**Appendix II Statistical Supplement**

*IPTW diagnostics*

The propensity score was estimated using binary logistic regression where treatment assignment (MACARTI vs. SOC) was regressed on the 12 covariates (11 nominal and 1 continuous) presented in **Supplemental Table 1**. Multiple specifications of the propensity model were considered, as suggested by Austin & Stuart,[1] and included a main effects-only model (simple model), as well as a complex model that utilized restricted cubic splines, with 4 knots, for the continuous age covariate. For the simple model, it was assumed that the age covariate was linearly related to the log-odds of receiving MACARTI treatment; likewise, the complex model considered the relationship between the spline-age covariate and the log-odds of treatment, adjusted for the 11 other nominal covariates. Average treatment effect (ATE) weights were calculated from the fitted values of each of the simple and complex models using the established formula: , where Z are participant-level treatment allocations (characterized as 1 (MACARTI) and 0 (SOC)), and *e* are the participant-level fitted values from the binary logistic regression models.[2] Weights were stabilized using the *ipw* package in CRAN R; moreover, due to small sample size and noted disparities in the treatment and control groups, stabilized weights were considered both untrimmed and trimmed at the 1% and 99%.[3] Unweighted standardized differences (effect sizes) are presented in **Supplemental Table 1** and calculated using the formulas presented by Austin & Stuart.[1] Weighted standardized mean differences, using both untrimmed and trimmed stabilized weights, are presented in **Supplemental Table 2** for each of the simple and complex logistic propensity models and were calculated using the *tableone* package in CRAN R. These values aid in the determination of study cohort balance, with regards to the 12 noted covariates, in the weighted samples. At present, there is no universal cutoff for establishing imbalance between cohorts. A number of authors have cited 10% as being reasonable;[4, 5] whereas, others have utilized cutoffs as high as 20% and 25%.[6] Cohen defined standardized differences (or effect sizes) loosely using the following criteria: 20% as small, 40% as moderate, and 60% as large.[7] Considering these values, and given our small size and level of cohort disparity, we utilized 25% as our weighted standardized difference cutoff for indication of covariate imbalance between the MACARTI and SOC groups. Finally, in an effort to create as much homogeneity between cohorts as possible, multiple forms of the simple model were considered using perceived, increasingly uniform sub-samples (i.e. men only black only black men), to determine if better covariate balance could be achieved within a subset of the sampled population (**Supplemental Table 2**).

*Stabilized Weight Diagnostics*

First considering each of the simple models (four options: all sampled participants and the three sub-samples, men only; black only; black men), means and standard deviations for the untrimmed, stabilized weights were all found to be close to 1, with similar standard deviations (All participants, Mean: 0.984 (SD: 1.101); Men Only, 0.994 (0.951); Black Only, 0.990 (1.000); Black Men, 1.008 (0.952)). In each case, means and standard deviations for the trimmed weights did not substantially differ from the untrimmed weights. After careful consideration of each of the four candidate simple models, the simple model containing all participants with untrimmed weights was selected, as it did not perform notably worse than any of the other candidate simple models, indicated by mean stabilized weight summaries and covariate standardized mean differences in **Supplemental Table 2**; moreover, this model permitted the use of all participant data, providing better inference (and coverage) for the target population, newly diagnosed HIV patients in MSA-Atlanta. The minimum and maximum for the selected simple model were 0.482 and 9.362, respectively. The complex model, also utilizing all participants and untrimmed stabilized weights, had a mean and standard deviation of 1.020 (1.399), as well as a minimum and maximum of 0.481 and 10.530, respectively. Given these values, non-positivity and misspecification of the propensity models did not appear to be of serious concern.

*Weighted Sample Mean Comparisons and Higher Order Terms*

The largest absolute weighted standardized differences for the simple model specification were for race (0.230) and condom usage (0.249), with the remaining standardized differences being under 0.20 (or 20%). For the complex model specification, the largest absolute weighted standardized differences were for race (0.267), education (0.230), sexual orientation (0.221), and condom usage (0.241), with the remaining standardized differences being under 0.20. In the simple model, all 12 covariates were below the weighted standardized difference cutoff of 25%; concurrently, 11 of 12 covariates met this criterion in the complex model. In contrast, unweighted standardized differences were large, reaching as high as 0.85, with the majority of other covariates differing between 0.25 and 0.70. Only three covariates met the cutoff threshold of 25%, unweighted (currently using drugs, abuse type, condom usage). **Supplemental Figure 1** presents absolute standardized mean differences for the unweighted sample and weighted samples, using the simple and complex model specifications. A dashed line is included at 25% to indicate the tolerance level for covariate imbalance between the MACARTI and SOC cohorts.

Having only one continuous covariate, continuous interactions were not considered in the simple model specification; however, higher order terms for age were further included as main effects (i.e. square and cube). In the unweighted sample, the standardized mean difference for the square and cube of age were 0.273 and 0.270, respectively. In separate simple model specifications, including each of Age2 and Age3 as main effects, the weighted mean standard differences were 0.103 and 0.100, respectively. These results suggest that balance has been achieved between MACARTI and SOC participants with regards to the higher order terms of age. Together, with the results presented in **Supplemental Figure 1**, it appears that the simple propensity specification is superior to the complex specification and will be utilized in weighted linear model analysis for the CD4 and Viral Load outcomes.

*Graphical and Statistical Comparisons for the Age Covariate*

As a final check for the continuous age covariate, empirical cumulative distribution functions are plotted for the unweighted and weighted samples (**Supplemental Figure 2**); additionally, Kolmogorov-Smirnov D-statistics were utilized to determine if the unweighted and weighted age distributions significantly differed between the MACARTI and SOC cohorts. A bootstrapped version of the traditional Kolmogorov-Smirnov test was implemented (with 10,000 boots), allowing for ties within the data. All resulting D-statistics were small, 0.036 in the unweighted sample, 0.088 and 0.176 in the simple and complex weighted samples, respectively, indicating MACARTI and SOC age distributions as insignificant from each other. Provided the D-statistics and visualizations of the data in **Supplemental Figure 2,** the simple model weight balances distributional differences better than the complex model weight, but neither the simple nor complex weighting improves age balance between the study groups, relative to the unweighted sample.

**Supplemental Table 1: Unweighted standardized differences for propensity model covariates, reproduced from the primary text for convenience**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Characteristic, N (%)** | **Overall**  **N = 98** | **SOC**  **N = 49** | **MACARTI**  **N = 49** | **P-Value** | **Unweighted**  **Std. Diff** | **Weighted**  **Std. Diff** |
| Gender |  |  |  |  |  |  |
| Male | 83 (84.7%) | 36 (73.5%) | 47 (95.9%) | 0.004 | 0.656 | 0.097 |
| Female | 15 (15.3%) | 13 (26.5%) | 2 (4.1%) |  |  |  |
| Race |  |  |  |  |  |  |
| Black | 89 (90.8%) | 47 (95.9%) | 42 (85.7%) | 0.159 | 0.359 | 0.230 |
| Other (White, Hispanic, Other) | 9 (9.2%) | 2 (4.1%) | 7 (14.3%) |  |  |  |
| Age (yr), Mean ± SD | 21.5 ± 1.8 | 21.3 ± 1.8 | 21.7 ± 1.7 | 0.175 | 0.276 | 0.083 |
| Work Status |  |  |  |  |  |  |
| Employed/In School | 74 (75.5%) | 32 (65.3%) | 42 (85.7%) | 0.019 | 0.489 | 0.139 |
| Neither | 24 (24.5%) | 17 (34.7%) | 7 (14.3%) |  |  |  |
| Education, N = 97 |  |  |  |  |  |  |
| High school or Less | 60 (61.9%) | 35 (72.9%) | 25 (51%) | 0.026 | 0.463 | 0.154 |
| College or More | 37 (38.1%) | 13 (27.1%) | 24 (49%) |  |  |  |
| Ever Abused Alcohol | 15 (15.3%) | 3 (6.1%) | 12 (24.5%) | 0.022 | 0.528 | 0.083 |
| Currently Using Drugs | 22 (22.5%) | 9 (18.4%) | 13 (26.5%) | 0.333 | 0.197 | 0.008 |
| Abused Type |  |  |  |  |  |  |
| No Abuse | 84 (85.7%) | 42 (85.7%) | 42 (85.7%) | 1.000 | <0.001 | <0.001 |
| Abused | 14 (14.3%) | 7 (14.3%) | 7 (14.3%) |  |  |  |
| Sexual Orientation |  |  |  |  |  |  |
| Straight | 22 (22.5%) | 19 (38.8%) | 3 (6.1%) | <0.001 | 0.850 | 0.198 |
| Gay/Bisexual/Queer | 76 (77.5%) | 30 (61.2%) | 46 (93.9%) |  |  |  |
| Condom Usage |  |  |  |  |  |  |
| Always/Usually | 71 (72.5%) | 33 (67.4%) | 38 (77.6%) | 0.258 | 0.230 | 0.249 |
| Sometimes/Never | 27 (27.5%) | 16 (32.6%) | 11 (22.4%) |  |  |  |
| Ever had STI – Patient Report, N = 97 | 47 (48.5%) | 28 (57.1%) | 19 (39.6%) | 0.084 | 0.357 | 0.071 |
| Any AIDS defining conditions, N = 94 | 34 (36.2%) | 25 (51%) | 9 (20%) | 0.002 | 0.685 | 0.112 |

**Supplemental Table 2: Weighted standard differences, trimmed and untrimmed, for simple and complex propensity models by covariate**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Characteristic, N (%)** | **Candidate Simple Models** | | | | | | | | **Complex Model** | |
| **Weighted Std. Diff (All)** | | **Weighted Std.**  **Diff (Men Only)** | | **Weighted Std.**  **Diff (Black Only)** | | **Weighted Std.**  **Diff (Black Men)** | | **Weighted Std.**  **Diff (All)** | |
| **UnTrim** | **Trim** | **UnTrim** | **Trim** | **UnTrim** | **Trim** | **UnTrim** | **Trim** | **UnTrim** | **Trim** |
| Gender | 0.097 | 0.131 | *NA* | | 0.109 | 0.120 | *NA* | | 0.132 | 0.140 |
| Race | 0.230 | 0.215 | 0.259 | 0.253 | *NA* | | *NA* | | 0.267 | 0.263 |
| Age (yr) | 0.083 | 0.055 | 0.138 | 0.126 | 0.063 | 0.054 | 0.102 | 0.108 | 0.068 | 0.061 |
| Work Status | 0.139 | 0.185 | 0.091 | 0.108 | 0.136 | 0.151 | 0.096 | 0.108 | 0.091 | 0.102 |
| Education | 0.154 | 0.061 | 0.177 | 0.145 | 0.118 | 0.088 | 0.162 | 0.141 | 0.230 | 0.209 |
| Ever Abused Alcohol | 0.083 | 0.090 | 0.037 | 0.027 | 0.099 | 0.039 | 0.054 | 0.066 | 0.065 | 0.023 |
| Currently Using Drugs | 0.008 | 0.054 | 0.006 | 0.012 | 0.022 | 0.037 | 0.048 | 0.006 | 0.126 | 0.140 |
| Abused Type | <0.001 | 0.031 | 0.038 | 0.027 | 0.066 | 0.077 | 0.049 | 0.057 | 0.035 | 0.028 |
| Sexual Orientation | 0.198 | 0.241 | 0.269 | 0.279 | 0.218 | 0.233 | 0.275 | 0.282 | 0.221 | 0.231 |
| Condom Usage | 0.249 | 0.200 | 0.163 | 0.144 | 0.266 | 0.249 | 0.181 | 0.168 | 0.241 | 0.228 |
| Ever had STI – Patient Report | 0.071 | 0.007 | 0.181 | 0.152 | 0.119 | 0.093 | 0.235 | 0.215 | 0.168 | 0.150 |
| Any AIDS defining conditions | 0.112 | 0.172 | 0.025 | 0.003 | 0.165 | 0.185 | 0.008 | 0.023 | 0.057 | 0.071 |

1All weights are ATE and stabilized by multiplying the values by the marginal logistic probability of receiving MACARTI

2Weighted Standard Differences are presented as ‘Untrimmed and Trimmed (at 1% and 99%)’ and considered non-impactful if less than 0.25

3Candidate simple models are shown including: All participants, men only, black only, and black men only; the all participant model was selected for sample weighting

**Supplemental Figure 1: Standardized differences in weighted and unweighted samples**

**C:\Users\sgille3\Projects\MACARTI\July 2016\Resubmission Papers\Weight Plot.tif**

**Supplemental Figure 2: Cumulative distributions of age between MACARTI and SOC**

**C:\Users\sgille3\Projects\MACARTI\July 2016\Resubmission Papers\Age CDF.tif**

*Weighted CD4 Linear Growth Model*

An individual linear growth model for CD4 count was constructed using the PROC MIXED procedure in SAS, allowing the growth parameters for each study participant to be examined as random effects. As noted in the statistical methods, the model specification included study arm (MACARTI versus SOC) as the fixed effect and participant-specific intercepts and study visit slopes as the random effects. An unstructured variance-covariance matrix was employed for repeated observations, and degrees of freedom were estimated using the between-within method. Derived stabilized weights were utilized, balancing noted confounding covariates at baseline as previously described. Per the project design, study visits (time) were targeted for 0, 30, 90, 180, and 365 days, but specifically calculated based upon actual observed visit dates for each participant (relative to baseline), and treated continuously. The outcome, CD4 count, was transformed via a square-root transformation and plotted, unweighted, against participant-specific study visits in each of the MACARTI and SOC arms (**Supplemental Figure 3a**). Study visit means of transformed CD4 counts were calculated and overlaid on the figure (dashed-line), demonstrating a curve-linear association with time. As such, the final growth model featured a quadratic term for time, as well as interactions between time, both quadratic and linear, and study arm. The fixed effects table for the model is presented in **Supplemental Table 3a,** and least squares mean estimates are in **Supplemental Table 3b**.

**Supplemental Table 3a: CD4 count linear growth model fixed effects estimates**

|  |  |  |
| --- | --- | --- |
| **Fixed Effect** | **Estimate (SE)** | **P-value** |
| Study Visit | 0.021 (0.005) | <0.001 |
| Study Visit2 | -3.0e-5 (1.3e-5) | 0.019 |
| Study Arm - MACARTI | 14.989 (0.860) | <0.001 |
| Study Arm - SOC | 18.157 (0.901) | <0.001 |
| Study Visit\*Study Arm |  |  |
| SOC | -0.020 (0.007) | 0.004 |
| MACARTI | Reference |  |
| Study Visit2\*Study Arm |  |  |
| SOC | 5.1e-5 (1.7e-5) | 0.004 |
| MACARTI | Reference |  |

1Intercept covariance: 30.86 (5.08), p<0.001; Slope covariance: 6.0e-5 (1.8e-5), p<0.001

2CD4 count has been square-root transformed

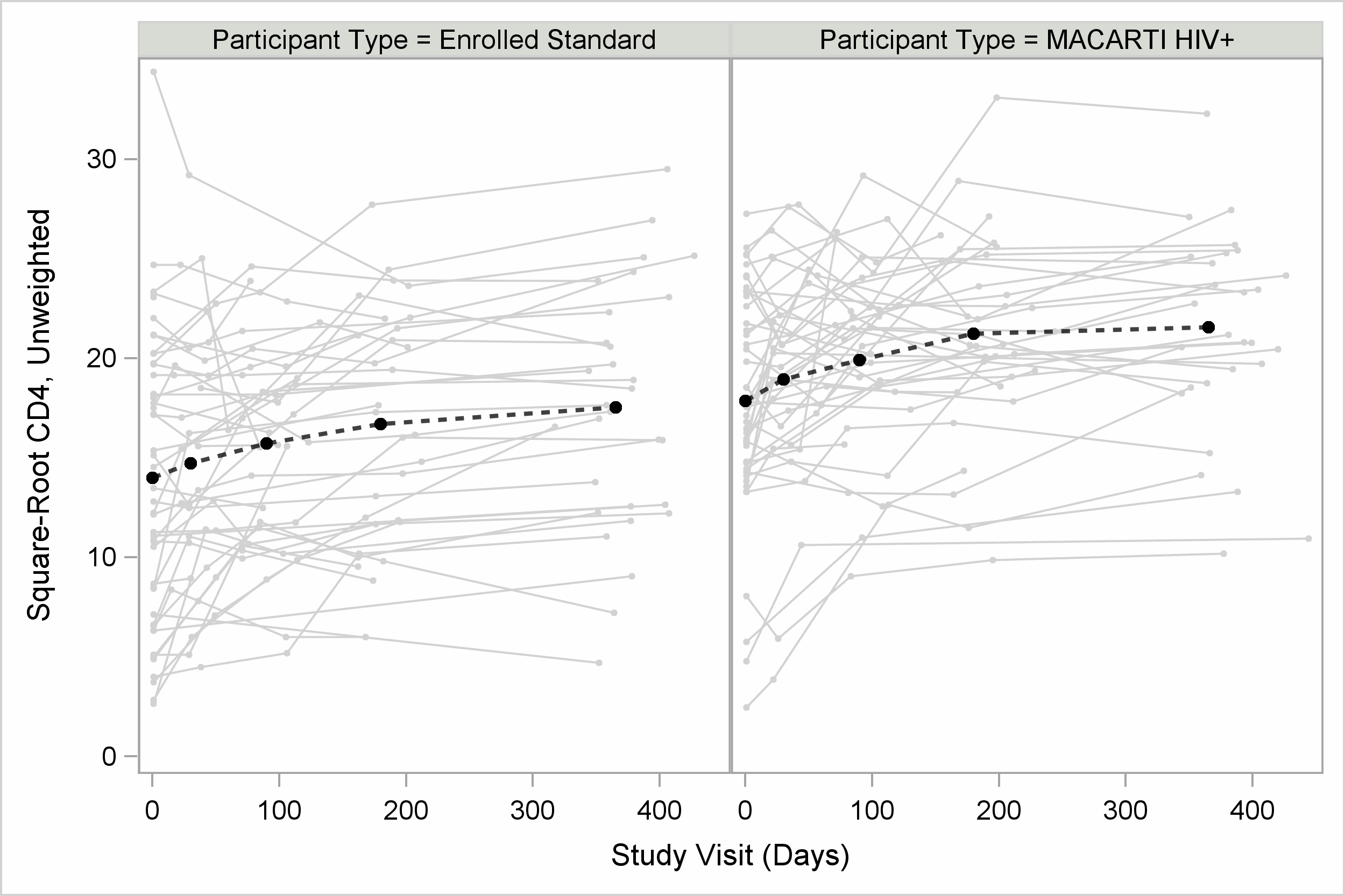
**Supplemental Table 3b: CD4 count linear growth model fixed effects estimates**

|  |  |  |  |
| --- | --- | --- | --- |
| **Study Visit** | **Standard Arm**  **Mean CD4 (95% CI)** | **MACARTI Arm**  **Mean CD4 (95% CI)** | **P-Value** |
| Baseline | 225 (176 - 279) | 330 (268 - 398) | 0.012 |
| 30 Days | 226 (179 - 278) | 352 (291 - 419) | 0.002 |
| 90 Days | 232 (186 - 283) | 393 (330 - 461) | <0.001 |
| 180 Days | 250 (203 - 302) | 441 (375 - 512) | <0.001 |
| 365 Days | 326 (270 - 387) | 479 (409 - 555) | 0.001 |

1CD4 count estimates and 95% CI back-transformed via squaring

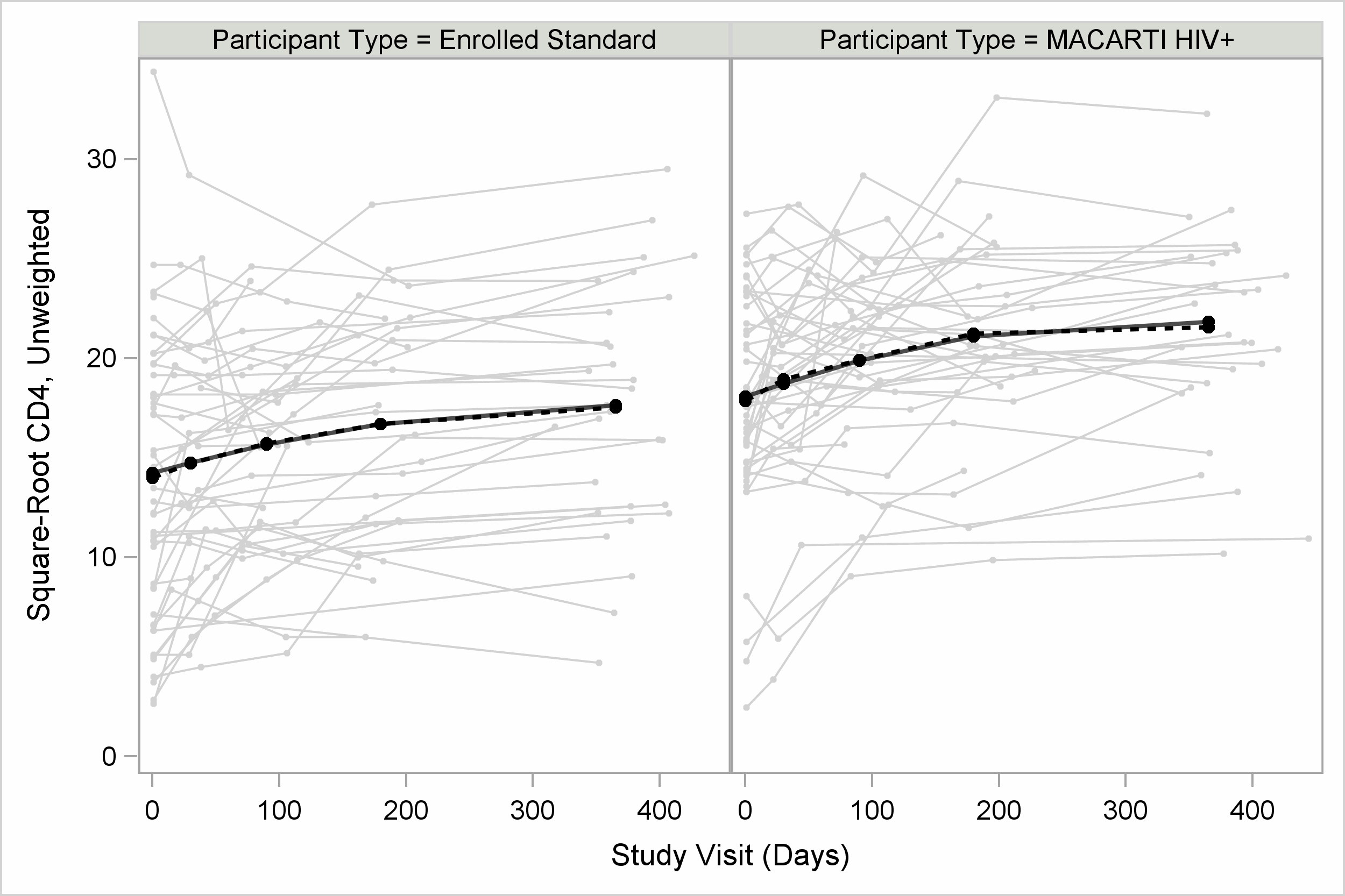
**Supplemental Table 3a** presents significant interactions and main effects, for both linear and quadratic terms; moreover, significant random effects are noted in the table footnote for both the intercept and slope. **Supplemental Figure 3b** shows unweighted, model-based least-squares mean values connected and overlaid on participant-level transformed CD4 observations (solid-line). Relative to the un-modeled means (dashed-line), the quadratic linear growth model appears to strongly describe the relationship between transformed CD4 counts and study visits across treatment arms. Finally, **Supplemental Figure 3c** shows weighted, model-based least-squares mean values connected and overlaid on predicted, participant-level transformed CD4 observations as described in the primary text.

**Supplemental Figure 3a: Participant-level CD4 counts means (square-root) over study visit follow-up, unweighted**



1Study days 0, 30, 90, 180, 365 are presented as mean points and connected with a dashed-line

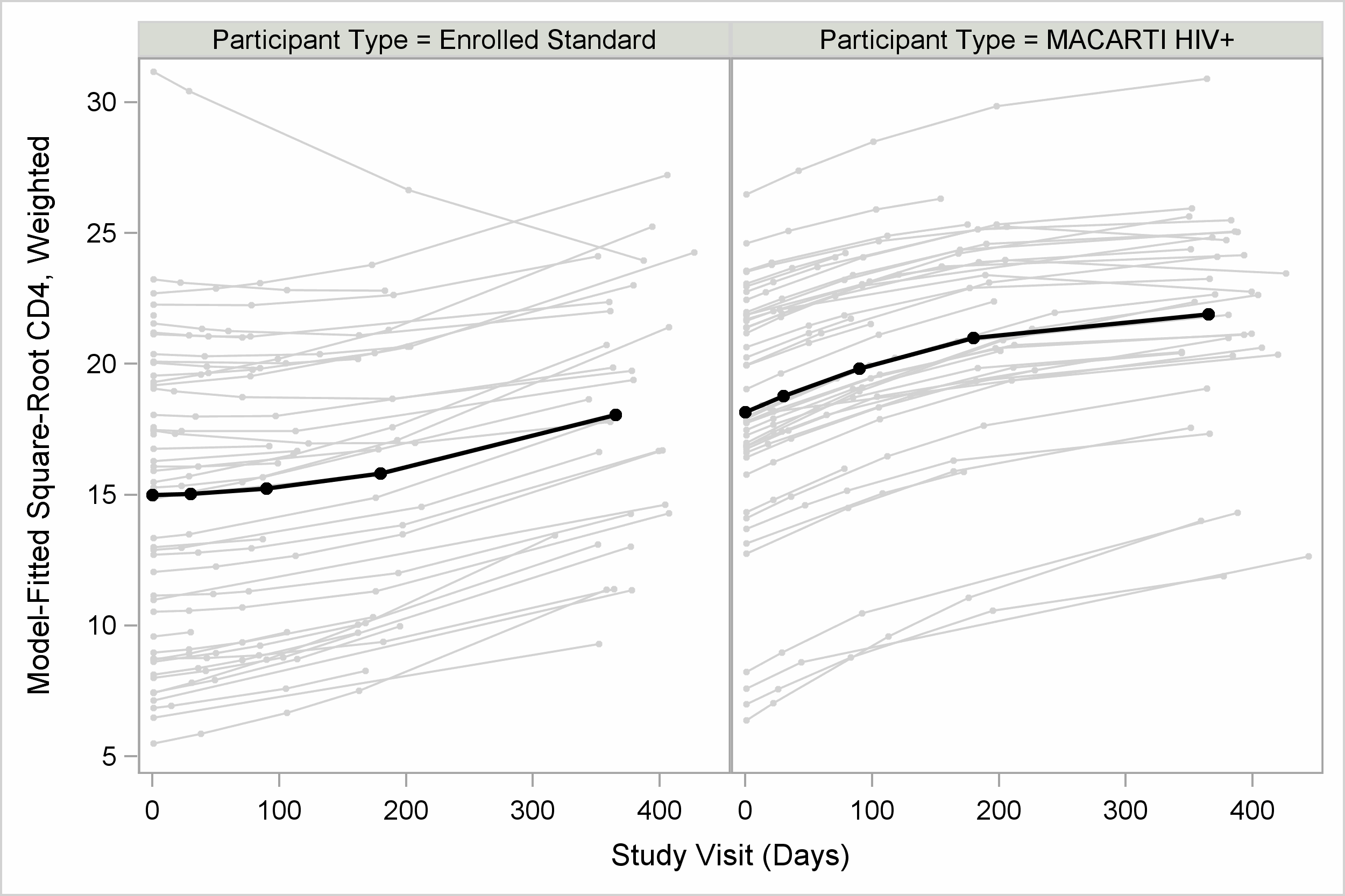
**Supplemental Figure 3b: Growth model-based mean CD4 count estimates (square-root) versus un-modeled mean estimates, unweighted**



1Study days 0, 30, 90, 180, 365 are presented as mean points and connected with a

dashed-line for un-modeled values and a connected-line for growth model estimates

**Supplemental Figure 3c: Growth model-based mean CD4 count estimates (square-root) and predicted patient-level observations, weighted**



1Weighted-growth model estimates for study days 0, 30, 90, 180, 365

*Weighted Viral Load Linear Growth Model*

As with CD4, an individual linear growth model for Viral Load was constructed using the PROC MIXED procedure in SAS. The model specification included study arm (MACARTI versus SOC) as the fixed effect and participant-specific intercepts and study visit slopes as the random effects. An unstructured variance-covariance matrix was employed for repeated observations, and degrees of freedom were estimated using the between-within method.[8] Derived stabilized weights were utilized to balance baseline covariate differences between study arms. Both the outcome, Viral Load, and study visit days were natural log-transformed and plotted, unweighted, in each of the MACARTI and SOC arms (**Supplemental Figure 4a**). As with CD4 counts, study visit means of transformed Viral Loads were calculated and overlaid on the figure (dashed-line), demonstrating a fairly curve-linear association with log-time. As such, the final growth model featured a quadratic term for log-time, as well as interactions between log-time, both quadratic and linear, and study arm. The fixed effects table for the model is presented in **Supplemental Table 4a,** and least squares mean estimates are in **Supplemental Table 4b.**

**Supplemental Table 4a: Viral load linear growth model fixed effects estimates**

|  |  |  |
| --- | --- | --- |
| **Fixed Effect** | **Estimate (SE)** | **P-value** |
| Log-Study Visit | 0.073 (0.307) | 0.813 |
| Log-Study Visit2 | -0.195 (0.053) | <0.001 |
| Study Arm - MACARTI | 9.985 (0.371) | <0.001 |
| Study Arm - SOC | 10.144 (0.407) | <0.001 |
| Log-Study Visit\*Study Arm |  |  |
| SOC | -0.372 (0.402) | 0.355 |
| MACARTI | Reference |  |
| Log-Study Visit2\*Study Arm |  |  |
| SOC | 0.115 (0.070) | 0.100 |
| MACARTI | Reference |  |

1Intercept covariance: 1.70 (0.87), p=0.025; Slope covariance: 0.081 (0.038), p=0.017

2Viral load and study visit are both natural-log transformed

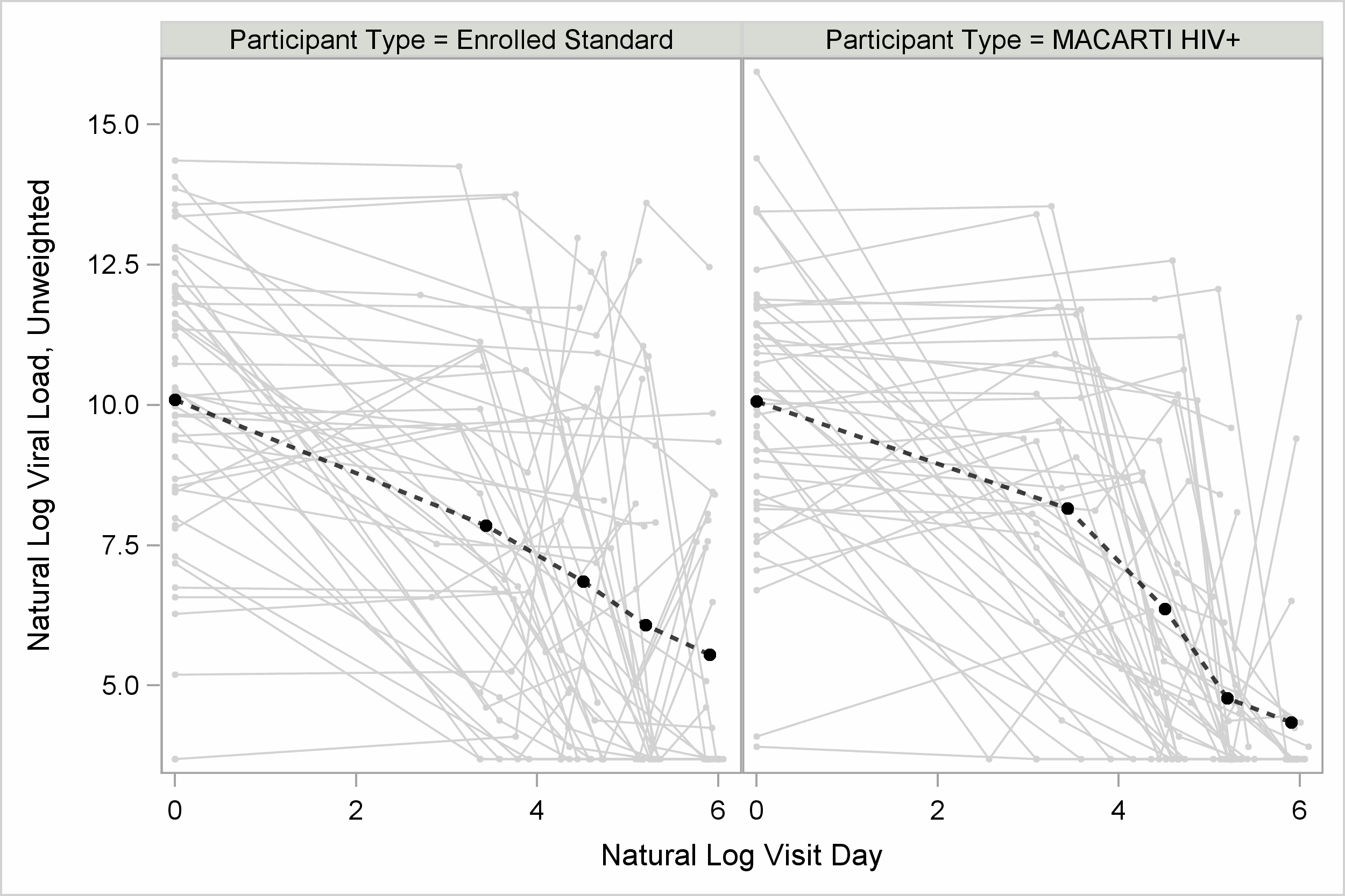
**Supplemental Table 4b: Viral load linear growth model fixed effects estimates**

|  |  |  |  |
| --- | --- | --- | --- |
| **Study Visit** | **Standard Arm**  **Mean VL (95% CI)** | **MACARTI Arm**  **Mean VL (95% CI)** | **P-Value** |
| Baseline | 21,694 (10,373 - 45,370) | 25,443 (11,328 - 57,142) | 0.773 |
| 30 Days | 3,018 (1,524 - 5,973) | 3,286 (1,544 - 6,992) | 0.868 |
| 90 Days | 1,102 (596 - 2,036) | 671 (353 - 1276) | 0.269 |
| 180 Days | 525 (275 - 1,002) | 192 (100 - 372) | 0.031 |
| 365 Days | 228 (99 - 524) | 44 (18 - 107) | 0.008 |

1VL estimates and 95% CI back-transformed via exponentiation

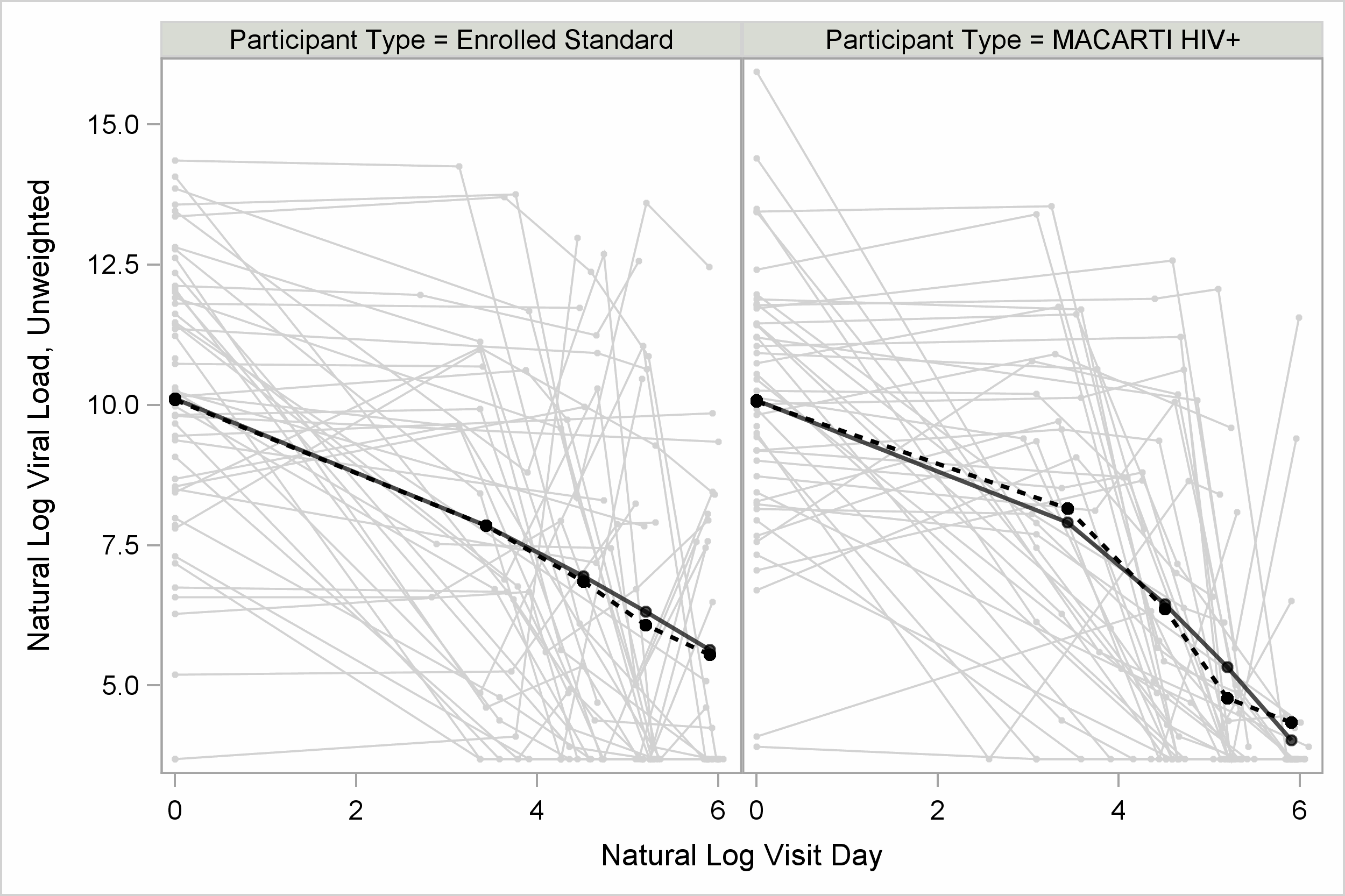
**Supplemental Table 4a** presents significant main effects for study arm and quadratic log-time. While insignificant, the interaction between quadratic log-time and study arm points to a marginal difference in growth trends between MACARTI and SOC and was retained in the analysis for least-squares means estimation. As with the CD4 count model, significant random effects are noted in the table footnote for both the intercept and slope. **Supplemental Figure 4b** shows unweighted, model-based least-squares mean values, connected and overlaid on participant-level transformed Viral Load observations (solid-line). Relative to the un-modeled means (dashed-line), the quadratic linear growth model appears to strongly describe the relationship between log-Viral Load and log-time across treatment arms. Finally, **Supplemental Figure 4c** shows weighted, model-based least-squares mean values connected and overlaid on predicted participant-level transformed Viral Load observations against log-time as described in the primary text.

**Supplemental Figure 4a: Participant-level Viral Load means (natural log) over study visit follow-up (natural log), unweighted**



1Study days 0, 30, 90, 180, 365 are presented as mean points and connected with a dashed-line

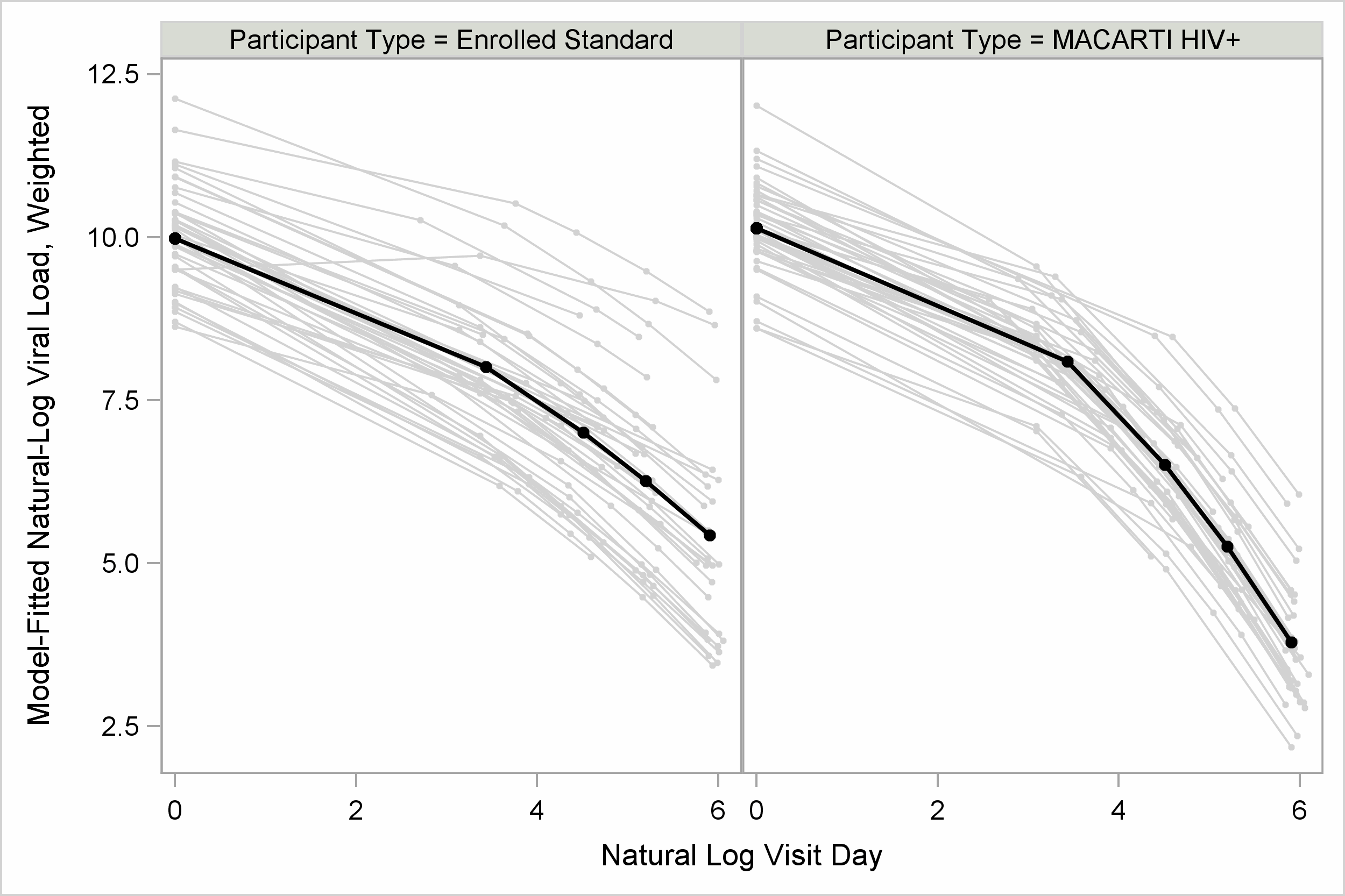
**Supplemental Figure 4b: Growth model-based mean Viral Load estimates (natural log) versus un-modeled mean estimates over log-time, unweighted**



1Study days 0, 30, 90, 180, 365 are presented as mean points and connected with a

dashed-line for un-modeled values and a connected-line for growth model estimates

**Supplemental Figure 4c: Growth model-based mean Viral Load estimates (natural log) and predicted patient-level observations over log-time, weighted**



1Weighted-growth model estimates for study days 0, 30, 90, 180, 365

**References**

1. Austin PC, Stuart EA. **Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies**. *Stat Med* 2015; 34(28):3661-3679.

2. PR. R. **Model-based direct adjustment.** *Journal of the American Statistical Association* 1987; 82:387–394.

3. Lee BK, Lessler J, Stuart EA. **Weight trimming and propensity score weighting**. *PLoS One* 2011; 6(3):e18174.

4. Austin PC. **Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples**. *Stat Med* 2009; 28(25):3083-3107.

5. Pirracchio R, Petersen ML, van der Laan M. **Improving propensity score estimators' robustness to model misspecification using super learner**. *Am J Epidemiol* 2015; 181(2):108-119.

6. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. **A tutorial on propensity score estimation for multiple treatments using generalized boosted models**. *Stat Med* 2013; 32(19):3388-3414.

7. Cohen J. **Statistical power analysis for the behavioral sciences (2nd ed.)***.* Lawrence Earlbaum Associates.; 1988.

8. Schluchter M, Elashoff J. **Small-sample adjustments to tests with unbalanced repeated measures assuming several covariance structures.** . *Journal of Statistical Computation and Simulation* 1990; 37(1-2):69-87.