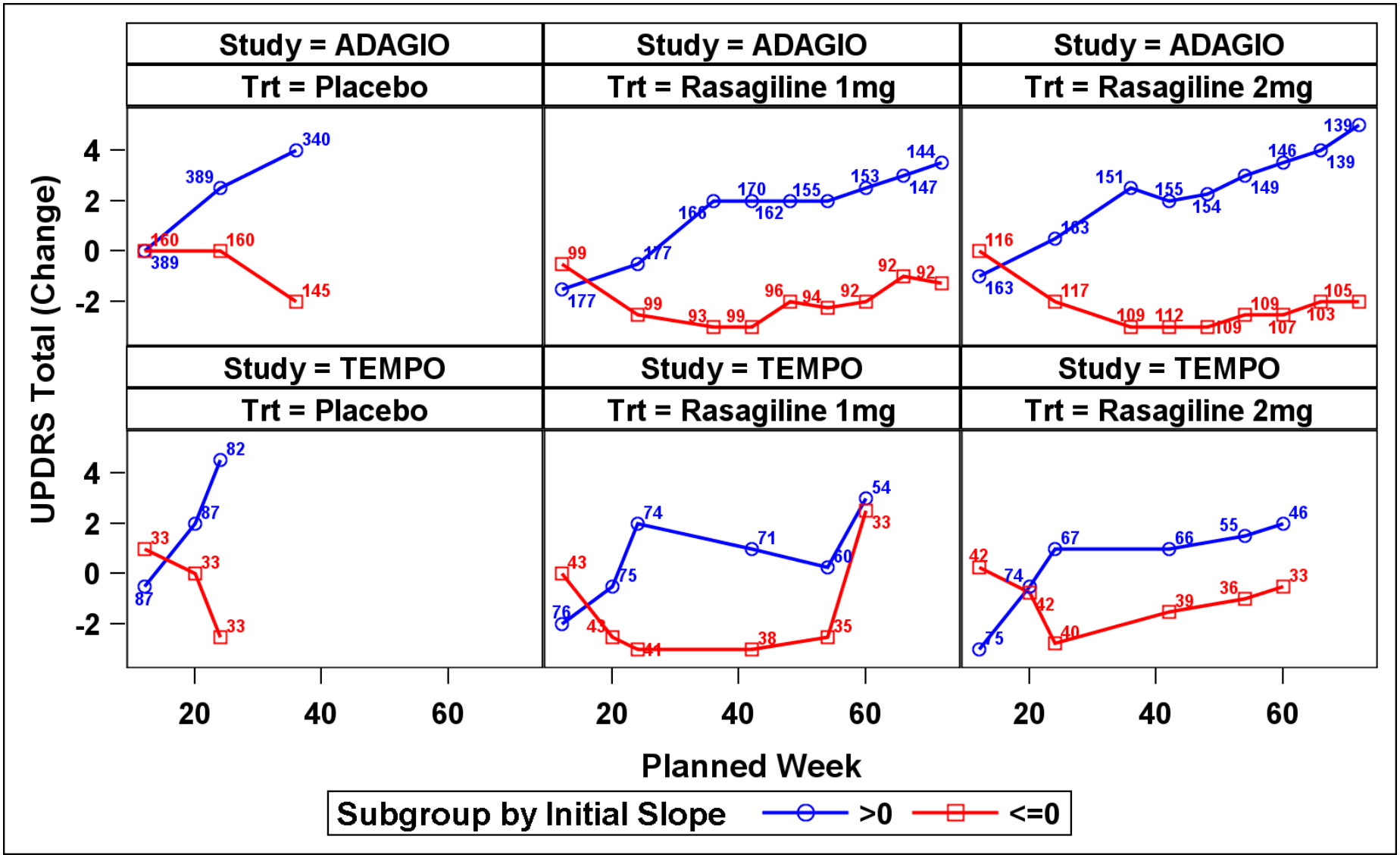
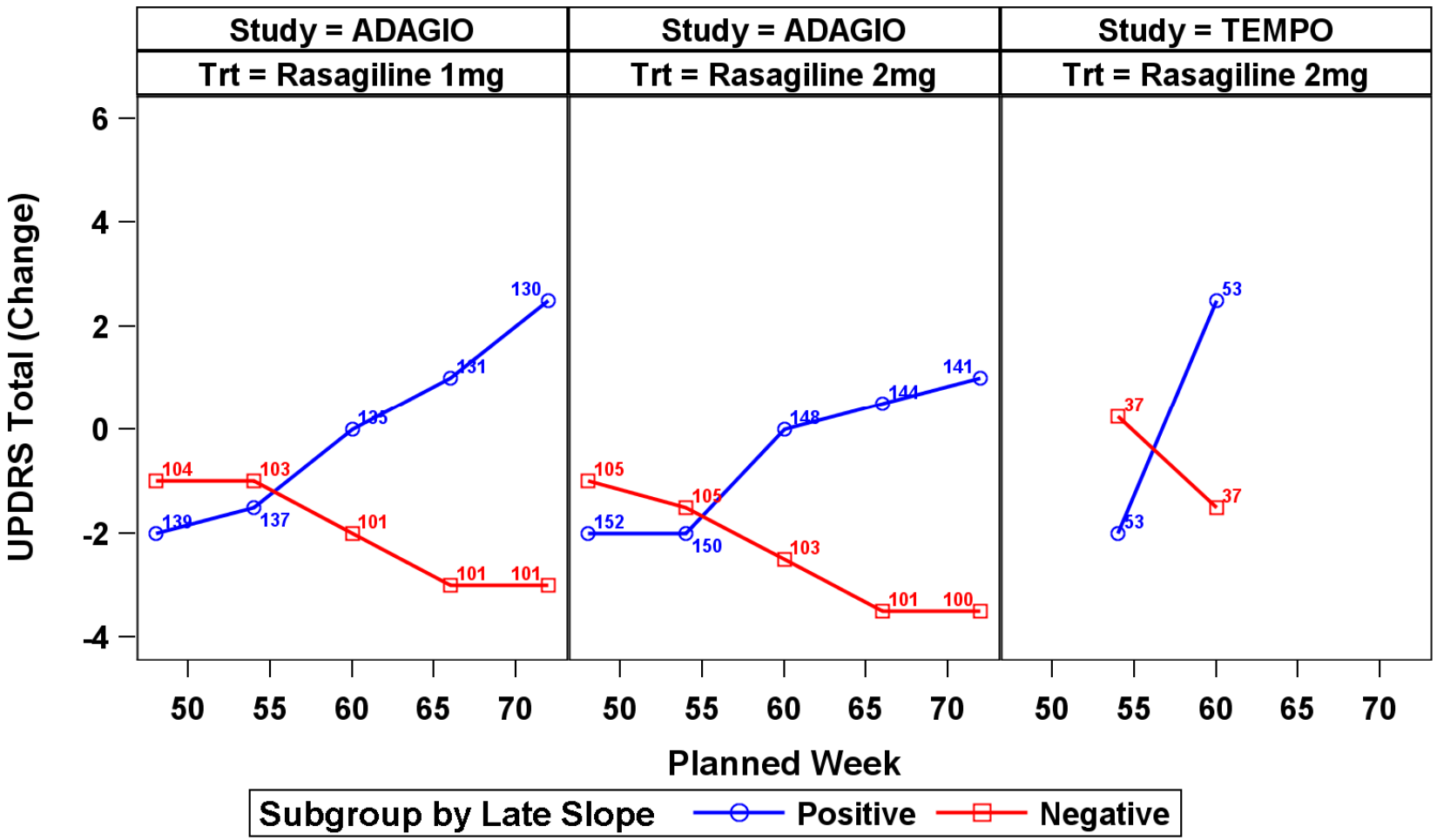
- ADAGIO 2 mg (High Dose)
  - No Early group benefit (i.e., treatment difference) at end of study (week 72)
  - 2 mg Early numerically worse than 2mg Delayed
    - Early – Delayed Difference = + 0.17, p = 0.60
- Other issues/concerns with ADAGIO 1 mg
  - Questions about switching Primary Analysis for Hypothesis 2
  - Baseline group imbalances in ACTE populations - biased sample of ITT
  - Nonlinearity of UPDRS change from baseline over time
  - Significant gender difference in 1 mg efficacy
  - Numerous interactions between treatment and other variables
    - Suggests if 1 mg treatment effect exists, effect is not consistent across subgroups
  - No effect in 2 mg group raises questions about biological plausibility of 1 mg effect

Conclusions (Continued)

- TEMPO 2 mg not considered definitive
- No robust finding for any dose in either study

Backup Slides

Different Time to Full Symptomatic Effect in Subgroups



Different Time to Full Symptomatic Effect in Subgroups (Delayed Groups in Phase 2)