



Published in final edited form as:

*Lancet Oncol.* 2017 March ; 18(3): 393–403. doi:10.1016/S1470-2045(17)30015-3.

## Phase III trial (EORTC 18071 / CA184-029) of post-operative adjuvant ipilimumab compared to placebo in patients with resected stage III cutaneous melanoma: Health Related Quality of Life (HRQoL) results

Corneel Coens<sup>1</sup>, Stefan Suciuc<sup>1</sup>, Vanna Chiarion-Sileni<sup>2</sup>, Jean-Jacques Grob<sup>3</sup>, Reinhard Dummer<sup>4</sup>, Jedd D. Wolchok<sup>5</sup>, Henrik Schmidt<sup>6</sup>, Omid Hamid<sup>7</sup>, Caroline Robert<sup>8</sup>, Paolo A. Ascierto<sup>9</sup>, Jon M. Richards<sup>10</sup>, Celeste Lebbe<sup>11</sup>, Virginia Ferraresi<sup>12</sup>, Michael Smylie<sup>13</sup>, Jeffrey S. Weber<sup>14</sup>, Michele Maio<sup>15</sup>, Srividya Kotapati<sup>16</sup>, Veerle de Prii<sup>17</sup>, Alessandro Testori<sup>18</sup>, and Alexander Eggermont<sup>8</sup>

<sup>1</sup>EORTC Headquarters, Brussels, Belgium <sup>2</sup>IOV-IRCCS, Melanoma Oncology Unit, Padova, Italy <sup>3</sup>Hôpital de la Timone, Marseille, France <sup>4</sup>University of Zürich Hospital, Zürich, Switzerland <sup>5</sup>Memorial Sloan-Kettering Cancer Center, New York, NY, United States <sup>6</sup>Aarhus University Hospital, Aarhus, Denmark <sup>7</sup>The Angeles Clinic and Research Institute, Los Angeles, CA, United States <sup>8</sup>Gustave Roussy Cancer Campus Grand Paris, Villejuif, France <sup>9</sup>Istituto Nazionale Tumori Fondazione “G. Pascale”, Naples, Italy <sup>10</sup>Oncology Specialists S.C., Park Ridge, IL, United States <sup>11</sup>Hôpital Saint-Louis, Paris, France <sup>12</sup>Istituti Fisioterapici Ospitalieri, Rome, Italy <sup>13</sup>Cross Cancer Institute, Edmonton, AB, Canada <sup>14</sup>H Lee Moffitt Cancer Center, Tampa, FL, United States <sup>15</sup>University Hospital of Siena, Istituto Toscano Tumori, Siena, Italy <sup>16</sup>Bristol-Myers Squibb, Wallingford, CT, United States <sup>17</sup>Bristol-Myers Squibb, Braine-l’Alleud, Belgium <sup>18</sup>European Institute of Oncology, Milan, Italy

### Abstract

**INTRODUCTION**—HRQoL was a secondary endpoint in the EORTC 18071 phase 3 trial in stage III melanoma patients comparing adjuvant Ipilimumab at 10 mg/kg (Ipi, N=475) versus placebo

Correspondence to: Corneel Coens.

#### CONFLICT OF INTEREST/DISCLOSURES

VCS has received personal fees from Bristol-Myers Squibb, GlaxoSmithKline, Roche, Merck, and Amgen for serving on an advisory board; J-JG has received personal fees and non-financial support from Bristol-Myers Squibb, Roche, GlaxoSmithKline, Merck, and Novartis for serving on an advisory board, lecturer, travel support, and grants from Bristol-Myers Squibb and GlaxoSmithKline; RD has received personal fees from Bristol-Myers Squibb, MSD, Roche, GlaxoSmithKline, and Novartis for serving on an advisory board; JDW has received other fees from Bristol-Myers Squibb, MSD, and MedImmune for serving as a consultant, and a grant from Bristol-Myers Squibb; OH has received personal fees from Bristol-Myers Squibb for lecturing and grants from Bristol-Myers Squibb; CR has received other fees from GlaxoSmithKline, Roche, Merck, Bristol-Myers Squibb, and Amgen for serving as a consultant; PAA has received personal fees from Bristol-Myers Squibb, Roche, Merck, GlaxoSmithKline, Ventana, Novartis, and Amgen for serving on an advisory board, and grants from Bristol-Myers Squibb, Roche, and Ventana; CL has received personal fees from Bristol-Myers Squibb, GlaxoSmithKline, Roche, Merck, Amgen, and Novartis for serving on an advisory board; MS has received personal fees from Bristol-Myers Squibb and Roche as honoraria and from Merck, Bristol-Myers Squibb, Roche, and GlaxoSmithKline for serving on an advisory board; JSW has received personal fees and grants from Bristol-Myers Squibb; MM has received personal fees from Bristol-Myers Squibb, Roche, MedImmune, and GlaxoSmithKline for serving on an advisory board and lecturing, and grants from Bristol-Myers Squibb and MedImmune; SK and VdP are employees and shareholders of Bristol-Myers Squibb. AE has received personal fees from Bristol-Myers Squibb, Amgen, Merck, and MedImmune for serving on an advisory board, and personal fee from GlaxoSmithKline for serving on a data and safety monitoring board; CC, SS, JMR, VF, HS, and AT have nothing to disclose.

(Pbo, N=476). The primary endpoint was recurrence-free survival [HR (95% CI) = 0.75 (0.64–0.90); P=0.0013]. Toxicity of Ipi consisted mainly of skin, gastrointestinal, endocrine and hepatic immune-related adverse events. Adjuvant treatment with Ipi was approved in October 2014 by the FDA based on the results of this trial.

**METHODS**—HRQoL was assessed using EORTC QLQ-C30 at baseline, week 4, 7, 10, 24; and every 12 weeks thereafter up to 2 years regardless of progression. Results were summarized per time point and per patient mean score during and after induction. A predefined threshold of 10 points was considered clinically meaningful. The primary HRQoL endpoint was the global health scale (GH) with the predefined hypothesis of no clinically relevant differences after induction between arms.

**RESULTS**—HRQoL compliance in this trial was 94% at baseline, >75% at week 24, 50% at year 2, with slightly higher rates in the Pbo arm. GH scores differed most at week 7 (77 vs 72) and week 10 (77 vs 70) with lower scores in the Ipi arm. Differences between treatment arms in diarrhea (8 vs 18) and insomnia (15 vs 26) were beyond 10 points at week 10. Patient mean GH scores during and after induction were statistically ( $p<.001$ ) but not clinically relevant between arms.

**CONCLUSIONS**—Despite increased toxicity which forced most patients off treatment during the induction phase of Ipi administration, overall HRQoL, as measured by the QLQ-C30, was similar between arms as no clinically relevant differences in GH scale were observed during or after induction. Clinical relevant deterioration for certain symptoms were observed at week 10 but after induction no clinically relevant differences remained.

## INTRODUCTION

Due to a rapidly increasing incidence rate and to sentinel node staging, more melanoma patients are diagnosed with stage III melanoma (positive regional lymph node(s), no distant metastases) than ever before (1, 2). Surgery, consisting of a regional lymph node dissection, followed by observation or radiotherapy and/or systemic adjuvant therapy depends on the extent of the lymph nodal involvement and differs in various countries (1). Stage III patients are at a very high risk of relapse (> 50% at 5 years), especially patients with Stage IIIB and Stage IIIC and in case of Stage IIIA when the tumor load diameter is greater than 1 mm. (3–5).

Ipilimumab is a recombinant, human monoclonal antibody which blocks cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4). It has led to improved overall survival (OS) in metastatic melanoma (6, 7). The EORTC 18071 (CA184-029), an international double-blind, randomised phase III trial, which compared the efficacy of ipilimumab 10 mg/kg in high risk patients with stage III cutaneous melanoma after having undergone a complete regional lymph node dissection, investigated HRQoL of the patients in the study in addition to recurrence-free survival (RFS) and OS. Additional details on the study design, treatment and conduct have been published (8).

This study demonstrated that RFS was significantly prolonged with ipilimumab (median 26.1 months) compared with placebo (median 17.1 months; HR=0.75; 95% CI 0.64–0.90, P<0.0013) (8). Most toxicities that were observed using Common Toxicity Criteria (CTC)

version 3, consisted of Ipi-induced immune-related adverse events (irAE) were observed, especially grade 3–4 gastrointestinal (15.9%), hepatic (10.6%) and endocrine (8.5%) irAEs. Although most toxicities were manageable, for 245 patients out of 471 (52%), Ipilimumab discontinuation was due to AEs..

A secondary objective of this study was to compare HRQoL outcomes between the two treatment arms, using the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire core 30 (EORTC QLQ-C30) (9). This patient-reported-outcome tool is composed of 30 questions, measuring various different aspects of health-related quality of life (HRQoL) specific to cancer. Its results allow to measure the patients self-assessed level of both disease and treatment burden so that improvements in RFS and a potential survival benefit can be placed in context. For this reason HRQOL was included as a secondary endpoint in the EORTC 18071 study.

## METHODS

### Patients

The EORTC 18071 (BMS CA184029) enrolled patients with complete and adequate resection of stage III melanoma. Histologically confirmed melanoma had to be metastatic to lymph nodes including stage IIIA melanoma (if N1a, at least 1 metastasis >1 mm), stage IIIB or stage IIIC (no in-transit metastasis). Full details of the eligibility criteria are reported elsewhere (8). All participating patients had to provide signed informed consent which included a description of the HRQoL assessments. The study protocol was approved by the local ethics committees and registered at ClinicalTrials.gov (NCT00636168) and EudraCT (2007-001974-10)

### Study Design

This was a randomized, double-blind, placebo-controlled, Phase III trial conducted across 18 countries worldwide. The primary objective was to determine whether post-operative adjuvant therapy with ipilimumab improves RFS (defined as the time between the date of randomization and the date of first recurrence (local, regional, distant metastasis) or death due to any cause) compared to placebo.

Secondary objectives included comparisons of distant metastasis-free survival (DMFS) overall survival (OS) and safety profiles. An additional objective was to compare HRQoL between treatment arms using the EORTC QLQ-C30 instrument. Patients were randomized in a 1:1 ratio to receive either ipilimumab 10 mg/kg or placebo every 3 weeks for 4 doses, then every 3 months for up to a maximum of 3 years. Treatment was discontinued early in case of disease progression, unacceptable toxicity, withdrawal of consent, or death.

As patients receiving Ipi might experience a temporary reduction in HRQoL, the main hypothesis was that there would be no clinically relevant differences between the two arms using the global QoL scale after the induction phase. A secondary objective was to evaluate the treatment effect on the various symptoms and functioning scales as treatment related side-effects may have a (temporary) negative influence on the health related domains of QoL of these patients.

## Assessments

The EORTC QLQ-C30 (Version 3.0) is a 30-item questionnaire that assesses HRQoL in cancer patients across nine multi-item scales: Global Health Status (GH/QoL), Physical Functioning, Role Functioning, Emotional Functioning, Cognitive Functioning, Social Functioning, Fatigue, Pain, and Nausea and Vomiting. It also contains single item measures of dyspnea, insomnia, anorexia, constipation, diarrhea, and financial impact (9). It has been psychometrically validated and translated into 93 languages (10). Patients provide their answers on a 4-point scale (from 1 “not at all” to 4 “very much”), except for the GH/QoL scale of the EORTC QLQ-C30, which has a 7-point scale (from 1 “very poor” to 7 “excellent”). A linear transformation is used to standardize the raw score, so that overall scores range from 0 to 100. For the EORTC QLQ-C30, a higher score in GH/QoL or a functioning scale represents a better level of quality of life and functioning; a higher score in a symptom scale represents a worse level of symptoms (11).

Administration of the HRQoL questionnaires followed the clinical assessment schedule of the trial: the EORTC QLQ-C30 was completed at baseline and at weeks 4, 7, 10, and 24, and every 12 weeks up to 2 years regardless of disease recurrence and treatment discontinuation. Guidelines for administering questionnaires were provided, ensuring standardization of HRQoL data by all personnel (12).

## Statistical Methods

A HRQoL-specific statistical analysis plan was put into place prior to starting the analyses. Prior to the analysis of any HRQoL outcome, the compliance of the received questionnaires was evaluated and the time windows were set accordingly. As there was indication of a delay shift (assessments tended to occur later than planned due to cumulative delays), the upper limit was extended by two weeks from week 10 onwards (appendix B). Compliance rates between the 2 arms were compared at each post-baseline time point and checked for informative patterns. The pre-specified primary HRQoL outcome was the GH/QoL scale of the EORTC QLQ-C30; the other scales were of secondary importance except for the financial difficulties item which was excluded completely from the analysis. Since the trial is overpowered for HRQoL differences and due to the multiple tests conducted, differences will be interpreted according to their magnitude rather than statistical significance alone. According to the work by Osoba et al (13) and King (14), changes in scores of 5–10 represent a small difference and 10–20 represent a moderate difference, with 10 points (on a 0–100 scale) being considered as the threshold for clinical relevant changes. All HRQoL analyses were performed according to the intention to treat (ie. all randomized patients compared according to the allocated treatment). The scores for all HRQoL domains and items in the two arms at each assessment time-point were presented descriptively (mean, standard deviation, median, quartiles), and graphically displayed over time.

Two summary scores were calculated per subject for each HRQoL scale: 1) the average score reported during induction (ADI; defined as the average of all outcomes received after the first day of the first induction treatment administration and on or before 21 days after the last day of actual induction treatment administration) and 2) the average score reported after induction (AAI; defined as the average of all outcomes received strictly after 21 days after

the last day of actual induction treatment administration but no later than 750 days afterwards). These two summary scores based on average per time interval (15–18) were motivated by the protocol objectives where different HRQoL treatment effects were expected with a HRQoL deterioration during the induction phase but no difference afterwards. Non-parametric rank-order tests at 2-sided significance level of 5% were used to test for statistically significant differences between the treatment arms for the ADI and AAI.

Sensitivity analyses to assess the robustness of the results were undertaken by comparing the change from baseline scores per time-point by treatment group, the change from baseline for the AAI and ADI summaries and repeating the main analysis after imputation of missing data. Imputation was done by predicting missing values from an explicit regression model that included time, treatment group, stratification factors (tumour stage and geographical region) and factors linked to the HRQoL outcomes (gender, age and ECOG performance score). In addition, a linear mixed effects model was constructed adjusting for age, gender and baseline HRQoL score with treatment, a time effect and time-treatment interactions as fixed effects and a patient specific random effect (appendix E). Score estimates, standard errors, associated confidence intervals and resulting test statistics were obtained from this model.

The relationship between clinical symptoms and patient-reported HRQoL was investigated post-hoc since the increase of side effects due to Ipi exposure did not translate in clinically worse overall HRQoL. In order to investigate this finding, patients were categorized according to whether they had a CTC adverse event for appetite loss, diarrhea, fatigue, insomnia and nausea/vomiting at the week 10 visit. The change from baseline score for the relevant QLQ-C30 scale was compared per treatment arm. The primary analysis was then replicated within each subgroup. All analyses were conducted using SAS v9.3 software (SAS Institute Inc. Cary, NC).

The requirements established by EORTC (19) and the CONSORT PRO guideline (20) for reporting HRQoL in RCT were used to report the HRQoL details in full.

## RESULTS

### Patients

A total of 951 patients (476 in the pbo arm, 475 in the Ipi arm) were randomized and represent the ITT population (Figure 1). Baseline demographics and clinical characteristics were well balanced between the two arms (Table 1; additional data reported elsewhere (8)).

### HRQoL completion rates and baseline scores

Completion rates for the HRQoL questionnaires in the ITT population are shown in Table 2. An evaluation of the compliance was done and due to delay shift, forms tended to be completed later than expected. Baseline forms were collected from 9 days prior to randomization until the day of randomization itself. The week 4 and 7 assessment could be completed up to 1 week earlier or later; the week 10 assessment up to 1 week earlier and 3 weeks later, and the week 24, 36, 48, 60, 72, 84, 96 and 108 assessments up to 3 weeks earlier and 5 weeks later. An overview is presented in Appendix A. Compliance with

HRQoL assessments was good, with 94% of the patients completing the questionnaires at baseline. Compliance rates slowly decreased over time with the lowest reported at week 108 (55% and 47% for the placebo and ipilimumab arm respectively) and tended to be higher in the placebo arm throughout. The mean and median HRQoL scores at baseline for the scales of the EORTC QLQ-C30 were comparable between the two treatment arms. Overall, when compared to other studies, the compliance rates of our trial were within acceptable limits (21).

### Primary analysis

The overall test for treatment difference between the two treatment arms for both the ADI and AAI in the GH/QoL scale were statistically significant but did not exceed the clinically relevant threshold (table 3). Differences in the GH/QoL scale between the two treatment arms assessed at each of the post-baseline assessments did not reach the pre-determined clinically relevant difference of 10 points (Table 4). The maximum observed difference occurred at week 10 and was 7.89 points (95% CI: 4.93; 10.85) indicating a lower score in the Ipi arm. In both arms, the mean GH/QoL scores remained stable over time.

### Missing data mechanism

An investigation into the reported reasons for missing data revealed that the main documented reason was administrative failure (either by patient or staff), accounting for 55% of all reported reasons (962/1752). Missing data was found to be related to treatment arm (higher in the Ipi arm), time (higher at later time points), American Joint Committee on Cancer (AJCC) stage (higher at higher stage) and geographical region (higher in Europe and Australia) but not to performance status, gender or age. There is no indication that patients' GH/QoL scores were systematically lower at their last completed assessment.

### Sensitivity analyses

Replication of the primary analysis adjusted for baseline score and imputing missing data yielded similar results, confirming the robustness of the findings. For the GH/QoL scale, changes from baseline summarized at each time point via a longitudinal linear mixed model adjusting for age, gender and baseline QoL score are presented graphically in figure 2. The results confirm the findings of the primary analysis, despite a strong statistical treatment effect, the observed differences remain within the clinical acceptable limits with the largest difference noted at week 10. When the longitudinal modelling was applied to the other scales of the EORTC QLQ-C30, those with significantly different scores between the two treatment arms, both statistically and clinically, were diarrhea, insomnia and fatigue (figure 3). Fatigue had a clinically relevant deterioration at week 10 compared to baseline in the Ipi arm; Insomnia had a clinically relevant deterioration at week 10 in the Ipi arm compared to placebo; Diarrhea had both a clinically relevant deterioration in the Ipi arm at week 10 when compared to baseline and when compared to the placebo arm at week 10. A systematic trend with worse outcomes in the Ipi arm peaking around week 10 was found for most other symptom and functioning scales, but none of these were clinically relevant. Results for all secondary scales are provided in appendix D.

### Exploratory post-hoc analyses

Exploratory analysis of the symptoms revealed that on average higher symptom ratings were reported when the clinical symptom had been reported during the induction period (see Table 5). The only exception is insomnia in the Ipi arm which may be due to underreporting of the AE and/or the low resulting numbers. For most of the scales the average difference between the group with versus without the AE event exceeds the 10 points clinical relevance threshold.

## DISCUSSION

HRQoL was an important secondary endpoint of the 18071 trial. The results obtained from the pre-specified analysis demonstrated that Ipi had no significant impact on the GH/QoL scale after induction. Various sensitivity analyses (using different populations, methodology or outcomes) supported the primary analysis, showing that Ipi can be administered in this patient population without clinical relevant deterioration in HRQoL as measured by the QLQ-C30.

Significantly worse outcomes in the Ipi arm were found for specific symptoms scales namely diarrhea, insomnia and fatigue. The most common reported severe irAE were gastrointestinal, hepatic and endocrine (8), and are similar to other trial reports (22–24).

The observed toxicity due to Ipi exposure places a major burden on the patient and, if untreated, can be severe and life threatening. With rigorous monitoring by the treating staff, these can be diagnosed in a timely manner and effectively managed. (25). In this trial, 54% of IPI patients experienced a grade 3–4 AEs, as compared to 25% in the placebo arm. Therefore it is remarkable that these side effects did not translate into worse patient reported GH/QoL compared to the placebo arm. Especially since this trial is placebo controlled, patients may experience side effects more positively as a confirmation of treatment efficacy. Post-hoc exploratory analyses revealed that the occurrence of side effects did result in a HRQoL decrease by the patient. However the absolute number of patients experiencing these specific side effects was relatively limited. Therefore the overall effect was diluted in the average profiles. Exploratory factor analysis (results not shown) revealed that the overall HRQoL endpoint, GH/QoL, was more related to the functional scales (except cognitive functioning) than the symptom scales (except for fatigue). This may indicate that patients value their ability to continue daily tasks more than the physical discomforts of the side-effects. However it should be noted that these results stem from an exploratory post-hoc analysis and may not be reliable. Results from a maintenance subgroup analyses (appendix C) suggest no evidence that maintenance treatment resulted in worse HRQoL outcomes due to longer treatment administration. However these analyses need to be interpreted with caution as early progression patients progressing early being allocated to the non-maintenance group.

Currently, there is limited information in the literature on the HRQoL impact of Ipi administration in melanoma. Revicki et al (26) did report HRQoL results from a double-blind, fixed dose study in 676 previously treated advanced unresectable stage III or IV melanoma patients. The results confirm our main findings with Ipi resulting in little to no

impairment in general HRQL despite an increase in severe adverse events. These adverse events are severe but largely transient, so that no lasting HRQoL differences are observed after induction (27).

Our study was not without limitations. A common challenge to HRQoL in clinical trials is missing data. Compliance in this trial was good and remained within acceptable limits to allow the analyses to be performed as intended. Despite the observed lower compliance and increased toxicity in the Ipi arm, no evidence of a relationship between increased toxicity and missing data was found. Sensitivity analyses confirmed the results and an investigation into the causes of missing data revealed no systematic bias. However, bias due to selective missing data can never be ruled out completely. This study was designed to detect an improvement in RFS and was therefore overpowered for the HRQoL endpoint with even small HRQoL treatment differences being statistically significant. We have mitigated this by using an absolute 10 point difference as benchmark to declare clinically relevant treatment effects (13, 14). The use of the 10-point threshold itself might be nuanced as work by Cocks et al validated the original cut-offs but advocated a smaller effect size of 4 points for treatment group comparisons in randomized clinical trials (28). Such lower threshold would result in several clinically relevant treatment differences for the primary GH/QoL scale, most notably for the ADI and AAI summary scores. Therefore, the observed deteriorations in the Ipi arm cannot be dismissed as trivial and should be interpreted with care.

Although the QLQ-C30 is one of the most commonly used and validated measures applied within the oncology clinical trials setting, when applied to this particular study, no immunotherapy specific validation exists and several symptoms common to irAE are lacking. Most notably absent are symptoms related to endocrine (hypothyroiditis; hypophysitis) or skin reactions which can represent significant limitations for the patient that are not always clinically apparent. However the QLQ-C30 core questionnaire has been successfully used to detect clinically relevant treatment differences in a melanoma specific population (29).

## CONCLUSION

HRQoL was an important secondary endpoint in this study in which Ipilimumab and placebo were compared. Its primary selected scale, the EORTC QLQ-C30 GH/QoL score, showed no clinically relevant differences (10 points or more) between the two treatment arms at any time point, confirmed by sensitivity analyses. We observed ipi-induced clinically relevant worsening for diarrhea and fatigue at week 10, and ipi-induced clinically relevant treatment differences for diarrhea and insomnia at week 10. After induction, no relevant treatment differences are observed. This study shows that Ipilimumab can achieve an improvement in RFS with little impairment in HRQoL as measured by the QLQ-C30 despite severe side effects.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.



## Acknowledgments

**FUNDING:** This study was sponsored by Bristol-Myers-Squibb.

We would like to thank all the patients who participated and all site staff involved in the trial. All participating investigators who contributed to the HRQoL part of this trial are listed in the appendix A. We also thank all EORTC Headquarters team members who have not been included among the co-author list of this publication, and who contributed to the study success: Isabelle Blangenois, Sandra Collette, Gaetan de Schaetzen, Valérie Dewaste, Thierry Gorlia, Sven Janssen, Niels Lema, Larissa Polders, Simon Vanderschaeghe. We would like to warmly thank the Bristol-Myers Squibb team members as well. This study was sponsored by Bristol-Myers Squibb.

### Role of the funding source

Study investigators of both the EORTC and Bristol-Myers Squibb were involved in writing the report and in the decision to submit for publication. Bristol-Myers Squibb employees (listed as authors) were involved in interpretation and writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## References

1. Eggermont A, Spatz A, Robert C. Cutaneous Melanoma. *Lancet*. 2014; 383:816–27. [PubMed: 24054424]
2. Balch CM, Gershenwald JE, Soong S-J, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol*. 2009; 27:6199–206. [PubMed: 19917835]
3. Balch CM, Gershenwald JE, Soong S-J, et al. Multivariate analysis of prognostic factors among 2,313 patients with stage III melanoma: comparison of nodal micrometastases versus macrometastases. *J Clin Oncol*. 2010; 28:2452–9. [PubMed: 20368546]
4. van der Ploeg AP, van Akkooi AC, Rutkowski P, et al. Prognosis in patients with sentinel node-positive melanoma is accurately defined by the combined Rotterdam tumor load and Dewar topography criteria. *J Clin Oncol*. 2011 Jun 1; 29(16):2206–14. [PubMed: 21519012]
5. van der Ploeg AP, van Akkooi AC, Haydu LE, et al. The prognostic significance of sentinel node tumour burden in melanoma patients: an international, multicenter study of 1539 sentinel node-positive melanoma patients. *Eur J Cancer*. 2014 Jan; 50(1):111–20. [PubMed: 24074765]
6. Hodi FS, O’Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010; 363:711–23. [PubMed: 20525992]
7. Robert C, Thomas L, Bondarenko I, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N Engl J Med*. 2011; 364:2517–26. [PubMed: 21639810]
8. Eggermont AM, Chiarion-Sileni V, Grob JJ, et al. Adjuvant ipilimumab versus placebo after complete resection of high-risk stage III melanoma (EORTC 18071): a randomised, double-blind, phase 3 trial. *Lancet Oncol*. 2015 May; 16(5):522–30. [PubMed: 25840693]
9. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993 Mar 3; 85(5):365–76. [PubMed: 8433390]
10. De Wolf L, Koller M, Velikova G, et al. EORTC translating procedures. Brussels: Quality of Life Study Group Publications. EORTC Publication; 2009.
11. Fayers P, Aaronson N, Bjordal K, et al. EORTC QLQ-C30 Scoring Manual. 3. Brussels: EORTC Publications; 2001.
12. Young T, de Haes H, Fayers P, et al. Guidelines for assessing quality of life in clinical trials. Brussels: EORTC Quality of Life Study Group Publications; 1999.
13. Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998; 16:139–144. [PubMed: 9440735]
14. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res*. 1996; 5:555–567. [PubMed: 8993101]
15. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *British Medical Journal*. 1990; 300:230–235. [PubMed: 2106931]

16. Fairclough, DL. Design and Analysis of Quality of Life Studies in Clinical Trials. Boca Raton: Chapman & Hall/CRC; 2002.
17. Fayers, P., Hays, R. Assessing Quality of Life in Clinical Trials: Methods and Practice. Oxford University Press; New York: 2005.
18. Young, T., de Haes, H., Curran, D., Fayers, P., Brandberg, Y., Vanvoorden, A., Bottomley, V. on behalf of the EORTC Quality of Life Group. EORTC Manual. Brussels: EORTC; 2002.
19. Efficace F, Bottomley A, Osoba D, et al. Beyond the development of Health-Related Quality of Life (HRQOL) Measures: A Checklist for evaluation HRQOL Outcomes in Cancer Clinical Trials- Does HRQOL evaluation in prostate cancer research inform clinical decision making. *Journal of Clinical Oncology*. 2003 Sep 15; 21(18):3502–11. [PubMed: 12972527]
20. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *JAMA*. 2013; 309(8):814–822. [PubMed: 23443445]
21. Bottomley A, Aaronson NK. European Organisation for Research and Treatment of Cancer. International perspective on health-related quality-of-life research in cancer clinical trials: the European Organisation for Research and Treatment of Cancer experience. *J Clin Oncol*. 2007 Nov 10; 25(32):5082–6. [PubMed: 17991925]
22. Lebbé C, Weber JS, Maio M, et al. Survival follow-up and ipilimumab retreatment for patients with advanced melanoma who received ipilimumab in prior phase II studies. *Ann Oncol*. 2014 Nov; 25(11):2277–84. [PubMed: 25210016]
23. Horvat TZ, Adel NG, Dang T, et al. Immune-Related Adverse Events, Need for Systemic Immunosuppression, and Effects on Survival and Time to Treatment Failure in Patients With Melanoma Treated With Ipilimumab at Memorial Sloan Kettering Cancer Center. *J Clin Oncol*. 2015 Oct 1; 33(28):3193–8. [PubMed: 26282644]
24. Wolchok JD, Neyns B, Linette G, et al. Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. *Lancet Oncol*. 2010; 11:155–64. [PubMed: 20004617]
25. Weber JS, Dummer R, de Pril V, et al. Patterns of onset and resolution of immune-related adverse events of special interest with ipilimumab: detailed safety analysis from a phase 3 trial in patients with advanced melanoma. *Cancer*. 2013 May 1; 119(9):1675–82. [PubMed: 23400564]
26. Revicki DA, van den Eertwegh AJ, Lorigan P, et al. Health related quality of life outcomes for unresectable stage III or IV melanoma patients receiving ipilimumab treatment. *Health Qual Life Outcomes*. 2012 Jun 13; 10:66. [PubMed: 22694829]
27. Johnson DB, Friedman DL, Berry E, et al. Survivorship in Immune Therapy: Assessing Chronic Immune Toxicities, Health Outcomes, and Functional Status among Long-term Ipilimumab Survivors at a Single Referral Center. *Cancer Immunol Res*. 2015 May; 3(5):464–9. [PubMed: 25649350]
28. Cocks K, King MT, Velikova G, Martyn St-James M, Fayers PM, Brown JM. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol*. 2011 Jan 1; 29(1):89–96. [PubMed: 21098316]
29. Bottomley A, Coens C, Suci S, et al. Adjuvant therapy with pegylated interferon alfa-2b versus observation in resected stage III melanoma: a phase III randomized controlled trial of health-related quality of life and symptoms by the European Organisation for Research and Treatment of Cancer Melanoma Group. *J Clin Oncol*. 2009 Jun 20; 27(18):2916–23. [PubMed: 19433686]
30. Corrie PG, Marshall A, Dunn JA, Middleton MR, Nathan PD, Gore M, Davidson N, Nicholson S, Kelly CG, Marples M, Danson SJ, Marshall E, Houston SJ, Board RE, Waterston AM, Nobes JP, Harries M, Kumar S, Young G, Lorigan P. Adjuvant bevacizumab in patients with melanoma at high risk of recurrence (AVAST-M): preplanned interim results from a multicentre, open-label, randomised controlled phase 3 study. *Lancet Oncol*. 2014 May; 15(6):620–30. [PubMed: 24745696]
31. Reuter K, Albrecht K, Seelig H, Meiss F, Mauch C, Kreuzberg N, Nashan D. Health-related quality of life, fatigue, and depression under low-dose IFN- $\alpha$  therapy in melanoma patients. *J Immunother*. 2014 Nov-Dec; 37(9):461–7. [PubMed: 25304729]

32. Mohr P, Hauschild A, Trefzer U, Enk A, Tilgen W, Loquai C, Gogas H, Haalck T, Koller J, Dummer R, Gutzmer R, Brockmeyer N, Hölzle E, Sunderkötter C, Mauch C, Stein A, Schneider LA, Podda M, Göppner D, Schadendorf D, Weichenthal M. Intermittent High-Dose Intravenous Interferon Alfa-2b for Adjuvant Treatment of Stage III Melanoma: Final Analysis of a Randomized Phase III Dermatologic Cooperative Oncology Group Trial. *J Clin Oncol*. 2015 Dec 1; 33(34):4077–84. [PubMed: 26503196]
33. Brandberg Y, Aamdal S, Bastholt L, Hernberg M, Stierner U, von der Maase H, Hansson J. Health-related quality of life in patients with high-risk melanoma randomised in the Nordic phase 3 trial with adjuvant intermediate-dose interferon alfa-2b. *Eur J Cancer*. 2012 Sep; 48(13):2012–9. [PubMed: 22196968]

## Research in context

### Evidence before this study

There is a large unmet medical need in patients with resectable advanced stage III melanoma, who are at high-risk of disease recurrence. Therapeutic interventions have to balance efficacy with toxicity in this relatively young patient population (median ~50 years) that is rapidly increasing in incidence. An international randomized controlled trial of Ipilimumab versus placebo in stage III cutaneous melanoma patients (excluding lymph node metastasis  $\geq 1$  mm or in-transit metastasis) showed that progression-free survival was significantly improved at the cost of increased toxicity (8). If Ipilimumab has a substantial impact on health-related quality of life, this might influence the choice to administer this immunotherapy in this population. We searched PubMed and our search terms were: ((melanoma AND adjuvant) AND (quality of life OR patient-reported-outcomes) AND (adult) AND (immunotherapy OR antibody)). We did not use language restrictions in our search. Through this search we found only 10 references. These references included only article pertaining to a randomized clinical trial, namely interim results of the AVAST-M trial which compares bevacizumab versus observation as adjuvant treatment for patients with resected melanoma at high risk of recurrence (30). They reported no significant differences between the treatment groups for any of the quality-of-life scales over 36 months despite a tripling of the incidence of grade 3–4 adverse events in the bevacizumab arm. Expanding our search by dropping the (immunotherapy OR antibody) restriction, immediately increased the number of found references to 54. Most of the trial reports among these were on interferon or radiotherapy treatment. Quality of life results from low, intermediate and high dose studies (29, 31–33) indicated that a negative impact on overall HRQoL could be observed using the QLQ-C30. These effects were mostly reversible after end of treatment. Unfortunately, no study could be identified that directly compared the impact of ipilimumab on HRQOL in stage III melanoma patients.

### Added value of this study

To our knowledge, our study is the first randomized international blinded evaluation of the impact on HRQOL of ipilimumab as adjuvant treatment in stage III cutaneous melanoma patients. Our analysis demonstrates no clinically relevant deterioration in global HRQOL due to Ipilimumab administration but there was clinically relevant deterioration for specific symptoms at specific timepoints.

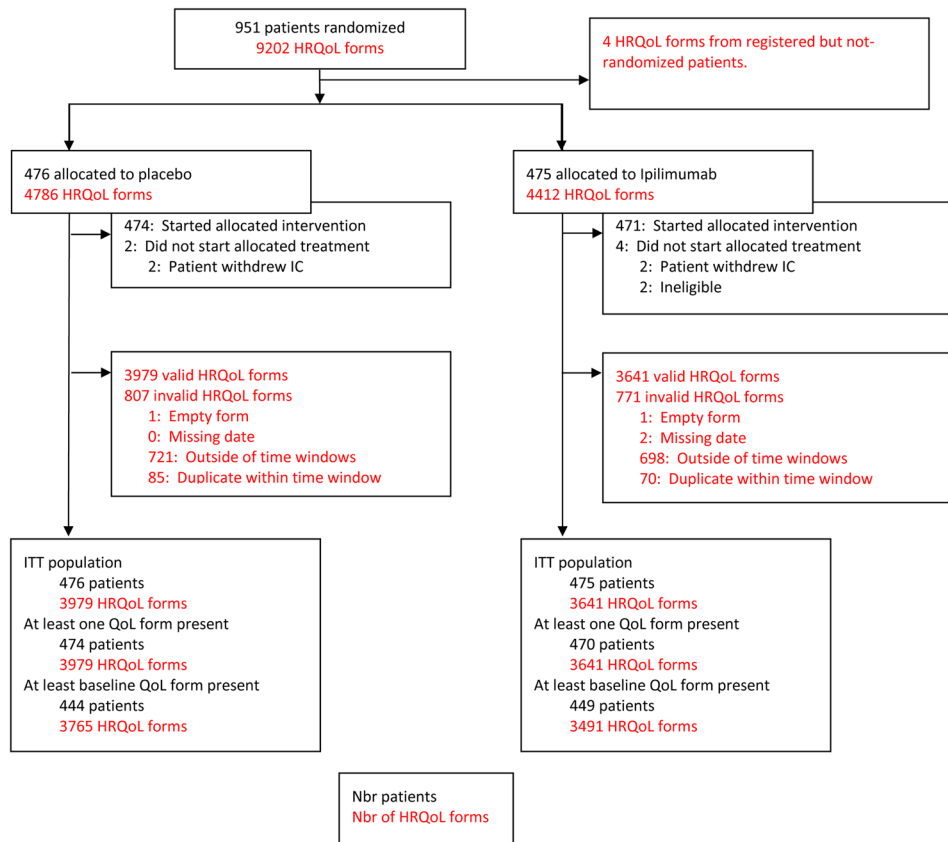
### Implications of all the available evidence

The HRQOL results support the overall clinical findings that Ipilimumab can achieve efficacy improvement at the cost of higher incidences of specific severe adverse events.

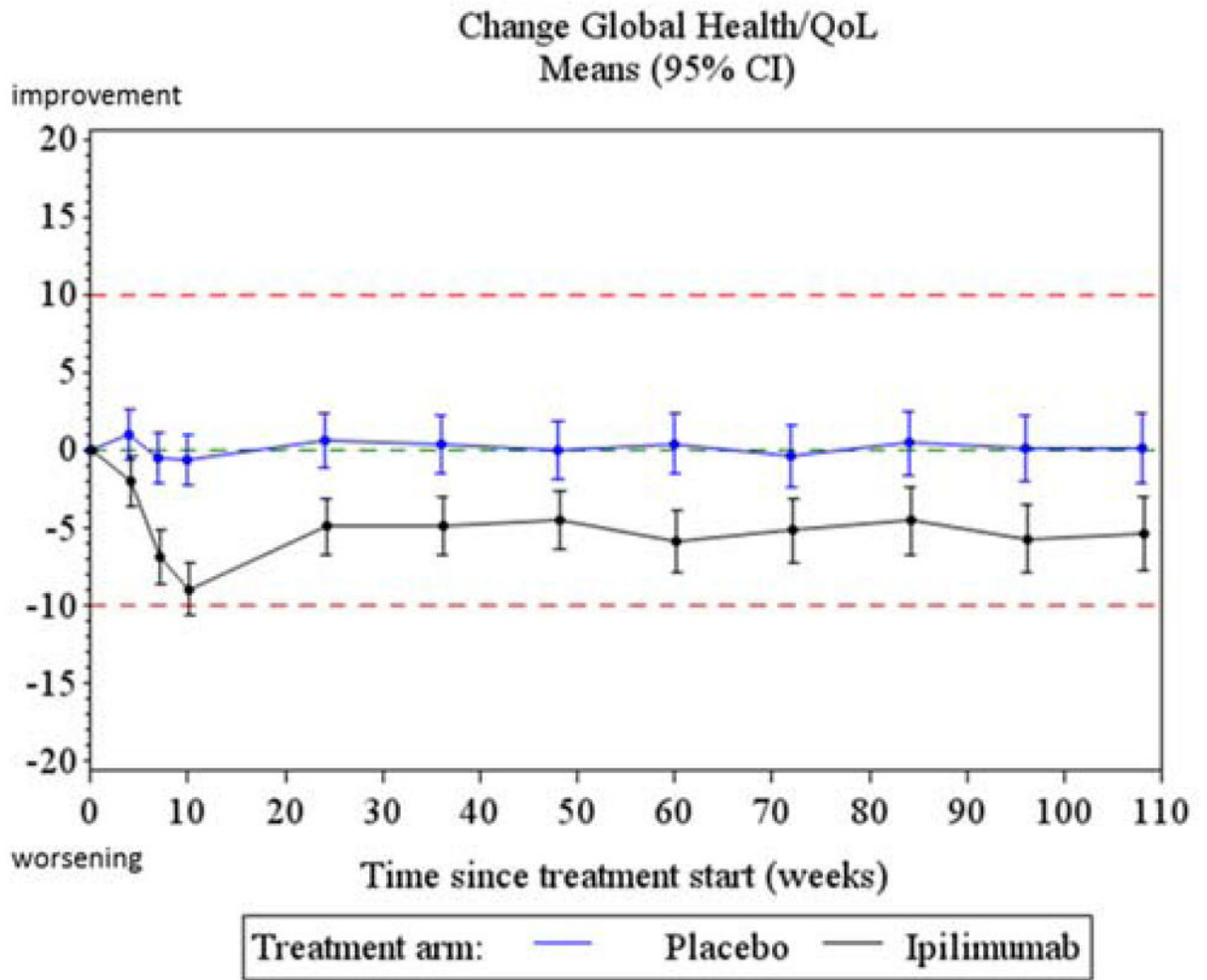
### Oversight

The protocol was approved by the EORTC protocol review committee and independent ethics committees. The study was conducted in accordance with the ethical principles originating from the Declaration of Helsinki and with Good Clinical Practice as defined by the International Conference on Harmonization. All participating patients gave written informed consent. The trial was designed jointly by the study coordinator, the EORTC Headquarters Melanoma team and the sponsor, Bristol-Myers Squibb. Data were collected and computerized at the EORTC Headquarters, and were transferred to the sponsor before the data-base lock. Data were analyzed independently at both the EORTC Headquarters and by the sponsor. An initial draft of this manuscript was prepared by the first author and reviewed by the study coordinator. All the authors contributed to subsequent drafts and made the decision to submit the manuscript for publication.

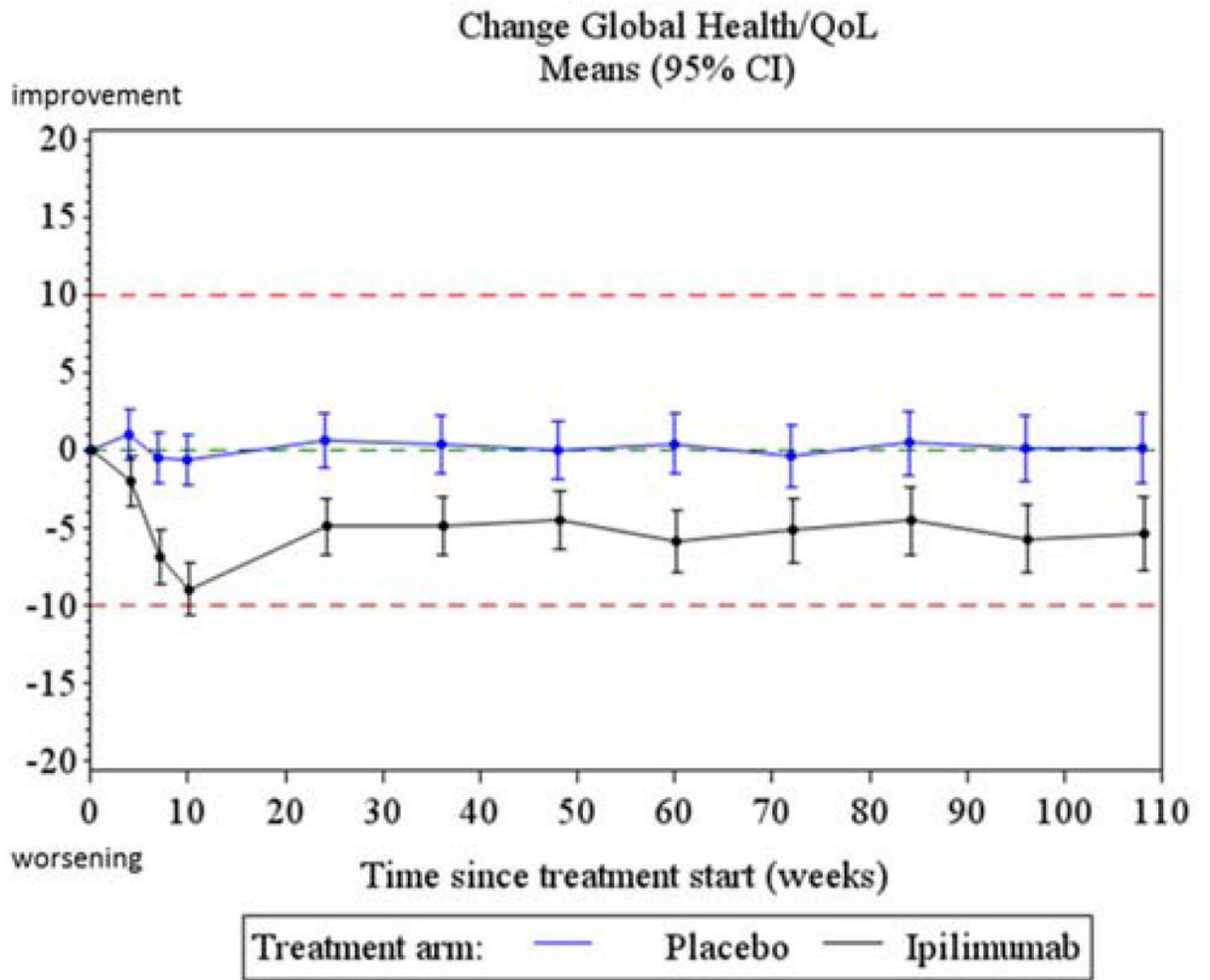
CC and SS participated in the study design, data collection, data analysis, data interpretation, and writing the manuscript; VC-S participated in writing the manuscript, and had a major contribution in patient accrual; J-JG participated in the data interpretation and writing the manuscript, and had a major contribution in patient accrual; RD, JDW, HS, OH, CR, PAA, JMR, CL, VF, MS, JSW and MM participated in writing the manuscript, and had a major contribution in patient accrual; SK participated in data interpretation and writing the manuscript; VdP participated in data collection, data analysis, data interpretation, and writing the manuscript; AT and AE participated in the study design, data collection, and writing the manuscript and had a major contribution in patient accrual.



**Figure 1.**  
CONSORT diagram



**Figure 2.**  
Treatment profiles for change from baseline in GH/QoL



**Figure 3.**  
Treatment profiles for change from baseline in Diarrhea, Insomnia and Fatigue



**Table 1**

Baseline clinical characteristics of the patients

	Placebo (N = 476)	Ipilimumab (N = 475)
Gender – no. (%)		
Male	293 (61.6)	296 (62.3)
Female	183 (38.4)	179 (37.7)
Age		
Median (range), yr	52 (18 – 78)	51 (20 – 84)
Stage provided at randomization – no. (%) <sup>†</sup>		
Stage IIIA	98 (20.6)	98 (20.6)
Stage IIIB	182 (38.2)	182 (38.3)
Stage IIIC (1–3 LN+)	121 (25.4)	122 (25.7)
Stage IIIC ( 4 LN+)	75 (15.8)	73 (15.4)
AJCC 2002 (CRF) – no. (%) <sup>‡</sup>		
Stage IIIA	88 (18.5)	98 (20.6)
Stage IIIB	207 (43.5)	213 (44.8)
Stage IIIC (1–3 LN+)	83 (17.4)	69 (14.5)
Stage IIIC ( 4 LN+)	98 (20.6)	95 (20.0)
Type of lymph node involvement – no. (%) <sup>‡</sup>		
Microscopic	193 (40.5)	210 (44.2)
Macroscopic	283 (59.5)	265 (55.8)
Nb of LN+ (pathological) – no. (%) <sup>‡</sup>		
1	220 (46.2)	217 (45.7)
2–3	158 (33.2)	163 (34.3)
4	98 (20.6)	95 (20.0)
Ulceration – no. (%) <sup>‡</sup>		
No	244 (51.3)	257 (54.1)
Yes	203 (42.6)	197 (41.5)
Unknown	29 (6.1)	21 (4.4)
Geographical region – no. (%) <sup>†</sup>		
North-America	119 (25.1)	117 (24.6)
Australia	19 (4.0)	18 (3.8)
Europe	337 (70.9)	341 (71.6)

<sup>†</sup> as provided at randomization<sup>‡</sup> as indicated on case report forms

LN+: positive lymph nodes

**Table 2**

Summary of questionnaire compliance rates

	Treatment arm	
	Pbo (N=476)	Ipi (N=475)
Baseline	93.3%	94.5%
Week 4	94.3%	89.1%
Week 7	92.8%	80.2%
Week 10	91.5%	79.5%
Week 24	81.2%	68.8%
Week 36	75.5%	69.5%
Week 48	70.4%	67.4%
Week 60	67.4%	61.1%
Week 72	68.1%	56.7%
Week 84	64.9%	54.1%
Week 96	63.2%	54.8%
Week 108	55.3%	46.5%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**

Primary HRQoL results

<i>Average score (standard deviation)</i>	<b>Pbo (N=476)</b>	<b>Ipi (N=475)</b>	<b>Difference in Means (95%CI)</b>	<b>Wilcoxon p-value</b>
ADI Global Health/QoL	77.32 (17.36)	72.96 (17.82)	4.35 (2.07–6.64)	<.001
AAI Global Health/QoL	76.48 (17.52)	72.32 (18.60)	4.16 (1.67–6.64)	<.001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4**

Descriptive GH/QoL score per time

<i>Mean (SD)</i>	<b>Pbo (N=476)</b>	<b>Ipi (N=475)</b>
Baseline	76.85 (18.26)	78.40 (18.83)
Week 4	78.45 (17.80)	76.07 (17.74)
Week 7	76.97 (19.39)	71.55 (21.75)
Week 10	77.41 (19.77)	69.52 (23.12)
Week 24	78.61 (17.90)	74.16 (21.43)
Week 36	78.80 (18.57)	74.16 (19.94)
Week 48	77.76 (19.21)	74.63 (20.50)
Week 60	79.01 (17.69)	74.39 (20.59)
Week 72	78.13 (18.80)	74.25 (20.80)
Week 84	78.89 (18.72)	75.73 (20.74)
Week 96	78.76 (18.55)	74.48 (20.94)
Week 108	79.08 (18.10)	74.10 (21.58)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5**

Summary of change from baseline in HRQoL score at Week 10 by Treatment and AE Status at week 10 visit.

	Placebo		Ipilimumab	
	AE absent	AE present	AE absent	AE present
<b>Diarrhea</b>				
Mean (SD)	-1.34 (14.66)	14.35 (30.98)	5.21 (20.40)	23.08 (39.34)
N obs	324	79	224	130
<b>Appetite loss</b>				
Mean (SD)	1.66 (14.63)	22.22 (27.22)	7.84 (23.35)	18.33 (33.29)
N obs	401	6	336	20
<b>Fatigue</b>				
Mean (SD)	1.83 (17.09)	11.39 (20.34)	8.65 (22.40)	16.62 (27.25)
N obs	309	98	239	119
<b>Insomnia</b>				
Mean (SD)	-3.35 (24.01)	14.81 (55.56)	7.43 (27.48)	4.17 (36.26)
N obs	398	9	341	16
<b>Nausea/vomiting</b>				
Mean (SD)	0.42 (6.68)	15.41 (21.40)	3.88 (14.10)	20.45 (27.70)
N obs	354	53	292	66