**Supplemental Online Content**

Unger JM, Darke A, Othus M, et al. Effectiveness of adjuvant pembrolizumab vs high-dose interferon or ipilimumab for quality-of-life outcomes in patients with resected melanoma: a secondary analysis of the SWOF S1404 randomized clinical trial. *JAMA Oncol*. Published online November 23, 2022. doi:10.1001/jamaoncol.2022.5486



This supplemental material has been provided by the authors to give readers additional information about their work.

**eTable 1:** Observed and Fitted Group Mean Results for the FACT-BRM Scales

| | | Mean (95% CI) | | | | | | Fitted Difference (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|
| **FACT-BRM Scale** | **Cycle** | **Ipilimumab/HDI** | | | **Pembrolizumab** | | | **Pembrolizumab vs. ipilimumab/HDI** | |
| | | N | Baseline | Follow-Up | N | Baseline | Follow-Up | Difference | P-value |
| Trial Outcome Index | 1 | 489 | 91.9 (90.8 to 93.0) | 83.3 (81.6 to 85.0) | 603 | 91.9 (90.9 to 92.9) | 91.8 (90.7 to 92.9) | 8.5 (7.1 to 10.0) | <.001 |
| | 3 | 267 | 91.9 (90.5 to 93.4) | 80.6 (78.4 to 82.9) | 565 | 92.1 (91.0 to 93.2) | 90.3 (89.2 to 91.5) | 9.6 (7.9 to 11.3) | <.001 |
| | 5 | 183 | 91.6 (89.7 to 93.5) | 83.2 (80.5 to 85.9) | 486 | 92.1 (91.0 to 93.3) | 89.2 (87.9 to 90.6) | 5.6 (3.7 to 7.6) | <.001 |
| | 7 | 141 | 90.9 (88.7 to 93.2) | 83.2 (80.2 to 86.2) | 425 | 92.5 (91.2 to 93.7) | 90.2 (88.8 to 91.6) | 5.7 (3.7 to 7.7) | <.001 |
| | 9 | 104 | 90.5 (87.8 to 93.2) | 83.3 (79.8 to 86.8) | 387 | 92.8 (91.5 to 94.0) | 90.3 (88.8 to 91.7) | 5.0 (2.6 to 7.4) | <.001 |
| Total Score | 1 | 489 | 135.7 (134.1 to 137.2) | 127.4 (125.4 to 129.5) | 602 | 135.5 (134.1 to 136.9) | 136.7 (135.2 to 138.1) | 9.4 (7.7 to 11.2) | <.001 |
| | 3 | 267 | 135.6 (133.5 to 137.6) | 124.4 (121.5 to 127.2) | 563 | 135.7 (134.3 to 137.2) | 134.9 (133.3 to 136.4) | 10.4 (8.1 to 12.6) | <.001 |
| | 5 | 183 | 135.0 (132.4 to 137.6) | 126.2 (122.8 to 129.7) | 484 | 135.8 (134.2 to 137.3) | 133.6 (131.9 to 135.4) | 6.8 (4.3 to 9.4) | <.001 |
| | 7 | 141 | 133.8 (130.8 to 136.9) | 127.3 (123.4 to 131.1) | 424 | 136.3 (134.7 to 138.0) | 134.9 (133.1 to 136.8) | 5.6 (2.9 to 8.3) | <.001 |
| | 9 | 104 | 133.4 (129.9 to 136.9) | 126.8 (122.3 to 131.4) | 385 | 136.6 (134.9 to 138.3) | 135.0 (133.1 to 136.9) | 5.4 (2.3 to 8.6) | <.001 |
| Physical Subscale | 1 | 495 | 24.4 (24.1 to 24.7) | 21.9 (21.4 to 22.3) | 612 | 24.3 (24.0 to 24.6) | 24.0 (23.7 to 24.3) | 2.2 (1.7 to 2.7) | <.001 |
| | 3 | 271 | 24.4 (24.0 to 24.8) | 21.2 (20.6 to 21.8) | 575 | 24.4 (24.1 to 24.6) | 23.7 (23.4 to 24.0) | 2.5 (2.0 to 3.0) | <.001 |
| | 5 | 186 | 24.3 (23.8 to 24.9) | 21.5 (20.8 to 22.2) | 496 | 24.4 (24.1 to 24.7) | 23.4 (23.0 to 23.7) | 1.9 (1.3 to 2.5) | <.001 |
| | 7 | 143 | 24.2 (23.5 to 24.8) | 21.6 (20.8 to 22.5) | 433 | 24.4 (24.1 to 24.7) | 23.6 (23.2 to 23.9) | 1.7 (1.1 to 2.3) | <.001 |
| | 9 | 107 | 24.1 (23.4 to 24.9) | 21.6 (20.6 to 22.7) | 393 | 24.5 (24.2 to 24.8) | 23.4 (23.0 to 23.8) | 1.5 (0.7 to 2.2) | <.001 |
| Cognitive/Emotional | 1 | 496 | 19.3 (18.9 to 19.6) | 18.6 (18.2 to 19.0) | 612 | 19.1 (18.8 to 19.5) | 19.6 (19.3 to 19.9) | 1.1 (0.7 to 1.4) | <.001 |
| | 3 | 271 | 19.4 (19.0 to 19.9) | 17.9 (17.3 to 18.5) | 575 | 19.2 (18.9 to 19.5) | 19.2 (18.8 to 19.5) | 1.5 (1.0 to 2.0) | <.001 |
| | 5 | 185 | 19.4 (18.8 to 20.0) | 18.4 (17.7 to 19.1) | 496 | 19.2 (18.9 to 19.6) | 18.8 (18.4 to 19.2) | 0.6 (0.0 to 1.1) | .04 |
| | 7 | 143 | 19.2 (18.6 to 19.9) | 18.1 (17.3 to 18.8) | 433 | 19.3 (18.9 to 19.7) | 19.1 (18.6 to 19.5) | 1.0 (0.4 to 1.6) | .002 |
| | 9 | 107 | 19.4 (18.7 to 20.0) | 17.9 (17.0 to 18.8) | 393 | 19.4 (19.0 to 19.7) | 19.0 (18.5 to 19.4) | 1.1 (0.4 to 1.8) | .003 |

HDI = high dose interferon alpha 2b

**eTable 2:** Determination of best model fit for longitudinal data analyses

| | Row 1 = Fit statistic (-2 Log Likelihood), model degrees of freedom<br>Row 2 = Difference in fit statistic compared to specified nested model, difference in degrees of freedom, p-value by chi-square<br>Row 3 = Difference in fit statistic compared to alternative specified nested model, difference in degrees of freedom, p-value by chi-square | | | | |
|---|---|---|---|---|---|
| | Function for Time in Model[#] | | | | |
| Domain | Linear Time, No Interaction | Linear Time, Interaction with Treatment | Quadratic Time, No Interaction | Quadratic Time, Interaction with Treatment | Best Model |
| FACT-BRM-TOI | 26951.8, 9 | 26937.6, 10<br>Vs linear time: 14.2, 1, p<.001 | 26932.1, 10<br>Vs linear time: 19.7, 1, p<.001 | 26913.8, 12<br>Vs linear time, no interaction: 38, 3, p<.001<br>Vs linear time, interaction with treatment: 23.8, 2, p<.001<br>Vs quadratic time, no interaction: 18.3, 2, p<.001 | Quadratic Time w/ Interaction* |
| FACT-BRM Total Score | 28543.8, 9 | 28534.9, 10<br>Vs linear time: 8.9, 1, p=.003 | 28524.2, 10<br>Vs linear time: 19.6, 1, p<.001 | 28512.0, 12<br>Vs linear time, no interaction: 31.8, 3, p<.001<br>Vs linear time, interaction with treatment: 22.9, 2, p<.001<br>Vs quadratic time, no interaction: 12.2, 2, p=.002 | Quadratic Time w/ Interaction* |
| FACT-BRM Physical Subscale | 18842.8, 9 | 18838.7, 10<br>Vs linear time: 4.1, 1, p=.04 | 18831.0, 10<br>Vs linear time: 11.8, 1, p=.001 | 18824.0, 12<br>Vs linear time, no interaction: 18.8, 3, p<.001<br>Vs linear time, interaction with treatment: 14.7, 2, p=.001<br>Vs quadratic time, no interaction: 7.0, 2, p=.03 | Quadratic Time w/ Interaction* |
| FACT-BRM Cognitive/Emotional Subscale | 17993.9, 9 | 17993.5, 10<br>Vs linear time: 0.4, 1, p=.53 | 17982.0, 10<br>Vs linear time: 11.9, 1, p=.001 | 17981.0, 12<br>Vs linear time, no interaction: 12.9, 3, p=.005<br>Vs linear time, interaction with treatment: 12.5, 2, p=.002<br>Vs quadratic time, no interaction: 1.0, 2, p=.61 | Quadratic Time, No Interaction* |
| FACT-G Total Score | 26263.0, 9 | 26252.7, 10<br>Vs linear time: 10.3, 1, p=.001 | 26250.0, 10<br>Vs linear time: 13.0, 1, p<.001 | 26236.8, 12<br>Vs linear time, no interaction: 26.2, 3, p<.001<br>Vs linear time, interaction with treatment: 15.9, 2, p<.001<br>Vs quadratic time, no interaction: 13.2, 2, p=.001 | Quadratic Time w/ Interaction* |
| FACT-G Physical Well-Being | 19175.3, 9 | 19153.0, 10<br>Vs linear time: 22.3, 1, p<.001 | 19156.1, 10<br>Vs linear time: 19.2, 1, p<.001 | 19128.1, 12<br>Vs linear time, no interaction: 47.2, 3, p<.001<br>Vs linear time, interaction with treatment: 24.9, 2, p<.001<br>Vs quadratic time, no interaction: 28.0, 2, p<.001 | Quadratic Time w/ Interaction* |
| FACT-G Social Well-Being | 19595.0, 9 | 19594.6, 10<br>Vs linear time: 0.4, 1, p=.53 | 19592.7, 10<br>Vs linear time: 2.3, 1, p=.13 | 19592.1, 12<br>Vs linear time, no interaction: 2.9, 3, p=.41<br>Vs linear time, interaction with treatment: 2.5, 2, p=.29<br>Vs quadratic time, no interaction: 0.6, 2, p=.74 | Linear Time |
| FACT-G Emotional Well-Being | 17606.7, 9 | 17606.4, 10<br>Vs linear time: 0.3, 1, p=.58 | 17599.4, 10<br>Vs linear time: 7.3, 1, p=.007 | 17598.5, 12<br>Vs linear time, no interaction: 8.2, 3, p=.04<br>Vs linear time, interaction with treatment: 7.9, 2, p=.02<br>Vs quadratic time, no interaction: 0.9, 2, p=.64 | Quadratic Time |
| FACT-G Functional Well-Being | 20324.7, 9 | 20313.7, 10<br>Vs linear time: 11.0, 1, p=.001 | 20322.1, 10<br>Vs linear time: 2.6, 1, p=.11 | 20308.5, 12<br>Vs linear time, no interaction: 16.2, 3, p=.001<br>Vs linear time, interaction with treatment: 5.2, 2, p=.07<br>Vs quadratic time, no interaction: 13.6, 2, p=.001 | Linear Time w/ Interaction |

**e Table 2:** Determination of best model fit for longitudinal data analyses (continued)

| | Row 1 = Fit statistic (-2 Log Likelihood), model degrees of freedom<br>Row 2 = Difference in fit statistic compared to specified nested model, difference in degrees of freedom, p-value by chi-square<br>Row 3 = Difference in fit statistic compared to alternative specified nested model, difference in degrees of freedom, p-value by chi-square | | | | |
|---|---|---|---|---|---|
| FACIT-D Trial Outcome Index | 12531.3, 9 | 12524.1, 10<br>Vs linear time: 7.2, 1, p=.007 | 12516.7, 10<br>Vs linear time: 14.6, 1, p<.001 | 12505.5, 12<br>Vs linear time, no interaction: 25.8, 3, p<.001<br>Vs linear time, interaction with treatment: 18.6, 2, p<.001<br>Vs quadratic time, no interaction: 11.2, 2, p=.004 | Quadratic Time w/ Interaction* |
| FACIT-D Total Score | 13657.7, 9 | 13653.6, 10<br>Vs linear time: 4.1, 1, p=.04 | 13643.3, 10<br>Vs linear time: 14.4, 1, p<.001 | 13635.5, 12<br>Vs linear time, no interaction: 22.2, 3, p<.001<br>Vs linear time, interaction with treatment: 18.1, 2, p<.001<br>Vs quadratic time, no interaction: 7.8, 2, p=.02 | Quadratic Time w/ Interaction* |
| FACIT-D Diarrhea Subscale | 8967.7, 9 | 8965.7, 10<br>Vs linear time: 2.0, 1, p=.16 | 8966.2, 10<br>Vs linear time: 1.5, 1, p=.22 | 8963.5, 12<br>Vs linear time, no interaction: 4.2, 3, p=.24<br>Vs linear time, interaction with treatment: 2.2, 2, p=.33<br>Vs quadratic time, no interaction: 2.7, 2, p=.26 | Linear Time |
| EQ-5D Index Score | -6446.1, 9 | -6446.8, 10<br>Vs linear time: 0.7, 1, p=.40 | -6447.1, 10<br>Vs linear time: 1.0, 1, p=.32 | -6448.0, 12<br>Vs linear time, no interaction: 1.9, 3, p=.59<br>Vs linear time, interaction with treatment: 1.2, 2, p=.55<br>Vs quadratic time, no interaction: 0.9, 2, p=.64 | Linear Time |
| EQ-5D Global Health Score | 28005.4, 9 | 27996.8, 10<br>Vs linear time: 8.6, 1, p=.003 | 27996.3, 10<br>Vs linear time: 9.1, 1, p=003 | 27985.4, 12<br>Vs linear time, no interaction: 20.0, 3, p<.001<br>Vs linear time, interaction with treatment: 11.4, 2, p=.003<br>Vs quadratic time, no interaction: 10.9, 2, p=.004 | Quadratic Time w/ Interaction* |

# Model parameterizations:

Linear, no interaction = intercept, baseline domain score, treatment, time, and the stratification factors [stage, IIIA vs. IIIB vs. IIIC vs. IV; PD-L1 status, positive vs. negative vs. indeterminate]
Linear, interaction = intercept, baseline domain score, treatment, time, treatment*time, and the stratification factors [stage, IIIA vs. IIIB vs. IIIC vs. IV; PD-L1 status, positive vs. negative vs. indeterminate]
Square, no interaction = intercept, baseline domain score, treatment, time, time*time, and the stratification factors [stage, IIIA vs. IIIB vs. IIIC vs. IV; PD-L1 status, positive vs. negative vs. indeterminate]
Square, interaction = intercept, baseline domain score, treatment, time, treatment*time, time*time, treatment*time*time, and the stratification factors [stage, IIIA vs. IIIB vs. IIIC vs. IV; PD-L1 status, positive vs. negative vs. indeterminate]

* Although the p-value for the comparison with the linear time model (alone or interacting with treatment) is statistically significant, the absence of statistical significance compared to the quadratic time model suggests that the quadratic function for time dominates the difference in nested terms, rather than the interaction term, and indicates that the best model fit is the quadratic time, no interaction model.

**eTable 3:** Best model results for longitudinal analyses using linear mixed models

| | | **Model fitted estimates for treatment effect**<br>Treatment effect (95% CI)<br>P-value | | | | |
|---|---|---|---|---|---|---|
| | | **Cycle** | | | | |
| **Domain** | **Best Model** | **1** | **3** | **5** | **7** | **9** |
| **FACT-BRM** | | | | | | |
| Trial Outcome Index | Quadratic interaction | 8.8 (7.4-10.2) | 7.9 (6.5-9.2) | 6.7 (5.1-8.2) | 5.6 (3.9-7.2) | 4.6 (2.4-6.8) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| Total Score | Quadratic interaction | 9.7 (8.0-11.4) | 8.9 (7.2-10.6) | 7.8 (5.8-9.7) | 6.5 (4.4-8.6) | 5.2 (2.4-8.0) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| Physical Subscale | Quadratic interaction | 2.2 (1.8-2.7) | 2.2 (1.8-2.6) | 2.1 (1.6-2.5) | 1.8 (1.3-2.3) | 1.4 (0.7-2.0) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| Cognitive/Emotional Subscale | Quadratic time | 1.1 (0.7-1.4) | 1.1 (0.7-1.4) | 1.1 (0.7-1.4) | 1.1 (0.7-1.4) | 1.1 (0.7-1.4) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| **FACT-G** | | | | | | |
| Total Score | Quadratic interaction | 6.2 (5.0-7.4) | 5.4 (4.3-6.6) | 4.5 (3.2-5.9) | 3.7 (2.3-5.2) | 3.0 (1.1-5.0) |
| | | <.001 | <.001 | <.001 | <.001 | .002 |
| Physical Well Being | Quadratic interaction | 3.0 (2.5-3.4) | 2.5 (2.1-3.0) | 2.0 (1.5-2.6) | 1.7 (1.1-2.2) | 1.4 (0.7-2.0) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| Social Well Being | Linear | 0.4 (0.0-0.7) | 0.4 (0.0-0.7) | 0.4 (0.0-0.7) | 0.4 (0.0-0.7) | 0.4 (0.0-0.7) |
| | | .05 | .05 | .05 | .05 | .05 |
| Emotional Well Being | Quadratic time | 0.5 (0.2-0.8) | 0.5 (0.2-0.8) | 0.5 (0.2-0.8) | 0.5 (0.2-0.8) | 0.5 (0.2-0.8) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| Functional Well Being | Linear interaction | 2.2 (1.7-2.7) | 1.9 (1.5-2.4) | 1.6 (1.2-2.1) | 1.3 (0.7-1.8) | 0.9 (0.2-1.6) |
| | | <.001 | <.001 | <.001 | <.001 | .008 |
| **FACIT-D** | | | | | | |
| Trial Outcome Index | Quadratic interaction | 5.0 (3.7-6.3) | 4.8 (3.5-6.0) | 4.0 (2.5-5.6) | 2.9 (1.3-4.5) | 1.4 (-0.8-3.6) |
| | | <.001 | <.001 | <.001 | <.001 | .22 |
| Total Score | Quadratic interaction | 6.1 (4.3-7.9) | 6.1 (4.3-7.8) | 5.4 (3.3-7.5) | 4.0 (1.7-6.3) | 2.0 (-1.2-5.1) |
| | | <.001 | <.001 | <.001 | <.001 | .22 |
| Diarrhea Subscale | Linear | 0.6 (0.2-1.0) | 0.6 (0.2-1.0) | 0.6 (0.2-1.0) | 0.6 (0.2-1.0) | 0.6 (0.2-1.0) |
| | | .002 | .002 | .002 | .002 | .002 |
| **EQ-5D-3L** | | | | | | |
| Index Score | Linear | 0.03 (0.02-0.04) | 0.03 (0.02-0.04) | 0.03 (0.02-0.04) | 0.03 (0.02-0.04) | 0.03 (0.02-0.04) |
| | | <.001 | <.001 | <.001 | <.001 | <.001 |
| Global Health Score | Quadratic interaction | 5.7 (4.3-7.2) | 4.8 (3.5-6.2) | 3.8 (2.1-5.4) | 2.9 (1.2-4.6) | 2.1 (-0.2-4.5) |
| | | <.001 | <.001 | <.001 | <.001 | .07 |

**eTable 4:** Observed and Fitted Group Mean Results for the FACT-G Scales

| | | Mean (95% CI) | | | | | | Fitted Difference (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|
| **FACT-G Scales** | **Cycle** | **Ipilimumab/HDI** | | | **Pembrolizumab** | | | **Pembrolizumab vs. ipilimumab/HDI** | |
| | | N | Baseline | Follow-Up | N | Baseline | Follow-Up | Difference | P-value |
| Total Score | 1 | 497 | 92.0 (90.9 to 93.1) | 87.0 (85.6 to 88.3) | 606 | 92.0 (91.0 to 93.0) | 93.0 (92.0 to 94.0) | 6.0 (4.9 to 7.2) | <.001 |
| | 3 | 269 | 91.8 (90.3 to 93.2) | 85.5 (83.6 to 87.3) | 566 | 92.1 (91.1 to 93.1) | 91.9 (90.8 to 93.0) | 6.2 (4.6 to 7.8) | <.001 |
| | 5 | 185 | 91.1 (89.2 to 93.0) | 86.4 (84.0 to 88.8) | 486 | 92.0 (91.0 to 93.1) | 91.3 (90.1 to 92.5) | 4.3 (2.5 to 6.1) | <.001 |
| | 7 | 142 | 90.5 (88.3 to 92.7) | 87.7 (85.2 to 90.2) | 426 | 92.5 (91.3 to 93.6) | 92.2 (91.0 to 93.4) | 3.0 (1.1 to 5.0) | .002 |
| | 9 | 104 | 90.0 (87.4 to 92.6) | 87.3 (84.3 to 90.3) | 387 | 92.6 (91.4 to 93.8) | 92.4 (91.1 to 93.7) | 3.2 (0.9 to 5.5) | .006 |
| Physical Well-Being | 1 | 499 | 25.8 (25.6 to 26.1) | 22.3 (21.8 to 22.8) | 609 | 26.1 (25.8 to 26.3) | 25.4 (25.1 to 25.6) | 2.9 (2.4 to 3.4) | <.001 |
| | 3 | 270 | 25.8 (25.5 to 26.2) | 21.5 (20.8 to 22.2) | 571 | 26.1 (25.9 to 26.3) | 24.8 (24.5 to 25.1) | 3.2 (2.6 to 3.7) | <.001 |
| | 5 | 186 | 25.8 (25.3 to 26.2) | 22.6 (21.8 to 23.4) | 492 | 26.1 (25.9 to 26.4) | 24.6 (24.2 to 24.9) | 1.7 (1.1 to 2.3) | <.001 |
| | 7 | 142 | 25.7 (25.2 to 26.2) | 22.4 (21.6 to 23.3) | 429 | 26.2 (25.9 to 26.4) | 24.8 (24.4 to 25.1) | 2.0 (1.3 to 2.6) | <.001 |
| | 9 | 105 | 25.6 (25.0 to 26.2) | 22.8 (21.9 to 23.7) | 390 | 26.2 (26.0 to 26.5) | 24.8 (24.4 to 25.2) | 1.5 (0.8 to 2.2) | <.001 |
| Social/Family Well-Being | 1 | 503 | 25.0 (24.7 to 25.4) | 24.5 (24.1 to 24.9) | 616 | 24.7 (24.4 to 25.1) | 24.8 (24.4 to 25.1) | 0.5 (0.1 to 0.9) | .02 |
| | 3 | 273 | 24.9 (24.4 to 25.4) | 24.3 (23.8 to 24.9) | 578 | 24.7 (24.4 to 25.1) | 24.4 (24.0 to 24.8) | 0.1 (-0.4 to 0.7) | .50 |
| | 5 | 187 | 24.7 (24.1 to 25.3) | 23.7 (23.0 to 24.5) | 499 | 24.7 (24.3 to 25.1) | 24.4 (24.0 to 24.8) | 0.7 (0.0 to 1.3) | .05 |
| | 7 | 144 | 24.5 (23.7 to 25.3) | 24.2 (23.5 to 24.9) | 436 | 24.9 (24.4 to 25.3) | 24.4 (24.0 to 24.8) | 0.0 (-0.7 to 0.7) | .94 |
| | 9 | 107 | 24.5 (23.5 to 25.4) | 23.9 (23.1 to 24.8) | 394 | 24.8 (24.3 to 25.3) | 24.3 (23.9 to 24.8) | 0.3 (-0.6 to 1.2) | .52 |
| Emotional Well-Being | 1 | 501 | 18.7 (18.4 to 19.0) | 19.6 (19.3 to 19.9) | 616 | 18.7 (18.4 to 19.0) | 20.0 (19.8 to 20.3) | 0.4 (0.1 to 0.7) | .01 |
| | 3 | 273 | 18.7 (18.3 to 19.2) | 19.3 (18.9 to 19.8) | 576 | 18.8 (18.5 to 19.1) | 20.0 (19.7 to 20.2) | 0.6 (0.2 to 1.0) | .007 |
| | 5 | 187 | 18.6 (18.1 to 19.2) | 19.2 (18.7 to 19.8) | 498 | 18.7 (18.4 to 19.0) | 19.8 (19.5 to 20.1) | 0.6 (0.1 to 1.1) | .03 |
| | 7 | 145 | 18.5 (17.8 to 19.1) | 19.9 (19.3 to 20.4) | 435 | 18.8 (18.4 to 19.1) | 20.1 (19.8 to 20.4) | 0.0 (-0.5 to 0.6) | .87 |
| | 9 | 106 | 18.5 (17.7 to 19.2) | 19.6 (18.9 to 20.2) | 396 | 18.8 (18.4 to 19.1) | 20.2 (19.8 to 20.5) | 0.4 (-0.2 to 1.0) | .21 |
| Functional Well-Being | 1 | 501 | 22.4 (21.9 to 22.9) | 20.6 (20.0 to 21.1) | 616 | 22.4 (21.9 to 22.8) | 22.7 (22.3 to 23.1) | 2.2 (1.7 to 2.7) | <.001 |
| | 3 | 273 | 22.3 (21.7 to 22.9) | 20.2 (19.4 to 20.9) | 577 | 22.4 (22.0 to 22.8) | 22.6 (22.1 to 23.0) | 2.3 (1.7 to 2.9) | <.001 |
| | 5 | 187 | 22.0 (21.1 to 22.8) | 20.8 (19.9 to 21.7) | 498 | 22.3 (21.8 to 22.8) | 22.3 (21.9 to 22.8) | 1.4 (0.6 to 2.1) | <.001 |
| | 7 | 145 | 21.9 (20.9 to 22.8) | 21.2 (20.3 to 22.1) | 435 | 22.5 (22.0 to 23.0) | 22.7 (22.2 to 23.1) | 1.1 (0.4 to 1.9) | .003 |
| | 9 | 106 | 21.5 (20.3 to 22.7) | 21.0 (19.9 to 22.1) | 396 | 22.5 (22.0 to 23.1) | 22.8 (22.3 to 23.3) | 1.3 (0.4 to 2.2) | .005 |

HDI = high dose interferon alpha 2b

**eTable 5:** Observed and Fitted Group Mean Results for FACIT-D Scales

| | | Mean (95% CI) | | | | | | Fitted Difference (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|
| **FACIT-D Scale** | **Cycle** | **Ipilimumab/HDI** | | | **Pembrolizumab** | | | **Pembrolizumab vs. ipilimumab/HDI** | |
| | | N | Baseline | Follow-Up | N | Baseline | Follow-Up | Difference | P-value |
| Trial Outcome Index | 1 | 254 | 91.1 (90.1 to 92.0) | 86.0 (84.5 to 87.5) | 310 | 91.3 (90.4 to 92.2) | 91.0 (90.1 to 92.0) | 4.9 (3.6 to 6.2) | <.001 |
| | 3 | 132 | 91.0 (89.7 to 92.4) | 83.5 (81.3 to 85.6) | 289 | 91.3 (90.4 to 92.3) | 89.6 (88.4 to 90.7) | 5.8 (4.0 to 7.6) | <.001 |
| | 5 | 96 | 91.6 (90.1 to 93.0) | 85.7 (83.2 to 88.2) | 253 | 91.5 (90.5 to 92.5) | 89.6 (88.4 to 90.7) | 3.8 (1.9 to 5.8) | <.001 |
| | 7 | 69 | 91.6 (89.9 to 93.4) | 86.5 (83.8 to 89.2) | 215 | 92.0 (90.9 to 93.1) | 90.5 (89.3 to 91.7) | 3.7 (1.6 to 5.7) | <.001 |
| | 9 | 47 | 89.9 (87.4 to 92.5) | 86.3 (83.3 to 89.4) | 198 | 92.0 (90.9 to 93.1) | 90.4 (89.0 to 91.7) | 2.5 (-0.1 to 5.0) | .06 |
| Total Score | 1 | 254 | 134.8 (133.2 to 136.4) | 130.2 (128.1 to 132.2) | 309 | 134.8 (133.4 to 136.3) | 136.1 (134.6 to 137.5) | 6.0 (4.3 to 7.7) | <.001 |
| | 3 | 132 | 134.9 (132.8 to 137.0) | 127.3 (124.4 to 130.3) | 287 | 134.8 (133.3 to 136.3) | 134.2 (132.6 to 135.9) | 6.8 (4.3 to 9.4) | <.001 |
| | 5 | 96 | 135.4 (133.0 to 137.9) | 128.6 (124.9 to 132.3) | 251 | 135.1 (133.5 to 136.6) | 134.1 (132.3 to 136.0) | 5.8 (3.0 to 8.6) | <.001 |
| | 7 | 69 | 134.9 (132.0 to 137.9) | 131.0 (127.1 to 134.9) | 214 | 135.9 (134.2 to 137.6) | 135.4 (133.4 to 137.3) | 3.5 (0.5 to 6.6) | .02 |
| | 9 | 47 | 132.9 (129.1 to 136.7) | 129.8 (124.9 to 134.6) | 196 | 136.0 (134.2 to 137.8) | 135.6 (133.6 to 137.7) | 3.6 (-0.1 to 7.3) | .06 |
| Diarrhea Subscale | 1 | 258 | 42.9 (42.7 to 43.2) | 42.0 (41.6 to 42.5) | 313 | 42.8 (42.5 to 43.1) | 42.7 (42.5 to 43.0) | 0.8 (0.3 to 1.3) | <.001 |
| | 3 | 136 | 43.0 (42.6 to 43.3) | 41.9 (41.3 to 42.5) | 292 | 42.8 (42.5 to 43.1) | 42.5 (42.1 to 42.9) | 0.7 (0.0 to 1.4) | .05 |
| | 5 | 98 | 43.1 (42.8 to 43.5) | 42.1 (41.5 to 42.7) | 256 | 42.9 (42.6 to 43.2) | 42.6 (42.3 to 42.9) | 0.6 (0.0 to 1.2) | .06 |
| | 7 | 71 | 43.1 (42.7 to 43.5) | 42.0 (41.3 to 42.7) | 218 | 43.0 (42.7 to 43.3) | 42.6 (42.3 to 43.0) | 0.6 (-0.1 to 1.3) | .07 |
| | 9 | 49 | 42.7 (42.1 to 43.4) | 42.4 (41.8 to 43.0) | 200 | 43.0 (42.6 to 43.3) | 42.5 (42.1 to 43.0) | -0.1 (-0.9 to 0.8) | .88 |

HDI = high dose interferon alpha 2b

**eTable 6:** Observed and Fitted Group Mean Results for the EQ-5D-3L Index Score

| | | Mean (95% CI) | | | | | | Fitted Difference (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|
| Scale | **Cycle** | **Ipilimumab/HDI** | | | **Pembrolizumab** | | | **Pembrolizumab vs. ipilimumab/HDI** | |
| | | N | Baseline | Follow-Up | N | Baseline | Follow-Up | Difference | P-value |
| EQ-5D-3L Index Score | 1 | 476 | 0.89 (0.88 to 0.90) | 0.87 (0.86 to 0.89) | 587 | 0.90 (0.89 to 0.91) | 0.90 (0.90 to 0.91) | 0.03 (0.02 to 0.04) | <.001 |
| | 3 | 250 | 0.89 (0.88 to 0.91) | 0.86 (0.84 to 0.88) | 546 | 0.90 (0.89 to 0.91) | 0.90 (0.89 to 0.91) | 0.04 (0.02 to 0.06) | <.001 |
| | 5 | 172 | 0.90 (0.88 to 0.92) | 0.88 (0.87 to 0.90) | 470 | 0.90 (0.89 to 0.91) | 0.90 (0.89 to 0.91) | 0.01 (0.00 to 0.03) | .12 |
| | 7 | 135 | 0.90 (0.87 to 0.92) | 0.87 (0.85 to 0.89) | 415 | 0.90 (0.89 to 0.91) | 0.90 (0.89 to 0.91) | 0.03 (0.01 to 0.05) | .01 |
| | 9 | 101 | 0.89 (0.86 to 0.92) | 0.87 (0.84 to 0.90) | 376 | 0.90 (0.89 to 0.91) | 0.90 (0.89 to 0.91) | 0.02 (0.00 to 0.05) | .04 |
| Global Health Score | 1 | 482 | 83.4 (82.1 to 84.7) | 78.9 (77.3 to 80.5) | 602 | 84.6 (83.6 to 85.7) | 84.7 (83.7 to 85.7) | 5.2 (3.7 to 6.7) | <.001 |
| | 3 | 264 | 83.6 (81.9 to 85.2) | 77.1 (74.8 to 79.3) | 565 | 84.9 (83.8 to 86.0) | 84.6 (83.6 to 85.6) | 6.7 (4.9 to 8.5) | <.001 |
| | 5 | 181 | 82.9 (80.9 to 85.0) | 80.3 (78.0 to 82.7) | 487 | 84.7 (83.5 to 85.9) | 83.5 (82.2 to 84.7) | 2.0 (0.0 to 4.1) | .05 |
| | 7 | 140 | 82.1 (79.6 to 84.5) | 79.8 (77.2 to 82.5) | 430 | 85.1 (83.9 to 86.3) | 84.1 (82.8 to 85.4) | 2.6 (0.2 to 4.9) | .03 |
| | 9 | 106 | 83.0 (80.2 to 85.8) | 81.2 (78.3 to 84.0) | 388 | 84.8 (83.5 to 86.1) | 85.3 (83.9 to 86.7) | 3.1 (0.5 to 5.6) | .02 |

HDI = high dose interferon alpha 2b

**eMethods**: Pattern mixture model analysis

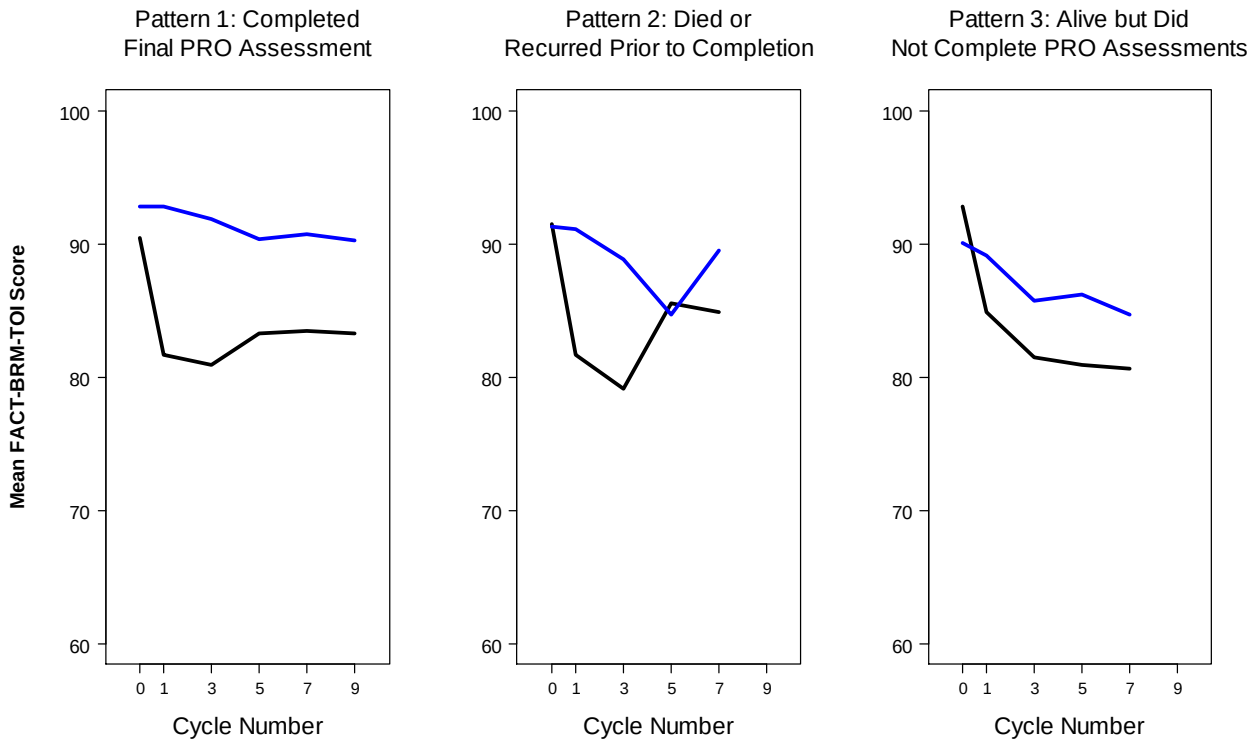
To address whether potential patterns of missing data influenced the primary analysis findings, we used pattern mixture models as a sensitivity analysis. Pattern mixture models condition on the type of missingness pattern through covariate adjustment in the regression model, with the appropriate missingness pattern determined by inspection of cohort plots.

Per protocol, we examined whether FACT-BRM TOI outcomes differed according to 6 potential patterns of missing data, by grouping patients according to whether patients:

- Completed their PRO assessments (through cycle 9)
- Died before the end of the cycle 9 PRO assessment
- Experienced disease recurrence before the end of the cycle 9 PRO assessment
- Were alive without disease recurrence before the end of the cycle 9 PRO assessment but did not complete 9 cycles of protocol therapy
- Were alive, recurrence-free, and completed 9 cycles of protocol therapy but stopped completing PRO assessments before Cycle 3
- Were alive, recurrence-free, and completed 9 cycles of protocol therapy but stopped completing PRO assessments after Cycle 3

Based on observed patterns (see Figure, below), 3 distinct patterns of missing data were identified, including 1) completed the final (cycle 9) PRO assessment; vs. 2) died or had recurrence prior to the completion of PRO assessments; vs. 3) were alive without recurrence but did not complete the PRO assessments through cycle 9.

**Figure:** Cohort Plots of Mean FACT-BRM-TOI Scores by Patterns of Missing Data by Arm



Adjusting for these patterns of missing data in a pattern mixture model, the best model fit for the FACT-BRM-TOI was a quadratic time model with interaction between treatment and time, as outlined in the table below:

**Table:** Determination of best model fit for longitudinal data analyses of FACT-BRM-TOI

| | Row 1 = Fit statistic (-2 Log Likelihood), model degrees of freedom<br>Row 2 = Difference in fit statistic compared to specified nested model, difference in degrees of freedom, p-value by chi-square<br>Row 3 = Difference in fit statistic compared to alternative specified nested model, difference in degrees of freedom, p-value by chi-square | | | | |
|---|---|---|---|---|---|
| | | | Function for Time in Model[#] | | |
| Domain | Linear Time, No Interaction | Linear Time, Interaction with Treatment | Quadratic Time, No Interaction | Quadratic Time, Interaction with Treatment | Best Model |
| FACT-BRM-TOI | 26926.0, 26 | 26900.0, 29<br>Vs linear time: 26.0, 3, p<.001 | 26896.0, 29<br>Vs linear time: 30.0, 3, p<.001 | 26865.8, 35<br>Vs linear time, no interaction: 60.2, 9, p<.001<br>Vs linear time, interaction with treatment: 34.2, 6, p<.001<br>Vs quadratic time, no interaction: 30.2, 6, p<.001 | Quadratic Time w/ Interaction* |

* Although the p-value for the comparison with the linear time model (alone or interacting with treatment) is statistically significant, the absence of statistical significance compared to the quadratic time model suggests that the quadratic function for time dominates the difference in nested terms, rather than the interaction term, and indicates that the best model fit is the quadratic time, no interaction model.

Using this model, the cycle-specific FACT-BRM TOI outcomes comparing pembrolizumab to ipilimumab/HDI are provided in the table below. Of note, for Cycle 3, the finding was X

**Table:** Cycle specific estimates of the effect of treatment on FACT-BRM TOI outcomes

| **Cycle No.** | **Estimate*** | **95% CI** | **p-value** |
|---|---|---|---|
| 1 | 9.7 | 7.1 to 12.2 | <.001 |
| 3 | 7.6 | 5.2 to 9.9 | <.001 |
| 5 | 5.7 | 3.2 to 8.1 | <.001 |
| 7 | 4.8 | 2.4 to 7.2 | <.001 |
| 9 | 4.9 | 2.2 to 7.6 | <.001 |

* Represents the added benefit of pembrolizumab compared to ipilimumab/HDI