Survival and biomarker analyses from the OpACIN-neo and OpACIN neoadjuvant immunotherapy trials in stage III melanoma

Rozeman, E.A.; Hoefsmit, E.P.; Reijers, I.L.M.; Saw, R.P.M.; Versluis, J.M.; Krijgsman, O.; ... ; Blank, C.U.

LETTERS
https://doi.org/10.1038/s41591-020-01211-7

nature medicine

Check for updates

Survival and biomarker analyses from the OpACIN-neo and OpACIN neoadjuvant immunotherapy trials in stage III melanoma

E. A. Rozeman[1], E. P. Hoefsmit[2,17], I. L. M. Reijers[1,17], R. P. M. Saw[3,4,5], J. M. Versluis[1], O. Krijgsman[2,6], P. Dimitriadis[2], K. Sikorska[7], B. A. van de Wiel[8], H. Eriksson[9,10], M. Gonzalez[3], A. Torres Acosta[7], L. G. Grijpink-Ongering[7], K. Shannon[3,5], J. B. A. G. Haanen[1,2], J. Stretch[3,4,5], S. Ch'ng[3,4,5], O. E. Nieweg[3,4,5], H. A. Mallo[1], S. Adriaansz[1], R. M. Kerkhoven[11], S. Cornelissen[12], A. Broeks[12], W. M. C. Klop[13], C. L. Zuur[13], W. J. van Houdt[13], D. S. Peeper[2,6], A. J. Spillane[3,4,14], A. C. J. van Akkooi[13], R. A. Scolyer[3,15], T. N. M. Schumacher[2,6], A. M. Menzies[3,16], G. V. Long[3,16] and C. U. Blank[1,2] ✉

Neoadjuvant ipilimumab plus nivolumab showed high pathologic response rates (pRRs) in patients with macroscopic stage III melanoma in the phase 1b OpACIN (NCT02437279) and phase 2 OpACIN-neo (NCT02977052) studies[1,2]. While the results are promising, data on the durability of these pathologic responses and baseline biomarkers for response and survival were lacking. After a median follow-up of 4 years, none of the patients with a pathologic response ($n = 7/9$ patients) in the OpACIN study had relapsed. In OpACIN-neo ($n = 86$), the 2-year estimated relapse-free survival was 84% for all patients, 97% for patients achieving a pathologic response and 36% for nonresponders ($P < 0.001$). High tumor mutational burden (TMB) and high interferon-gamma-related gene expression signature score (IFN-γ score) were associated with pathologic response and low risk of relapse; pRR was 100% in patients with high IFN-γ score/high TMB; patients with high IFN-γ score/low TMB or low IFN-γ score/high TMB had pRRs of 91% and 88%; while patients with low IFN-γ score/low TMB had a pRR of only 39%. These data demonstrate long-term benefit in patients with a pathologic response and show the predictive potential of TMB and IFN-γ score. Our findings provide a strong rationale for a randomized phase 3 study comparing neoadjuvant ipilimumab plus nivolumab versus standard adjuvant therapy with antibodies against the programmed cell death protein-1 (anti-PD-1) in macroscopic stage III melanoma.

With the current standard of care consisting of surgery followed by adjuvant anti-PD-1 or targeted therapy with B-Raf proto-oncogene (BRAF) and mitogen-activated protein kinase (MEK) inhibitors, clinical trial data suggest that 40% of the macroscopic stage III melanoma patients still relapse within 3 years[3–5]. Moreover, a substantial subset of patients (15–25%) relapse soon after surgery and before starting adjuvant therapy[6,7], resulting in a 3-year relapse-free survival (RFS) of the intent-to-treat population of less than 50%.

Preclinical studies demonstrated improved survival and stronger antitumor immunity when immune checkpoint inhibition (ICI) was given before surgery as compared to adjuvant application[8–12]. Neoadjuvant therapy also has the advantage of providing information on pathologic response, which is valuable to estimate prognosis, and to guide the choice of adjuvant therapy and follow-up. Moreover, the availability of tumor tissue before and following therapy enables efficient exploration of possible mechanisms of resistance and response, and identification of baseline biomarkers.

The OpACIN study investigated neoadjuvant ICI in stage III melanoma patients, comparing four cycles of adjuvant ipilimumab plus nivolumab to two cycles of neoadjuvant and two cycles of adjuvant ipilimumab plus nivolumab (Extended Data Fig. 1a). This study showed that neoadjuvant ipilimumab plus nivolumab was feasible, induced an unexpectedly high pRR (78%) and expanded more tumor resident T cell clones compared to adjuvant treatment[1]. However, toxicity was high in both arms, with 90% grade 3–4 immunotherapy-related adverse events (irAEs). Similar high response and toxicity rates upon neoadjuvant ipilimumab plus nivolumab were observed in another study[13].

[1]Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [2]Department of Molecular Oncology and Immunology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [3]Melanoma Institute Australia, The University of Sydney, Sydney, New South Wales, Australia. [4]Department of Surgery, Mater Hospital, Sydney, New South Wales, Australia. [5]Department of Surgery, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia. [6]Oncode Institute, Utrecht, the Netherlands. [7]Department of Biometrics, Netherlands Cancer Institute, Amsterdam, the Netherlands. [8]Department of Pathology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [9]Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden. [10]Department of Oncology/Skin Cancer Center, Theme Cancer, Karolinska University Hospital, Stockholm, Sweden. [11]Genomics Core Facility, Netherlands Cancer Institute, Amsterdam, the Netherlands. [12]Core Facility Molecular Pathology and Biobanking, Netherlands Cancer Institute, Amsterdam, the Netherlands. [13]Department of Surgical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [14]Breast and Melanoma Surgery Unit, Royal North Shore hospital, Sydney, New South Wales, Australia. [15]Department of Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital and New South Wales Health Pathology, Sydney, New South Wales, Australia. [16]Department of Medical Oncology, Royal North Shore and Mater Hospitals, Sydney, New South Wales, Australia. [17]These authors contributed equally: E. P. Hoefsmit, I. L. M. Reijers. ✉e-mail: c.blank@nki.nl

The subsequent multicenter randomized phase 2 OpACIN-neo study tested the efficacy and toxicity of three different dosing schedules of neoadjuvant ipilimumab plus nivolumab (A: two cycles ipilimumab 3 mg kg^{-1} plus nivolumab 1 mg kg^{-1} every 3 weeks; B: two cycles ipilimumab 1 mg kg^{-1} plus nivolumab 3 mg kg^{-1} every 3 weeks; C: two cycles ipilimumab 3 mg kg^{-1} every 3 weeks followed by two cycles nivolumab 3 mg kg^{-1} every 2 weeks) in 86 stage III melanoma patients with ≥1 measurable lymph node metastasis according to the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria (Extended Data Fig. 2a,b and Supplementary Table 1). Primary analysis identified the treatment regimen consisting of two cycles of ipilimumab 1 mg kg^{-1} plus nivolumab 3 mg kg^{-1} (arm B) as the most favorable schedule with a pRR of 77% and 20% grade 3–4 irAEs[2]. Here we present updated survival data and biomarker analyses of OpACIN-neo and long-term follow-up data of OpACIN.

For OpACIN-neo, the median RFS and event-free survival (EFS) were not reached in any of the arms after a median follow-up of 24.6 months (interquartile range (IQR) 21.6–27.6 months). At data cutoff (7 Feb 2020), two (2.3%) patients had progressed before surgery and 12 (14%) patients had relapsed after surgery (Supplementary Table 2). The estimated 2-year RFS was 84% (95% confidence interval (CI) 76–92%) for all patients who did not progress before surgery (Fig. 1a), and 90% (95% CI 80–100%), 78% (95% CI 63–96%) and 83% (95% CI 70–100%) for arm A, B and C, respectively (log-rank $P=0.58$; Fig. 1b). The 2-year RFS remained significantly higher for patients achieving a pathologic response (97%; 95% CI 93–100%) versus patients without pathologic response (36%; 95% CI 17–74%, log-rank $P<0.0001$; Fig. 1c). There were only two events in the group of patients with pathologic response; one patient died due to toxicity and one patient relapsed. The estimated 2-year EFS was 82% (95% CI 74–91%) and also did not differ between arms (arm A: 90%; 95% CI 80–100%; arm B: 74%; 95% CI 59–94%; arm C: 81%; 95% CI 67–97%; log-rank $P=0.42$; Extended Data Fig. 3a,b). None of the patients were treated with adjuvant systemic therapy and three nonresponding patients were treated with adjuvant radiotherapy. Patients that progressed or relapsed were treated according to standard practice (Supplementary Table 3). The 2-year overall survival (OS) was 95% (95% CI 90–100%) for the total population and 93% (95% CI 85–100%) in arm A, 95% (95% CI 87–100%) in arm B and 96% (95% CI 89–100%) in arm C (log-rank $P=0.85$; Extended Data Fig. 3c,d).

In the OpACIN trial, none of the seven patients who achieved a pathologic response upon neoadjuvant therapy relapsed (median follow-up of 48.0 months). Within the neoadjuvant arm, the two nonresponding patients and the patient who was not evaluable relapsed, while four patients relapsed in the adjuvant arm. Median RFS, EFS and OS were not reached for both arms (Fig. 1d and Extended Data Fig. 1b,c). The 4-year EFS rate was 80% for the neoadjuvant arm and 60% for the adjuvant arm (Extended Data Fig. 1b), with 4-year OS rates of 90% and 70%, respectively (Extended Data Fig. 1c).

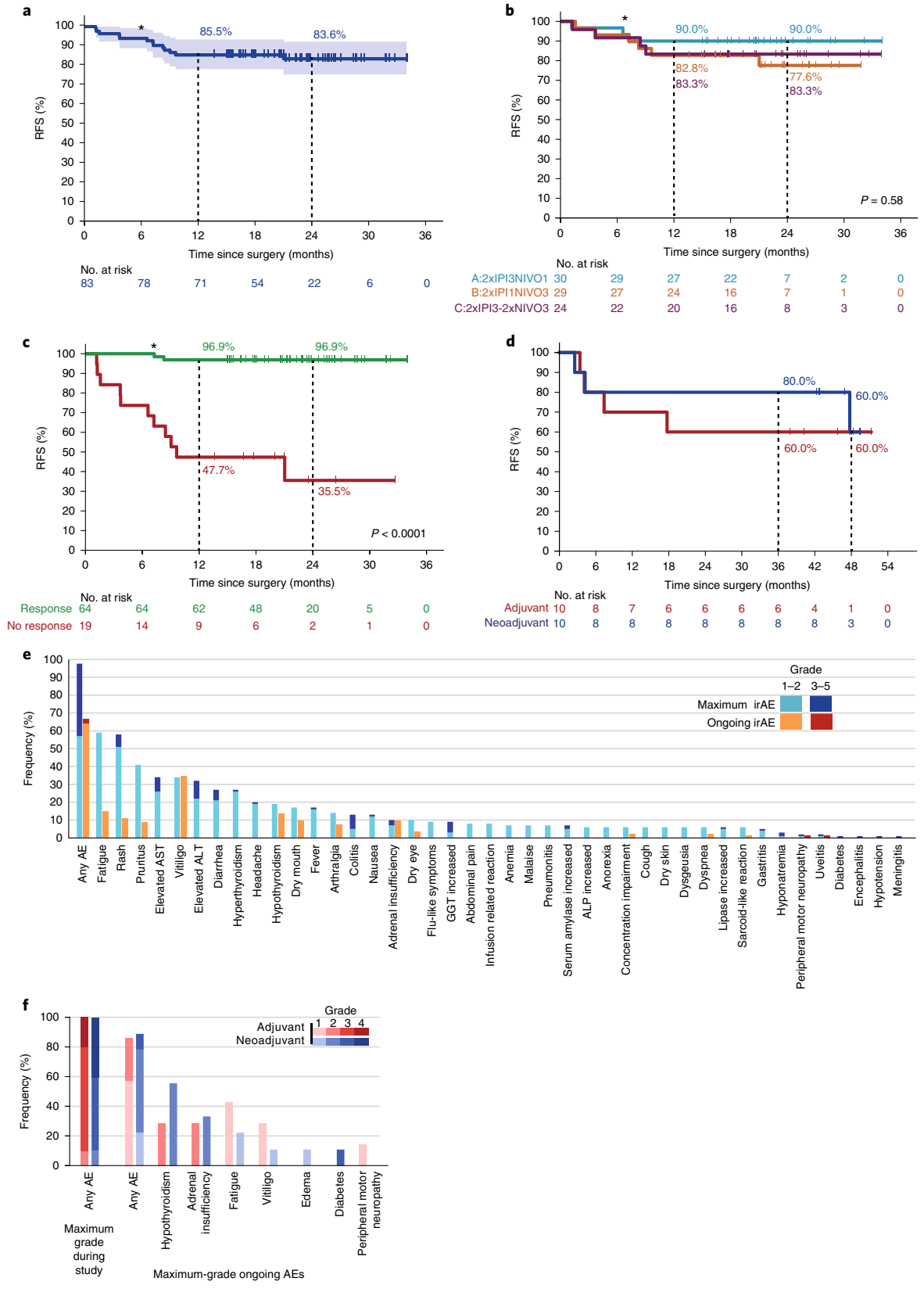
Grade 3–5 irAEs were observed in 43% (95% CI 25–63%) of patients in arm A, 27% (95% CI 12–46%) in arm B and 54% (95% CI 33–73%) in arm C (Supplementary Table 4). The majority of toxicities occurred within the first 12 weeks; only four patients developed their first high-grade irAE beyond 12 weeks. High-grade toxicities were more frequent in female than male patients (51.4% versus 32.7%; $P=0.081$). There was no difference in toxicity rate between older (>60 years; 35.3%) and younger (≤60 years; 44.2%, $P=0.41$) patients. Of 81 patients who were still alive, 55 (68%) had ongoing irAEs at data cutoff (Supplementary Table 5 and Fig. 1e). The majority of these irAEs were low grade; only two (3%) patients had grade 3 irAEs. The most frequent ongoing irAEs were vitiligo (35%), endocrine toxicities (21%), fatigue (15%), rash (11%), dry mouth (10%) and arthralgia (7%). In total, 17 (21%) patients need hormone replacement therapy: 11 (14%) received thyroid hormone and eight (10%) received corticosteroids. The frequency of ongoing toxicities was similar across arms. Surgery-related adverse events (AEs) were observed in 83% of patients including 14% grade 3–4 AEs. Frequencies were not different between treatment arms (Supplementary Table 6). Ongoing surgery-related AEs were observed in 31 (38%) patients, with lymphedema being most frequent (17 patients; 21%; Supplementary Table 5 and Extended Data Fig. 4). Similarly, in the OpACIN trial, all high-grade toxicities recovered to grade 1 or lower, with the exception of grade 2 endocrinopathies (Fig. 1f). Eight of the 16 patients alive (50%) require ongoing hormone replacement therapy: 7 (44%) require thyroid hormone, 5 (31%) require corticosteroids and 1 (6%) patient is insulin dependent. Two patients who require corticosteroids developed central adrenal insufficiency after long-term steroid use for treatment of irAEs.

Since pathologic response appeared to be a surrogate marker for RFS, baseline biomarkers for response were analyzed. We previously showed that clinical characteristics such as ulceration, maximum diameter of target lesions (assessed radiologically) and PD-1 expression were not significantly associated with response in OpACIN-neo[2] (Extended Data Fig. 5), while baseline interferon-gamma (IFN-γ) gene signature expression[14] was associated with absence of relapse in the OpACIN study[1]. In OpACIN-neo, we performed RNA sequencing on pretreatment tumor biopsies (data available for 65 patients) and confirmed that a high IFN-γ score (average z-score of all genes within the IFN-γ signature described by Ayers et al.[14]) was associated with pathologic response and low risk of relapse (Fig. 2a).

Fig. 1 | Relapse-free survival and ongoing toxicities. a, RFS for the total population of the OpACIN-neo study. A Kaplan–Meier curve for RFS of all patients who underwent surgery ($n=83$) was generated (two patients progressed before surgery and one patient did not undergo surgery because of toxicity). The corresponding 95% CI is displayed and was computed using log transformation. **b**, RFS of the OpACIN-neo study by treatment arm. All patients from arm A ($n=30$), 29 patients from arm B (one patient progressed before surgery) and 24 patients from arm C (one patient progressed before surgery and one patient had no surgery because of toxicity) were included. **c**, RFS of the OpACIN-neo study by pathologic response after neoadjuvant treatment. All patients of whom pathologic response was assessed and did not progress before surgery ($n=83$) were included and categorized by pathologic response (responders: $n=64$; nonresponders: $n=19$). Pathologic response was defined as <50% viable tumor cells in the tumor bed and was scored by two independent pathologists blinded to the treatment arm. P values were calculated using the log-rank test (two-sided). An asterisk denotes the patient who died due to irAEs. **d**, RFS curve of patients treated in the OpACIN study by treatment arm. A Kaplan–Meier curve of RFS is displayed for all 20 patients (10 per arm) who were included in the study. **e**, Frequency of maximum-grade and ongoing irAEs of the OpACIN-neo study. Frequencies of maximum grade irAEs are displayed in light blue (grade 1–2) and dark blue (grade 3–5), and frequencies of ongoing irAEs in orange (grade 1–2) and red (grade 3–5). irAEs that were reported at a frequency of >5% and all grade 3–5 irAEs were included. All patients ($n=86$) were included in the analysis of maximum-grade irAEs; for ongoing irAEs, only patients alive at the time of data cutoff ($n=81$) were included. ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase. **f**, Frequency of maximum-grade and ongoing irAEs of the OpACIN study. Frequencies of maximum-grade irAEs during the study are displayed in the leftmost columns and frequencies of maximum-grade ongoing irAEs in the other columns, both in ascending shades of red (adjuvant arm) and blue (neoadjuvant arm) for grades 1–4. irAEs that were reported at a frequency of >5% and all grade 3–5 irAEs were included. All patients ($n=20$) were included in the analysis of maximum-grade irAEs; for ongoing irAEs, only patients alive at the time of data cutoff ($n=17$) were included.

Using the upper tertile as the threshold, 20 of 21 (95%) patients with a high IFN-γ score achieved a pathologic response compared to 27 of 43 (62%) patients with a low IFN-γ score ($P=0.014$). We also observed a numerically longer EFS for patients with a high IFN-γ score compared with patients with a low IFN-γ score (2-year EFS: 90% versus 71%; Fig. 2b).

Whole-exome sequencing data were available for 60 patients. $BRAF^{V600}$ mutations were observed in 28 (47%) patients (Extended



Data Fig. 6a). No significant difference in pRRs was observed according to $BRAF^{V600}$ status (64% for patients who were positive for $BRAF^{V600}$ mutation versus 83% for those with *BRAF* wild-type status; $P=0.11$). Patients with a pathologic response had a higher TMB than nonresponders (median 860 versus 293; $P=0.0013$; Fig. 2c). Baseline TMB was strongly associated with EFS: estimated 2-year EFS was 93.3% for patients with TMB > median versus 58.8% for patients with TMB < median (log-rank $P=0.0027$; Fig. 2d). There was no correlation between IFN-γ score and TMB (log scale; $R=-0.1$; $P=0.44$; Extended Data Fig. 6b), and both biomarkers were significantly associated with pathologic response (IFN-γ score $P=0.0066$; TMB $P=0.021$) in a multivariate logistic regression model that included age, gender and continent (Supplementary Table 7).

Based on this observation, we evaluated whether the combination of IFN-γ score and TMB could better discriminate responders from nonresponders. Optimal cutoffs were defined based on summary receiver operating characteristic (sROC) curves, identifying 0.4665 as the optimal cutoff for IFN-γ score (exactly corresponding to the upper tertile) and 747 for TMB (Fig. 2e,f). Using this strategy, a group of patients who were substantially less likely to respond to neoadjuvant ipilimumab plus nivolumab was identified; these patients with both a low IFN-γ score and low TMB ($n=23$) had a pRR of only 39% (Fig. 2g and Extended Data Fig. 6c). For patients with a high IFN-γ score/high TMB ($n=9$), the pRR was 100%, while for patients with only a high IFN-γ score ($n=11$), the pRR was 91%, and for patients with only high TMB ($n=16$), pRR was 88% (Fig. 2g and Extended Data Fig. 6c). Correspondingly, the group with a low IFN-γ score/low TMB had a significantly lower 2-year EFS of only 49.5%, compared to 83.3%, 93.8% and 100% for patients with IFN-γ-low/TMB-high, IFN-γ-high/TMB-low and IFN-γ-high/TMB-high tumors, respectively ($P=0.0018$; Fig. 2h). The area under the sROC curve (AUC) was higher for the combination of the IFN-γ score and TMB (0.83) compared to either biomarker alone (IFN-γ score: 0.67; TMB: 0.76; Fig. 2e,f and Extended Data Fig. 6d).

Applying the microenvironment cell populations (MCP) counter signature[15], we assessed different cell populations that could discriminate patients according to pathologic response. Higher levels of all immune cell populations were found in responders (Extended Data Fig. 6e). Gene-set enrichment analysis (GSEA) on differentially expressed genes revealed that several immune pathways were upregulated in responders, as well as proliferation and signaling pathways. Interestingly, the hallmark angiogenesis and epithelial-to-mesenchymal transition gene sets were upregulated in nonresponders (Extended Data Fig. 6f).

To identify potential peripheral blood biomarkers, we performed the Olink proteomic assay, evaluating 92 immuno-oncology-related markers in plasma samples of patients treated in OpACIN-neo ($n=85$). We found a significant increase in almost all markers after neoadjuvant treatment (Fig. 3a). The highest post-treatment increases were observed for PD-1 ($P<0.0001$), CXCL9 ($P<0.0001$) and CXCL10 ($P<0.0001$), irrespective of response (Extended Data Fig. 7a). Significantly higher levels of vascular endothelial growth factor receptor 2 (VEGFR-2; $P<0.0001$), CX3CL1 ($P=0.0020$) and PD-L2 ($P=0.0018$) were found in pretreatment samples of nonresponders versus responders (Fig. 3b and Extended Data Fig. 7b).

A post hoc analysis demonstrated a trend toward a higher pRR for Australian versus European patients (84.2% versus 64.7%; OR 2.50, $P=0.092$; Supplementary Table 7 and Fig. 4a). Australian patients were older (median age 60 versus 53 years; $P=0.017$) and more likely to be male (65.8% versus 50.0%; $P=0.14$) compared to European patients (Supplementary Table 8). Multivariate analysis including continent, age, gender, IFN-γ score and TMB revealed that only IFN-γ score and TMB were significantly associated with response (OR 3.76, $P=0.0066$ and OR 14.19, $P=0.021$, respectively; Supplementary Table 7). We observed no continental difference in IFN-γ score (Fig. 4b), but TMB was higher in Australian patients ($P=0.0003$; Fig. 4c). There was a positive correlation between age and TMB (Fig. 4d), indicating that the higher TMB in Australian patients might be explained by higher age and/or cumulative effect of more UV exposure.

To our knowledge, OpACIN-neo is so far the largest and OpACIN the most mature study that evaluated neoadjuvant ipilimumab plus nivolumab, both showing a high pRR after only 6 weeks of therapy. Here we demonstrated that neoadjuvant ipilimumab plus nivolumab, without subsequent adjuvant systemic therapy, induces a durable RFS benefit with a 2-year RFS of more than 80%.

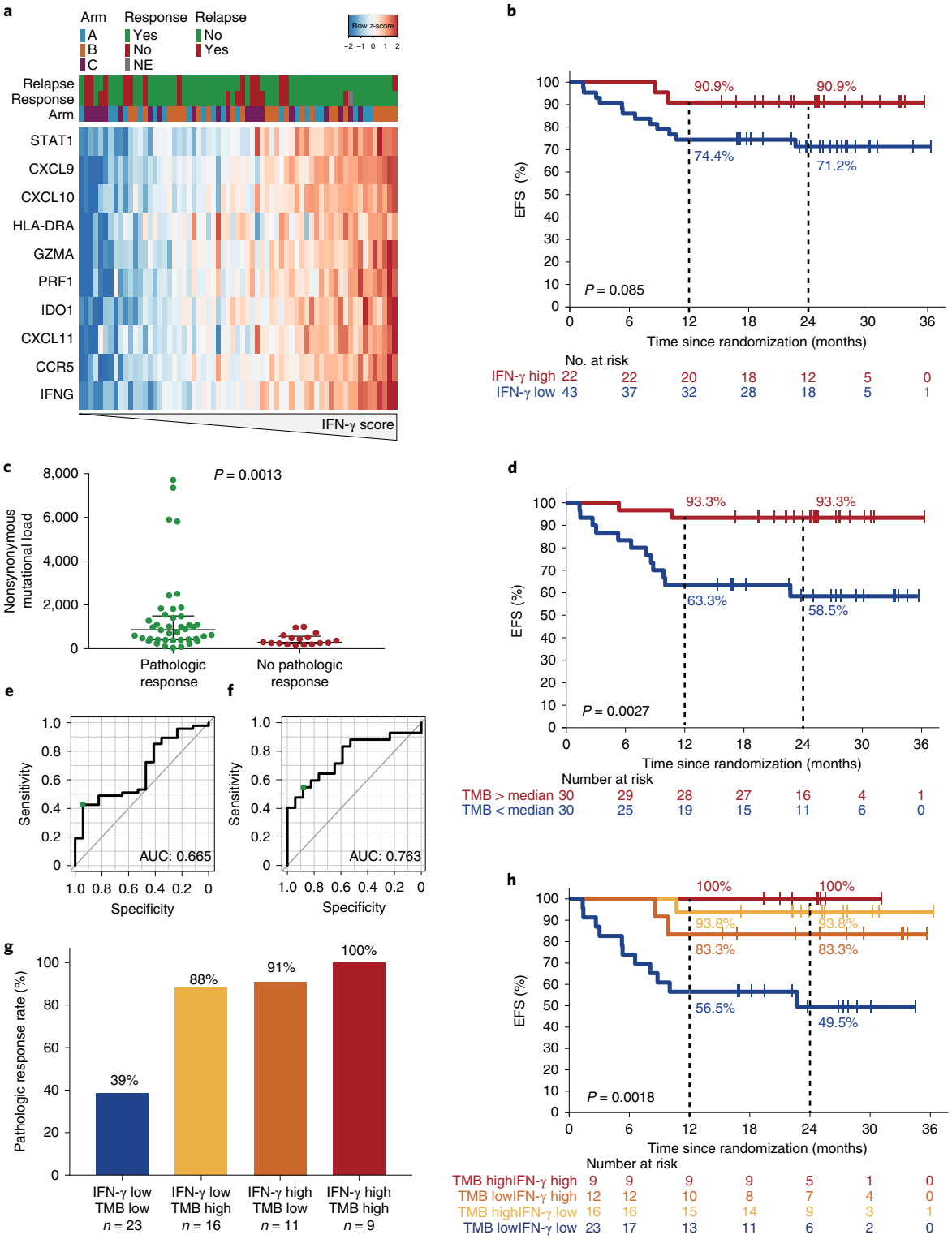
Since only 1 of the 71 (1.4%) patients with a pathologic response versus 16 of 23 (69.6%) nonresponding patients progressed or relapsed, pathologic response appears to be a strong surrogate marker for long-term benefit. These data are corroborated by the other trials included in the pooled analysis (40 additional patients, total $n=184$) of the International Neoadjuvant Immunotherapy in Melanoma consortium (INMC) that showed a 2-year RFS of 96% in all patients achieving a pathologic complete response upon neoadjuvant ICI versus 64% in patients without a pathologic complete response[16]. The extensive RFS difference between responding and nonresponding patients highlights the need for identification of baseline biomarkers predictive for response.

To our knowledge, this study reports the first large-scale biomarker analysis in the neoadjuvant setting in stage III melanoma. We identified that IFN-γ signature expression and TMB were associated with pathologic response and survival. Both biomarkers

Fig. 2 | Baseline IFN-γ signature and tumor mutational burden associated with response and relapse. a, RNA sequencing of pretreatment lymph node tumor biopsies of the OpACIN-neo study. The heat map of the IFN-γ[14] RNA gene signatures included 65 patients for whom baseline material was available, ordered according to average expression of IFN-γ gene signature. The IFN-γ score of each patient is the average *z*-score of all the genes within the IFN-γ signature, which was calculated on the gene expression counts normalized by DESeq2. Each column represents one patient (green: pathologic response/no relapse; red: no pathologic response/relapse; gray: not evaluable (NE); blue: treatment arm A; orange: treatment arm B; purple: treatment arm C), and rows display genes. Positive values (red) indicate higher gene expression and negative values (blue) indicate lower gene expression. **b**, EFS for all patients with available RNA-sequencing data ($n=65$) by IFN-γ score. A Kaplan–Meier curve displays patients with a high (red) and low (blue) IFN-γ score using the upper tertile of the average IFN-γ score as the cutoff. **c**, Whole-exome sequencing of pretreatment lymph node tumor biopsies of the OpACIN-neo study. The nonsynonymous mutational load (TMB) of 60 patients with available baseline materials was calculated for patients with pathologic response (green) versus no pathologic response (red). The median and IQR are shown. **d**, EFS for all patients with available WES data ($n=60$) by nonsynonymous mutational load. A Kaplan–Meier curve displays patients with a TMB > median (red) and TMB < median (blue). **e**,**f**, sROC curves for defining the optimal cutoff (green) of IFN-γ score and TMB. **e**, The AUC for the IFN-γ score (0.67); optimal cutoff was 0.4665 ($n=64$). **f**, AUC for the TMB (0.76); optimal cutoff was 747 ($n=59$). **g**,**h**, Patients were grouped according to IFN-γ score (0.4665 as cutoff) and TMB (747 as cutoff) resulting in a group with a low IFN-γ score and low TMB (dark blue), low IFN-γ score and high TMB (yellow), high IFN-γ score and low TMB (orange) and high IFN-γ score and high TMB (red). Patients for whom both RNA-sequencing data and WES data were available were included ($n=59$ for response analysis (**g**) and $n=60$ for EFS analysis (**h**)). **g**, pRRs for the different patient groups. Numbers of patients per group are indicated. **h**, EFS of OpACIN-neo by different patient groups. **b**–**d**,**h**, *P* values were calculated using the log-rank test (two-sided).

have already been shown to be prognostic in stage III melanoma patients[17] and are associated with response to ICI[18]. Our results are not completely in line with those of the COMBI-AD trial investigating adjuvant BRAF plus MEK inhibition versus placebo, which demonstrated that IFN-γ gene expression, but not TMB, was predictive for RFS[17]. We found that only patients with a low IFN-γ score and low TMB are less likely to achieve a pathologic response and have inferior survival. This group represents an interesting target population to test new treatment combinations in the neoadjuvant setting.

Although the combination of IFN-γ score and TMB appears to be a promising biomarker for response, the requirement of a tumor biopsy and need for DNA and RNA sequencing might call for easier and faster biomarker approaches. Our previous analyses showed, however, that clinical characteristics and PD-1 expression were not significantly associated with response[2]. In advanced lung cancer,



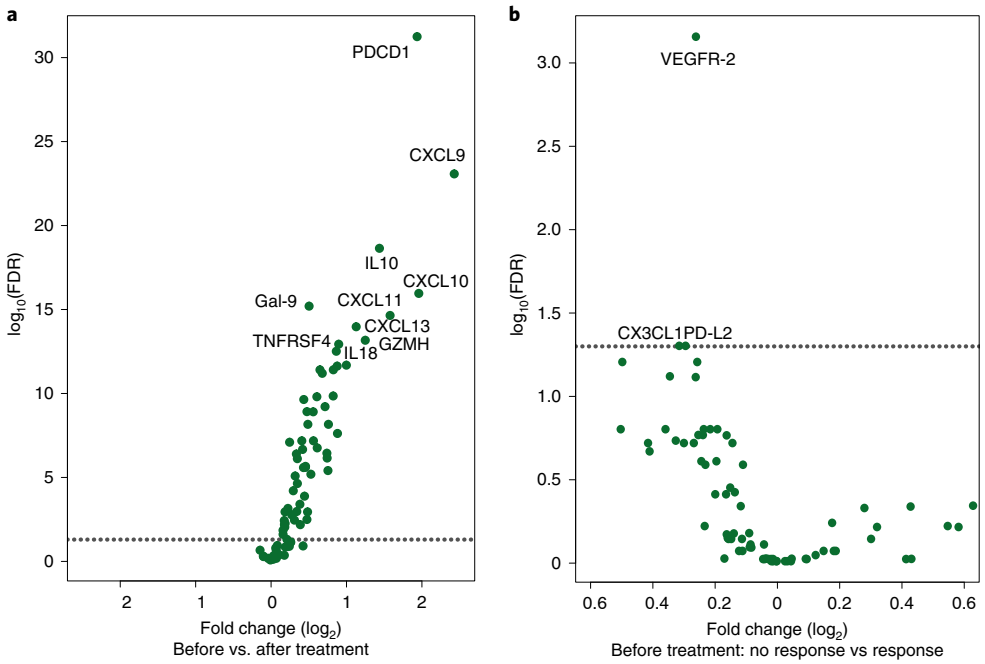


Fig. 3 | Plasma analysis using Olink proteomic assay. a,**b**, Volcano plots showing differential expression of plasma markers measured with Olink immunoassay. The horizontal axis displays the magnitude of a protein's fold change on a $\log_2$ scale; the vertical axis displays the significance scale by the $-\log_{10}$ (*P* adjusted), which increases with statistical significance. The black dotted line corresponds to the adjusted *P*-value cutoff (false discovery rate (FDR) = 0.05). **a**, Differential expression of plasma proteins between matched pre- and post-treatment samples ($n = 85$) are displayed. Proteins displayed on the right were higher in the post-treatment samples. A two-tailed paired Student's *t*-test was used to determine statistical significance between pre- and post-treatment samples. **b**, Differential expression of plasma proteins between pretreatment samples of patients without pathologic response ($n = 21$) and pretreatment samples of patients with a pathologic response ($n = 64$). Proteins displayed on the right were higher in patients who had a pathologic response. A two-tailed Welch's *t*-test was used to determine statistical significance between the two groups.

plasma-based TMB is used as a predictive biomarker for response[19]. Nevertheless, in clinical stage III melanoma patients, only 40% have detectable levels of circulating tumor DNA[20,21]. To reduce costs and time, targeted gene panel sequencing to determine TMB[22], panel RNA sequencing or techniques like the nCounter digital molecular barcoding technology to assess gene expression might be of interest. Recently the US Food and Drug Administration approved a panel sequencing-based TMB assay as diagnostic compendium for pembrolizumab therapy. Thus, biomarkers such as those identified in this study could be implemented in the clinic in the near future. The DONIMI trial (NCT04133948) is the first biomarker-driven neoadjuvant immunotherapy trial, randomizing patients between different treatment combinations based on their baseline IFN-γ score, with the aim to induce the IFN-γ signature and other immune pathways by adding an HDAC inhibitor. Alternative strategies, based on our findings of the GSEA, might be to target epithelial-to-mesenchymal transition and angiogenesis. The latter is supported by our observation that VEGFR-2 was also higher in pretreatment plasma of non-responders and corroborated by other studies linking angiogenesis to ICI resistance[23–26]. Baseline soluble VEGFR-2 might be a potential noninvasive biomarker to select patients for combinations of angiogenesis inhibitors and ICI, for example, in the NeoPeLe trial (NCT04207086).

A limitation of our study is the relatively small sample size, which precludes formal comparisons of response and survival between patient subgroups. Larger patient cohorts are needed for validation of the biomarkers and to define optimal cutoffs.

Because of the curative intent of therapy in stage III patients, (long-term) AEs and their impact on quality of life should be considered carefully, especially because some patients may be cured by surgery alone. Specifically, the risk of (long-term) toxicity should be weighed against the higher chance of response and cure. We observed that almost all grade 3–4 irAEs resolved, although a substantial proportion of patients still experienced low-grade ongoing toxicities. This is in line with safety data of ipilimumab and nivolumab in stage IV melanoma[27]. Adjuvant anti-PD-1 studies showed lower high-grade AE rates, but endocrine toxicities were observed in 23% of patients, which is in the same range as we observed[7,28]. Post hoc analyses of both adjuvant anti-PD-1 trials showed that the incidence of irAEs increased during the entire year of adjuvant therapy[29,30]. In our study, only 4 of 96 patients developed a first high-grade AE beyond 12 weeks after treatment initiation, which allows more focused irAE management.

To reduce the frequency of AEs and to increase RFS in patients without pathologic response, personalized treatment strategies are needed. The PRADO extension cohort of OpACIN-neo investigated a response-driven treatment strategy based on pathologic response in a marked index node (largest involved lymph node at baseline) and evaluated whether a therapeutic lymph node dissection could be omitted in patients with a major pathologic response[31]. Another strategy to reduce toxicity would be to explore whether the IFN-γ score and TMB are also biomarkers for response to neoadjuvant anti-PD-1 monotherapy, such that it might be possible to identify a group of patients that benefit from monotherapy. In the DONIMI trial, we are investigating this strategy in patients with a high IFN-γ score.

In conclusion, extended follow-up data of OpACIN and OpACIN-neo show that two cycles of neoadjuvant ipilimumab plus nivolumab without additional adjuvant therapy induces durable RFS in more than 80% of patients and further endorse pathologic response as a strong surrogate outcome marker for RFS. With this longer follow-up, almost all high-grade irAEs have resolved to ≤grade 1, except for endocrine toxicities requiring hormone

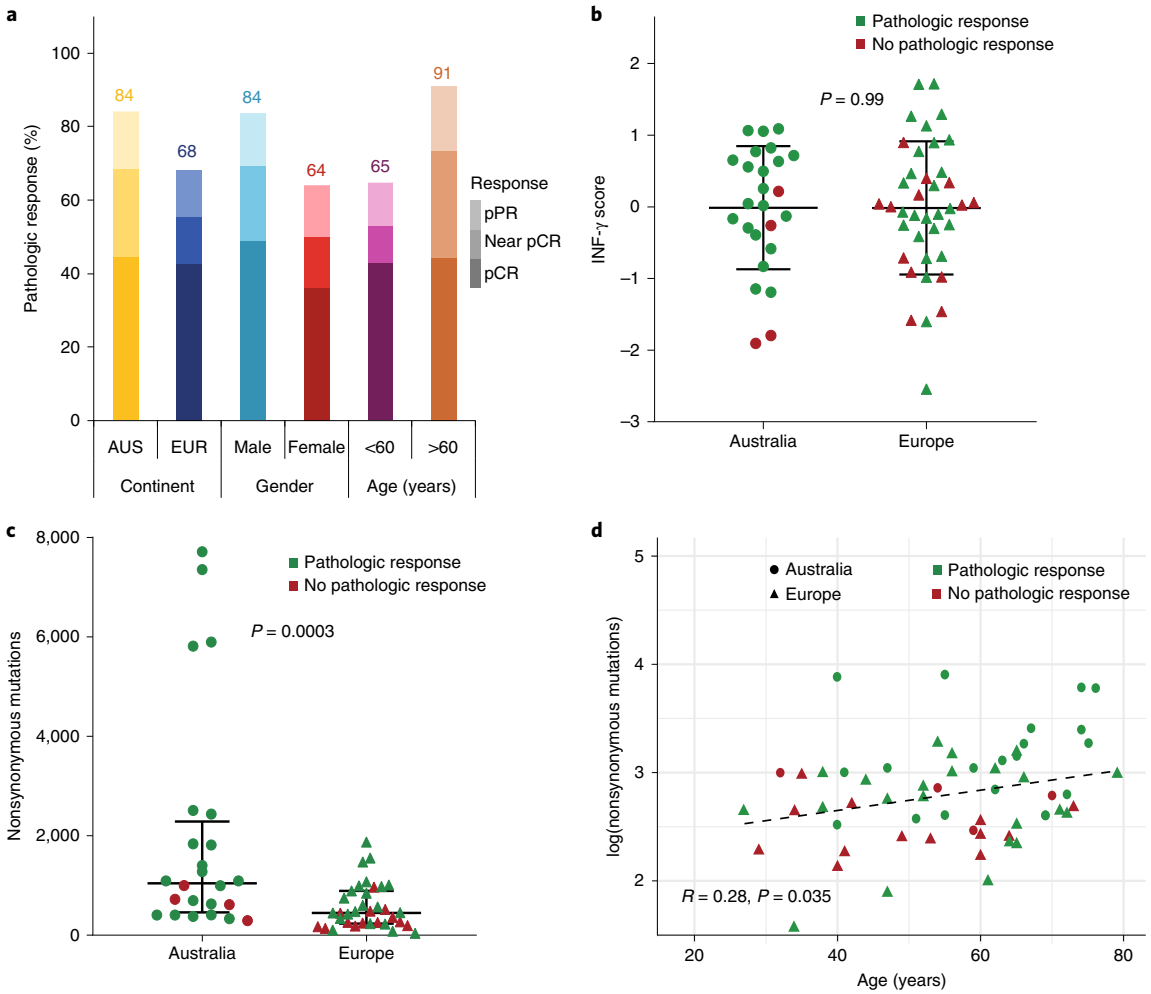


Fig. 4 | Continental differences between European and Australian patients. a, pRR of different subgroups of OpACIN-neo according to continent (Australia: yellow; Europe: dark blue), gender (male: light blue; female: red) and age (<60 years: purple; >60 years: orange). Ascending shades of colors represent patients with a pathologic partial response (pPR; >10–50% viable tumor cells), pathologic near-complete response (near-pCR; 1–10% viable tumor cells) and a pathologic complete response (pCR; 0% viable tumor cells). **b**, RNA sequencing of pretreatment lymph node tumor biopsies ($n = 64$) of the OpACIN-neo study. The IFN-γ score was calculated and compared between patients from Europe and Australia. Mean and standard deviation (s.d.) values are shown. *P* value was calculated using the student's *t*-test (two-sided). **c**, Whole-exome sequencing of pretreatment lymph node tumor biopsies of the OpACIN-neo study. The nonsynonymous mutational load (TMB) of 60 patients with available baseline materials was calculated for patients from Australia versus Europe. The median and IQR are shown. *P* values were calculated using the Mann–Whitney *U* test (two-sided). **d**, Correlation between age and TMB (displayed in log scale). The correlation coefficient and *P* value were calculated using the Pearson correlation method (two-sided). In **b** and **c**, every triangle (Europe) or circle (Australia) represents one patient with pathologic response (green) and no pathologic response (red).

replacement therapy. These findings provide a strong rationale to test two cycles of neoadjuvant ipilimumab plus nivolumab versus adjuvant anti-PD-1 in a randomized phase 3 study. Biomarker analyses revealed that patients with a low TMB and low IFN-γ gene signature expression are less likely to respond and therefore are a target population for new neoadjuvant treatment combinations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-020-01211-7.

Received: 3 August 2020; Accepted: 14 December 2020;
Published online: 8 February 2021

References

1. Blank, C. U. et al. Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. *Nat. Med.* **24**, 1655–1661 (2018).
2. Rozeman, E. A. et al. Identification of the optimal combination dosing schedule of neoadjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma (OpACIN-neo): a multicentre, phase 2, randomised, controlled trial. *Lancet Oncol.* **20**, 948–960 (2019).
3. Hauschild, A. et al. Longer follow-up confirms relapse-free survival benefit with adjuvant dabrafenib plus trametinib in patients with resected $BRAF^{V600}$-mutant stage III melanoma. *J. Clin. Oncol.* https://doi.org/10.1200/JCO.18.01219 (2018).
4. Eggermont, A. M. M. et al. Longer follow-up confirms recurrence-free survival benefit of adjuvant pembrolizumab in high-risk stage III Melanoma: updated results from the EORTC 1325-MG/KEYNOTE-054 Trial. *J. Clin. Oncol.* https://doi.org/10.1200/jco.20.02110 (2020).
5. Ascierto, P. A. et al. Adjuvant nivolumab versus ipilimumab in resected stage IIIB-C and stage IV melanoma (CheckMate 238): 4-year results from a multicentre, double-blind, randomised, controlled, phase 3 trial. *Lancet Oncol.* https://doi.org/10.1016/s1470-2045(20)30494-0 (2020).

6. Bloemendal, M. et al. Early recurrence in completely resected IIIB and IIIC melanoma warrants restaging prior to adjuvant therapy. *Ann. Surg. Oncol.* **26**, 3945–3952 (2019).
7. Weber, J. et al. Adjuvant nivolumab versus ipilimumab in resected stage III or IV melanoma. *N. Engl. J. Med.* **377**, 1824–1835 (2017).
8. Liu, J. et al. Improved efficacy of neoadjuvant compared to adjuvant immunotherapy to eradicate metastatic disease. *Cancer Discov.* **6**, 1382–1399 (2016).
9. Brockwell, N. K. et al. Neoadjuvant interferons: critical for effective PD-1–based immunotherapy in TNBC. *Cancer Immunol. Res.* **5**, 871–884 (2017).
10. Bourgeois-Daigneault, M.-C. et al. Neoadjuvant oncolytic virotherapy before surgery sensitizes triple-negative breast cancer to immune checkpoint therapy. *Sci. Transl. Med.* **10**, eaao1641 (2018).
11. Brooks, J. et al. Perioperative, spatiotemporally coordinated activation of T and NK cells prevents recurrence of pancreatic cancer. *Cancer Res.* **78**, 475–488 (2018).
12. O'Donnell, J. S., Hoefsmit, E. P., Smyth, M. J., Blank, C. U. & Teng, M. W. The promise of neoadjuvant immunotherapy and surgery for cancer treatment. *Clin. Cancer Res.* **25**, 5743–5751 (2019).
13. Amaria, R. N. et al. Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nat. Med.* **24**, 1649–1654 (2018).
14. Ayers, M. et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017).
15. Becht, E. et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 1–20 (2016).
16. Menzies, A. M. et al. Pathological response and survival neoadjuvant therapy in melanoma: a pooled analysis from the Neoadjuvant Melanoma Consortium. *J. Clin. Oncol.* **37**, 9503 (2019).
17. Dummer, R. et al. Adjuvant dabrafenib plus trametinib versus placebo in patients with resected, $BRAF^{V600}$-mutant, stage III melanoma (COMBI-AD): exploratory biomarker analyses from a randomised, phase 3 trial. *Lancet Oncol.* **21**, 358–372 (2020).
18. Lu, S. et al. Comparison of biomarker modalities for predicting response to PD-1/PD-L1 checkpoint blockade: a systematic review and meta-analysis. *JAMA Oncol.* **5**, 1195–1204 (2019).
19. Chae, Y. K. et al. Clinical implications of circulating tumor DNA tumor mutational burden (ctDNA TMB) in non-small cell lung cancer. *Oncologist* **24**, 820–828 (2019).
20. Lee, J. H. et al. Pre-operative ctDNA predicts survival in high-risk stage III cutaneous melanoma patients. *Ann. Oncol.* **30**, 815–822 (2019).
21. Rowe, S. P. et al. From validity to clinical utility: the influence of circulating tumor DNA on melanoma patient management in a real-world setting. *Mol. Oncol.* **12**, 1661–1672 (2018).
22. Fancello, L., Gandini, S., Pelicci, P. G. & Mazzarella, L. Tumor mutational burden quantification from targeted gene panels: major advancements and challenges. *J. Immunother. Cancer* **7**, 183 (2019).
23. Yuan, J. et al. Pretreatment serum VEGF is associated with clinical response and overall survival in advanced melanoma patients treated with ipilimumab. *Cancer Immunol. Res.* **2**, 127–132 (2014).
24. Ott, P. A., Hodi, F. S. & Buchbinder, E. I. Inhibition of immune checkpoints and vascular endothelial growth factor as combination therapy for metastatic melanoma: an overview of rationale, preclinical evidence and initial clinical data. *Front. Oncol.* **5**, 202 (2015).
25. Chen, P.-L. et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov.* **6**, 827–837 (2016).
26. Voron, T. et al. VEGF-A modulates expression of inhibitory checkpoints on $CD8^+$ T cells in tumors. *J. Exp. Med.* **212**, 139–148 (2015).
27. Hodi, F. S. et al. Nivolumab plus ipilimumab or nivolumab alone versus ipilimumab alone in advanced melanoma (CheckMate 067): 4-year outcomes of a multicentre, randomised, phase 3 trial. *Lancet Oncol.* **19**, 1480–1492 (2018).
28. Eggermont, A. M. M. et al. Adjuvant pembrolizumab versus placebo in resected stage III melanoma. *N. Engl. J. Med.* **378**, 1789–1801 (2018).
29. Eggermont, A. M. M. et al. Association between immune-related adverse events and recurrence-free survival among patients with stage III melanoma randomized to receive pembrolizumab or placebo: a secondary analysis of a randomized clinical trial. *JAMA Oncol.* https://doi.org/10.1001/jamaoncol.2019.5570 (2020).
30. Mandalà, M. et al. An analysis of nivolumab-mediated adverse events and association with clinical efficacy in resected stage III or IV melanoma (CheckMate 238). *J. Clin. Oncol.* **37**, 9584–9584 (2019).
31. Blank, C. U. et al. First safety and efficacy results of PRADO: A phase II study of personalized response-driven surgery and adjuvant therapy after neoadjuvant ipilimumab and nivolumab in resectable stage III melanoma. *J. Clin. Oncol.* **38**, 10002–10002 (2020).

Methods

Study design and patients. The investigator-initiated, multicenter randomized OpACIN-neo trial evaluated the toxicity and efficacy of three different dosing schedules of neoadjuvant ipilimumab and nivolumab in macroscopic stage III melanoma. The study was conducted at the Netherlands Cancer Institute, Melanoma Institute Australia and Karolinska Institute. The investigator-initiated, randomized OpACIN trial compared the safety and feasibility of neoadjuvant versus adjuvant application of ipilimumab and nivolumab in macroscopic stage III melanoma and was conducted in the Netherlands Cancer Institute. For both trials, eligible patients were 18 years or older, were diagnosed with histologically confirmed resectable stage III melanoma without in-transit metastasis, were naïve for systemic therapy, needed to have at least one measurable lymph node metastasis according to the RECIST 1.1 criteria and normal lactate-dehydrogenase level at baseline. The full trial designs, eligibility criteria and assessments have been reported previously[1,2].

Treatment and assessments. In OpACIN-neo, patients were randomized in a 1:1:1 ratio between one of the three treatment arms stratified by treatment center. In arm A, patients were treated with two cycles of ipilimumab 3 mg kg^{-1} plus nivolumab 1 mg kg^{-1} every 3 weeks. In arm B, patients received two cycles of ipilimumab 1 mg kg^{-1} plus nivolumab 3 mg kg^{-1} every 3 weeks. In arm C, patients received two cycles of ipilimumab 3 mg kg^{-1} every 3 weeks, directly followed by two cycles nivolumab 3 mg kg^{-1} every 2 weeks. Therapeutic lymph node dissection was planned in week 6.

In OpACIN, patients were randomized in a 1:1 ratio between four cycles of ipilimumab 3 mg kg^{-1} plus nivolumab 1 mg kg^{-1} every 3 weeks starting 6 weeks after therapeutic lymph node dissection (adjuvant arm), or two cycles of ipilimumab 3 mg kg^{-1} plus nivolumab 1 mg kg^{-1} every 3 weeks before surgery, followed by total lymph node dissection and two cycles of ipilimumab plus nivolumab after surgery (neoadjuvant arm).

Patients were evaluated for AEs before every cycle, during follow-up visit and upon indication when having symptoms assessed by physical examination and laboratory tests. AEs were reported and scored according to Common Terminology Criteria for Adverse Events version 4.03 by the treating physician who determined whether they were related to immunotherapy and/or surgery. Radiologic response was assessed by a local radiologist and scored according to RECIST 1.1 criteria.

Pathologic response in OpACIN was reviewed by one blinded pathologist who evaluated the surgery material on vital tumor cell percentage. In OpACIN-neo, pathologic response was centrally evaluated by two experienced and blinded pathologists according to pathologic response criteria of the INMC as previously described[32]. A pathologic response was defined as 0% viable tumor cells, a near-complete pathologic response as ≤10% viable tumor cells, a pathologic partial response as ≤50% viable tumor cells and a pathologic nonresponse as >50% viable tumor cells in the tumor bed area[32].

In OpACIN-neo, patients were evaluated for relapse and toxicities by physical examination and laboratory tests every 3 months from week 12 until development of distant metastasis, death, loss to follow-up or withdrawal of consent for up to 3 years. Radiology evaluation was performed according to institutional standards; at the Netherlands Cancer Institute, a computed tomography (CT) scan was obtained every 3 months for patients without a pathologic response and every 6 months for patients with a pathologic response; at Melanoma Institute Australia, a CT scan and MRI of the brain were performed every 3 months and a positron emission tomography–CT every year, while at the Karolinska Institute, a CT scan was performed every 6 months.

In OpACIN, patients in both trial arms were evaluated for relapse every 3 months starting at week 18 until 3 years by physical examination and laboratory testing. CT scans were performed every 6 months according to the Netherlands Cancer Institute standard. Subsequent follow-up was carried out according to the current Dutch melanoma guidelines (years 4 and 5: physical examination and laboratory testing every 6 months; years 6–10: once a year).

Collection of blood and tumor samples. In OpACIN-neo, blood samples for isolation of plasma and peripheral blood mononuclear cells were collected at baseline, after immunotherapy and before and during surgery at week 6 and at week 12. Pretreatment tumor biopsies were taken from a lymph node metastasis by a radiologist using ultrasound. The obtained samples were immediately snap frozen or formalin fixed and paraffin embedded.

DNA and RNA sequencing. DNA and RNA were isolated from patients that had sufficient tumor material based on the pathologist's scoring (at least 30% tumor cells on an H&E-stained cryostat frozen section of the tumor sample). A total of 65 of 86 patients had sufficient tumor cells in their frozen biopsies. DNA and RNA were simultaneously isolated from fresh frozen pretreatment tumor sections (10 µm) with the AllPrep DNA/RNA/miRNA Universal isolation kit (Qiagen, 80224) using the QIAcube, according to the manufacturer's protocol. DNA was isolated from peripheral blood mononuclear cells using AllPrep DNA/RNA/miRNA Universal isolation kit (Qiagen, 80224) to be able to filter out single-nucleotide polymorphisms when determining TMB.

Strand-specific libraries were generated using the TruSeq Stranded mRNA sample preparation kit (Illumina) according to the manufacturer's instructions. In brief, total RNA was fragmented, randomly primed and reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen) with the addition of actinomycin D. The synthesis of second strand was performed using polymerase I and RNaseH with replacement of dTTP for dUTP. The generated cDNA fragments were 3′-end adenylated and ligated and amplified by 12 cycles of PCR. The libraries were validated on a 2100 Bioanalyzer using a 7500 chip (Agilent) and pooled. We pooled and sequenced the libraries using single-end sequencing with 65-bp reads on a HiSeq 2500 System in a high output mode using V4 chemistry (Illumina). FASTQ files were mapped to the human reference genome (Homo.sapiens.GRCh38.v82) using STAR(2.7)[33] with default settings. Count data generated with HTseq-count (version 0.11.1)[34] were analyzed with DESeq2 (version 1.24.0)[35]. Centering of the normalized gene expression data for each dataset was performed by subtracting the row means and scaling by dividing the columns by the s.d. values. Next, the previously defined gene expression immune signatures, IFN-γ signature[14] and MCP counter (using MCPcounter version 1.1.0)[15] were analyzed. GSEA (using fGSEA version 1.10.1) was performed to identify gene sets from the Hallmark database[36] that were significantly enriched in responders compared with nonresponders. FDR values were computed as previously described[37]. An FDR value of <0.05 was considered as significant.

Whole-exome sequencing was performed by CeGaT. Exome libraries were generated using the Twist Human Core Exome Plus (Twist Biosciences), according to the manufacturer's instructions. The libraries were sequenced with 100-bp reads on a NovaSeq 6000 System according to the manufacturer's protocols, with a sequence quality Q30 value of 92.4%. Data were analyzed in CeGaT exome analysis pipeline. Briefly, demultiplexing of the sequencing reads was performed with Illumina bcl2fastq (version 2.20). Adaptors were trimmed with Skewer (version 0.2.2). The quality of FASTQ files was analyzed with FastQC (version 0.11.5-cegat). Subsequent FASTQ files were aligned to the human reference genome (GRCh38) using Burrows–Wheeler aligner (version 0.7.12)[38], followed by marking of duplicate reads by the MarkDuplicates tool in Picard (version 1.140). Subsequently, base quality scores were recalibrated using BaseRecalibrator in GATK (version 4.0.6.0) and single-nucleotide variants were called using MuTect2 in GATK[39]. The TMB was calculated by summarizing the total number of nonsynonymous, somatic mutations per sample with minimal variant allele frequency of 0.05 (5%).

Plasma proteomics profiling. A multiplex assay to profile the plasma proteomics was performed using proximity extension assay technology (Olink Bioscience AB) on pre- and post-treatment plasma samples. The assay was performed at the Department of Clinical Chemistry & Hematology at the University Medical Center Utrecht. We selected the Olink Immuno-Oncology panel, including 92 oligonucleotide-labeled antibody probe pairs that can bind to their respective targets in the sample and can be detected and quantified using standard real-time PCR. Additional details about the 92 markers, detection range, data normalization and standardization are available at https://www.olink.com/resources-support/document-download-center/. Analysis of the samples was performed in R (R Foundation for Statistical Computing).

Endpoints. In OpACIN-neo, primary endpoints were grade 3–4 toxicity rate in the first 12 weeks, radiologic response rate (according to RECIST 1.1.) and pRR (according to INMC criteria)[32]. Secondary endpoints included RFS, description of late and ongoing AEs and associations between mutational load and RNA signatures with response. EFS was an exploratory endpoint.

In OpACIN, primary endpoints were safety and feasibility and the comparison of the immune-activating capacity of neoadjuvant versus adjuvant arm. Safety was measured by the frequency of suspected unexpected serious adverse reactions in both arms, and feasibility (of the neoadjuvant arm only) was measured by execution of the total lymph node dissection at the preplanned time point. Secondary endpoints included RFS and rate and type of AEs.

RFS was defined as time from surgery until date of first relapse (local or distant metastasis) or death from any cause, whichever occurred first. EFS was defined as time from randomization until date of progression during neoadjuvant therapy, precluding surgery (distant metastases or local progression when unresectable), relapse or death of any cause.

Statistics. Primary endpoints were summarized by frequency (per treatment arm) with corresponding two-sided 95% CI calculated using the Clopper–Pearson method. Comparisons of frequencies between treatment arms were performed using Fisher's exact test. Median follow-up time was calculated using the inverse Kaplan–Meier method. We used the Kaplan–Meier method to estimate RFS and EFS. The log-rank test was used to compare differences between treatment arms and between patients with and without a pathologic response. The 95% CIs were computed using log transformation. The phase 1b OpACIN trial was not powered to compare response or RFS between the neoadjuvant and adjuvant arm.

The probability of achieving a pathologic response based on IFN-γ signature expression score and/or TMB was examined by univariable and multivariable logistic regression analyses based on whichever sROC curves were computed. The relative AUC values were determined as a global metric of the ability of

each biomarker to discriminate between patients with and without a pathologic response. Optimal cutoffs were computed using the cutpointr package in R (maximize metric).

Differences in translational endpoints between patients with a pathologic response and those without a response and patients with and without a relapse were analyzed using the Mann–Whitney *U* test (Wilcoxon's rank-sum test), and the correlation between biomarkers and EFS was calculated using a log-rank test. Changes in cytokines between pre- and post-treatment samples were analyzed using a paired Student's *t*-test and the Welch's *t*-test (*P* value and FDR were calculated).

Analyses were performed in R (version 3.6.3) and R Studio (version 1.2.1335) using the packages tidyverse (version 1.3.0), survival (version 2.44-1.1.), ggplot2 (version 3.2.1), survminer (version 0.4.5), stats (version 3.6.1), pROC (version 1.16.2), cutpointr (version 1.0.32), heatmap.plus (version 1.3) and RColorBrewer (version 1.1.-2). Dot plots and bar plots were generated in GraphPad Prism (version 7.03).

Trial oversight. The protocol and amendments were reviewed and approved by the appropriated review boards and ethics committees of each of the three participating centers (OpACIN-neo) or only the Netherlands Cancer Institute (OpACIN). The studies were conducted in accordance with Good Clinical Practice guidelines as defined by the International Conference of Harmonization and the Declaration of Helsinki. All patients provided written informed consent before enrollment. Both investigator-initiated trials were funded by Bristol Myers Squibb with the Netherlands Cancer Institute as the sponsor. Data were collected by the sponsor and analyzed in collaboration with all authors. The OpACIN-neo was monitored by a data safety monitoring board. A two-stage Simon design was applied to stop the trial early for futility in case of a low proportion of patients with a pathologic response. After the report of serious AEs from a patient with severe colitis in arm C who required a colectomy, the data safety monitoring board required an interim safety analysis. Based on this analysis, they advised premature closure of arm C because of a high incidence of >grade 3 irAEs.

The authors declare the completeness and accuracy of the data and adherence to the trial protocol. The database lock for the presented analysis took place on 6 February 2020 (OpACIN-neo) and 8 May 2020 (OpACIN).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

DNA-sequencing and RNA-sequencing data generated during the study will be deposited in the European Genome-phenome Archive (EGA) under the accession codes EGAS00001004832 (DNA) and EGAS00001004833 (RNA), and will be made available on reasonable request. Data requests will be reviewed by the institutional review board of the Netherlands Cancer Institute and applying researchers will need to sign a data access agreement after approval.

References

32. Tetzlaff, M. T. et al. Pathological assessment of resection specimens after neoadjuvant therapy for metastatic melanoma. *Ann. Oncol.* **29**, 1861–1868 (2018).
33. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
34. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
35. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
36. Liberzon, A. et al. The molecular signatures database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
37. Subramanian, A. et al. Gene-set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
38. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
39. Van der Auwera, G. A. et al. From FASTQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10. 33 (2013).

Acknowledgements

We thank the patients and their families for participating in the study. We thank all investigators and members of the clinical trial teams in Melanoma Institute Australia, the Netherlands Cancer Institute and the Karolinska Institutet; the Netherlands Cancer Institute-Antoni van Leeuwenhoek Core Facility Molecular Pathology & Biobanking for supplying biobank material and/or laboratory support; the Genomics Core Facility for their support regarding sequencing; and S. Vanhoutvin for financial management. We acknowledge A. Evans and B. Stegenga from Bristol Myers Squibb for scientific input and support. G.V.L. is supported by an Australian National Health and Medical Research Council (NHMRC) Practitioner Fellowship and the Medical Foundation at the University of Sydney. A.M.M. is supported by Cancer Institute New South Wales fellowship and Melanoma Institute Australia. R.A.S. is supported by NHMRC Practitioner Fellowship. Support from an Australian NHMRC program grant (to G.V.L. and R.A.S.), the Ainsworth Foundation, the Fairfax Foundation and the Cameron family is also gratefully acknowledged. The authors also acknowledge assistance from colleagues at their various institutions.

Author contributions

C.U.B. designed the study and wrote the study protocol. A.C.J.v.A. and T.N.M.S. provided additional input to the study design. E.A.R. coordinated the trial, analyzed and interpreted clinical and translational data and wrote the first draft of the manuscript with E.P.H., I.L.M.R. and J.M.V. C.U.B. co-wrote the manuscript. E.A.R., R.P.M.S., H.E., K. Shannon, J.B.A.G.H., J.S., S. Ch'ng, O.E.N., H.A.M., S.A., W.M.C.K., C.L.Z., W.J.v.H., A.J.S., A.C.J.v.A., A.M.M., G.V.L. and C.U.B. recruited and treated patients and collected data. B.A.v.d.W. and R.A.S. reviewed and scored the pathology of all cases, including grading pathologic responses. P.D. and O.K. (under the supervision of D.S.P.) performed the bioinformatics analysis. K. Sikorska performed the statistical analysis. A.T.A. and L.G.G.-O. were responsible for central and local data management. M.G. was a clinical project manager involved in the trial. E.P.H. performed the analysis of the plasma proteomics data. R.M.K. was responsible for the sequencing. S. Cornelissen performed DNA and RNA isolations. A.B. coordinated and contributed to translational laboratory logistics and immunohistochemistry and molecular laboratory work. Every author contributed to the initial draft of the manuscript and agreed on submission for publication. All authors interpreted the data, reviewed the manuscript and approved the final version.

Competing interests

E.P.H., I.L.M.R., J.M.V., O.K., P.D., K. Sikorska, H.E., M.G., A.T.A., L.G.G.-O., K. Shannon, J.S., S. Ch'ng, O.E.N., H.A.M., S.A., R.M.K., S. Cornelissen, A.B., W.M.C.K., C.L.Z., W.J.v.H. and A.J.S. declared no competing interests. E.A.R. reports travel support from NanoString Technologies and MSD. R.P.M.S. has served on advisory boards for MSD Novartis Qbiotics and received honoraria from BMS. B.A.v.d.W. has served on advisory boards for BMS. J.B.A.G.H. has served on advisory boards for AIMM, Achilles, AZ/MEDimmune, Amgen, Bayer, BMS, GSK, Ipsen, Immunocore, MSD, Merck Sorono, Neon Therapeutics, Neogene Therapeutics, Novartis, Pfizer, Roche/Genentech, Sanofi, Seattle Genetics, Third Rock Ventures and Vaximm, and reports research fees paid to the institute from BMS, MSD, Novartis and Neon Therapeutics. D.S.P. is cofounder, shareholder and advisor of Immagene BV. A.C.J.v.A. has served on advisory boards for Amgen, BMS, Novartis, MSD-Merck, Merck-Pfizer, Sanofi and 4SC, and reports research fees paid to the institute from Amgen, BMS and Merck-Pfizer. R.A.S. reports financial support from Qbiotics, Novartis, NeraCare, AMGEN, BMS, Myriad Genetics, GlaxoSmithKline and Merck Sharp & Dohme. T.N.M.S. has served on advisory boards for Adaptive Biotechnologies, AIMM Therapeutics, Allogene Therapeutics, Merus, BioNTech, Scenic Biotech; reports financial support from Merck KGaA; and is stockholder in AIMM Therapeutics, Allogene Therapeutics, Merus, Neogene Therapeutics, BioNTech and Scenic Biotech. A.M.M. has served on advisory boards for BMS, MSD, Novartis, Roche, Pierre Fabre and QBiotics. G.V.L. has served on advisory boards for Aduro, Amgen, BMS, Highlight Therapeutics S.L., Mass-Array, Merck, MSD, Novartis, OncoSec Medical, Pierre Fabre, Roche, Qbiotics and Sandoz. C.U.B. has served on advisory boards for BMS, MSD, Roche, Novartis, GSK, AZ, Pfizer, Lilly, GenMab, Pierre Fabre and Third Rock Ventures, for which the institute received funding; received research funding from BMS, Novartis and NanoString, all paid to the institute; stock ownership: Uniti Cars; and is cofounder of Immagene BV.

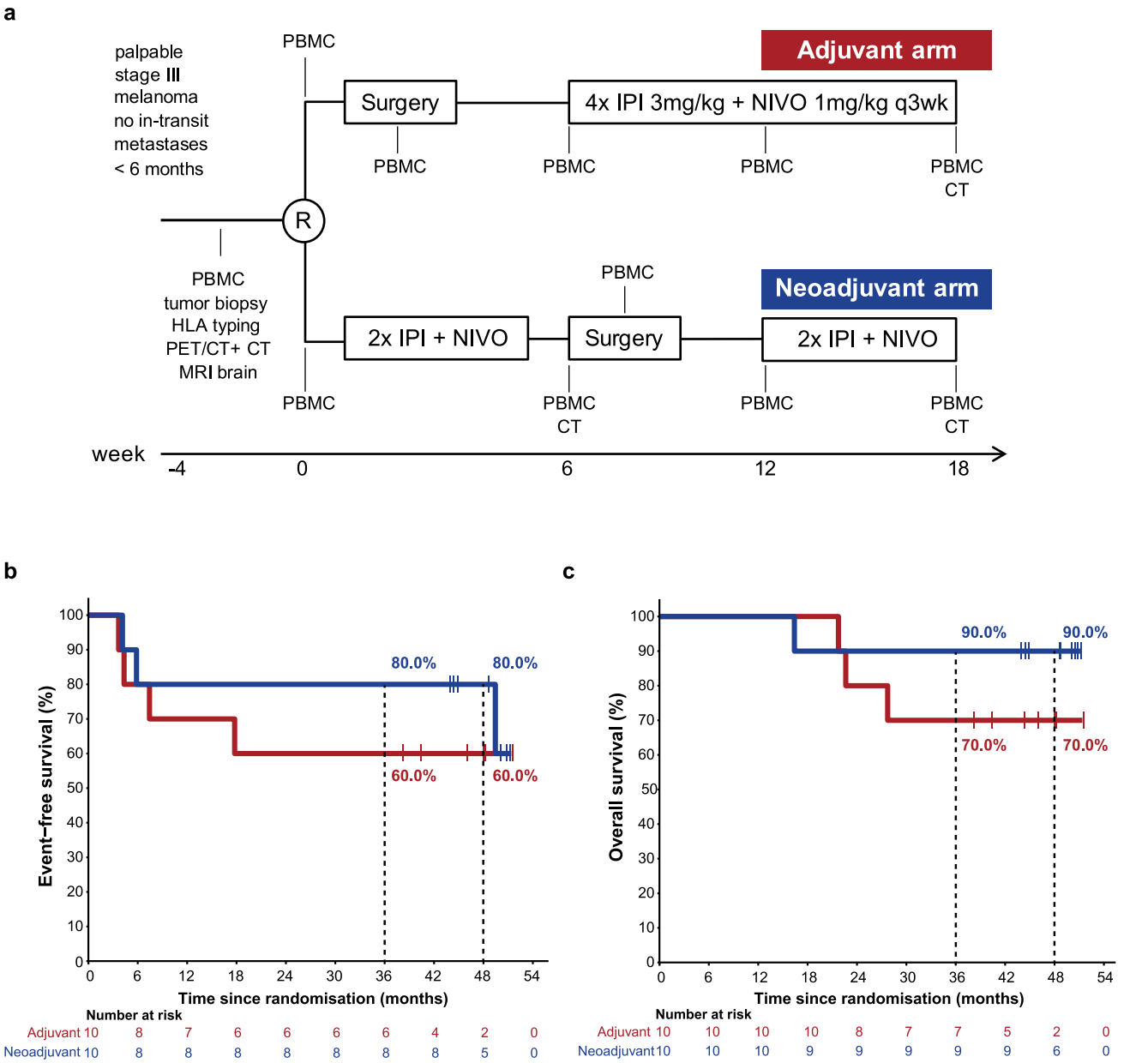
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-020-01211-7.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-020-01211-7.

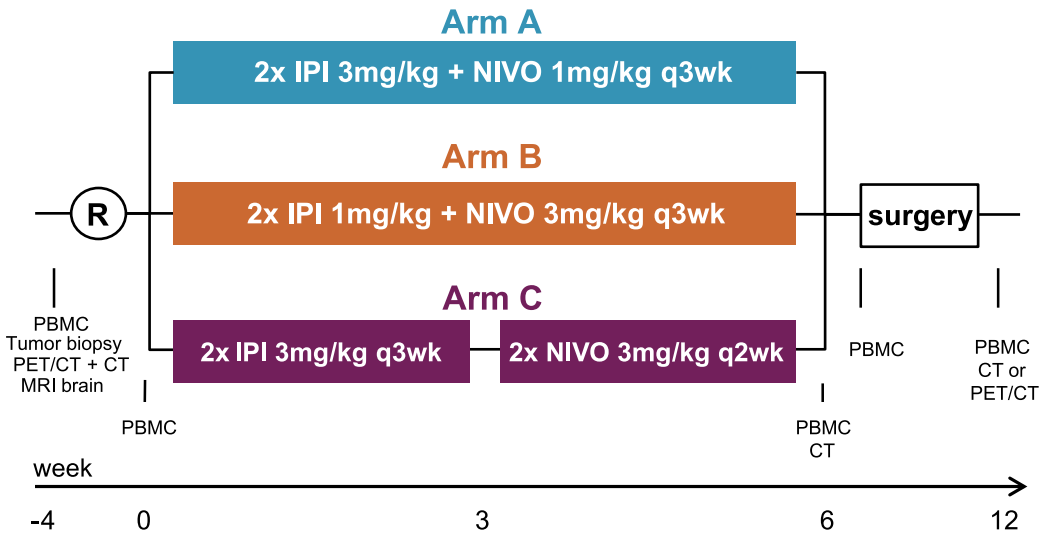
Correspondence and requests for materials should be addressed to C.U.B.

Peer review information Javier Carmona was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

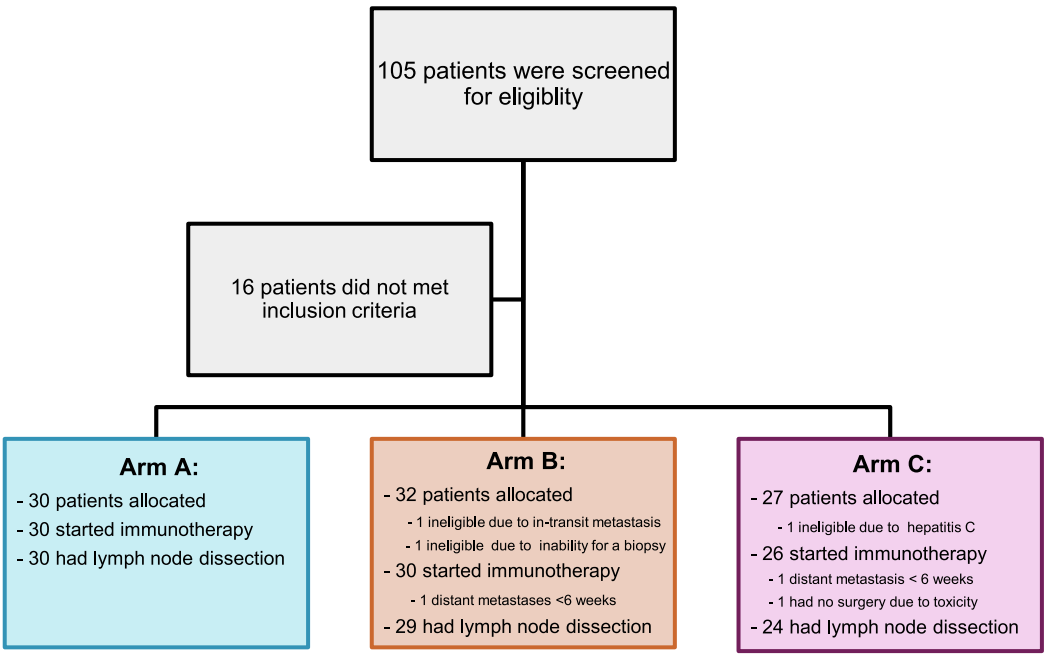


Extended Data Fig. 1 | Study design OpACIN, event-free survival and overall survival of OpACIN. a, Study design of the OpACIN study. Patients were randomized to receive 4 cycles of ipilimumab 3 mg/kg + nivolumab 1 mg/kg every 3 weeks after surgery (adjuvant arm, n = 10) or 2 cycles of ipilimumab 3 mg/kg + nivolumab 1 mg/kg every 3 weeks followed by surgery and thereafter again 2 cycles of ipilimumab 3 mg/kg + nivolumab 1 mg/kg (neoadjuvant arm, n = 10). A biopsy was taken at screening and blood samples were taken at screening, baseline, week 6, week 12 and week 18. IPI; ipilimumab, NIVO; nivolumab, PBMC; peripheral blood mononuclear cells. **b**, Event-free survival by treatment arm and **c**, Overall survival by treatment arm of the OpACIN study. Kaplan-Meier curves were generated including all patients from the adjuvant arm (red, n = 10) and neoadjuvant arm (blue, n = 10).

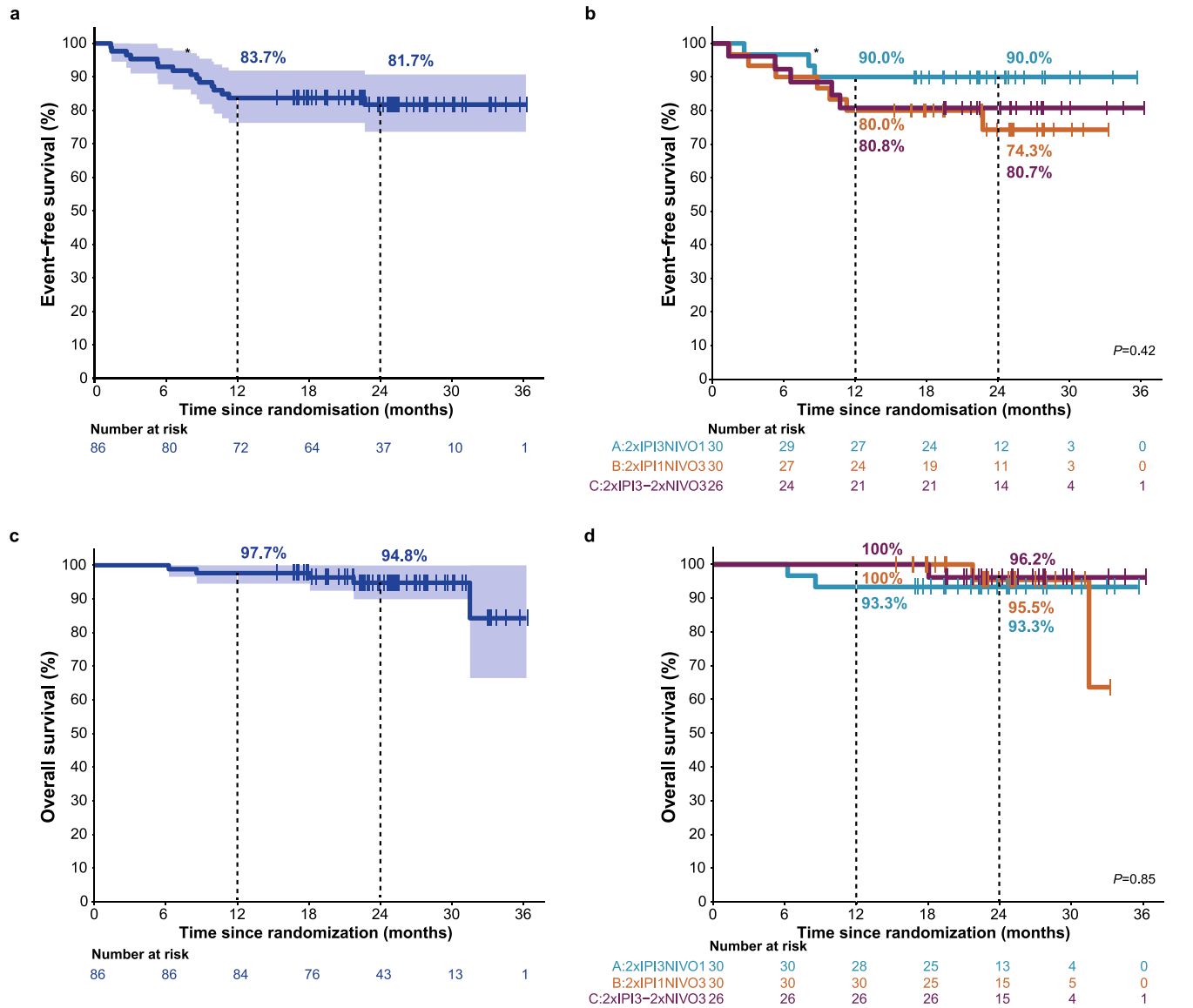
a



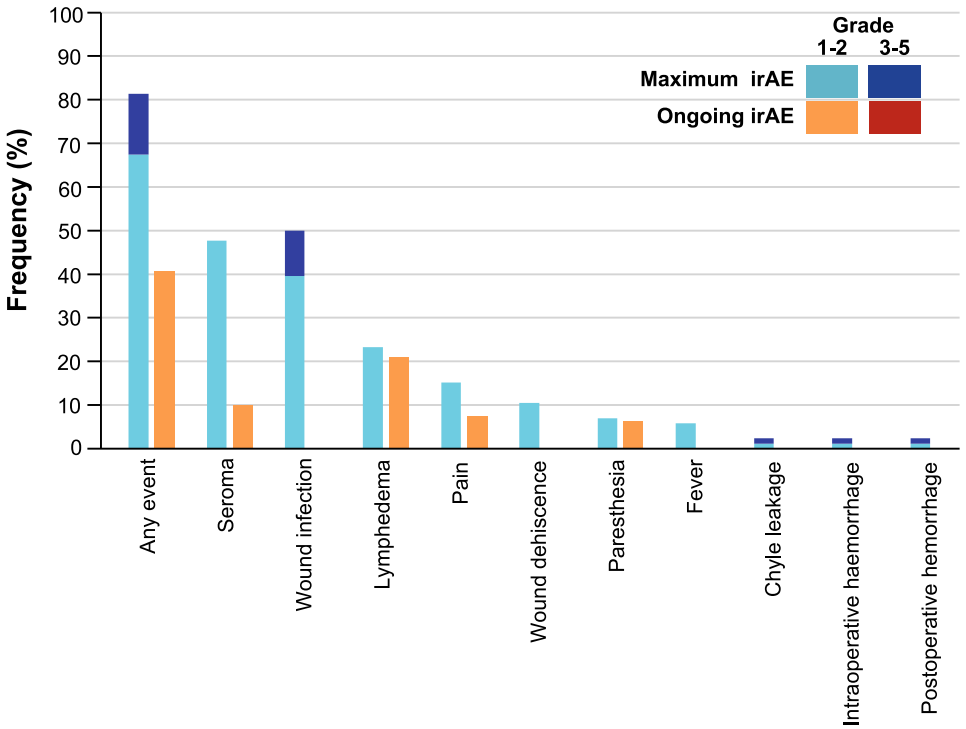
b



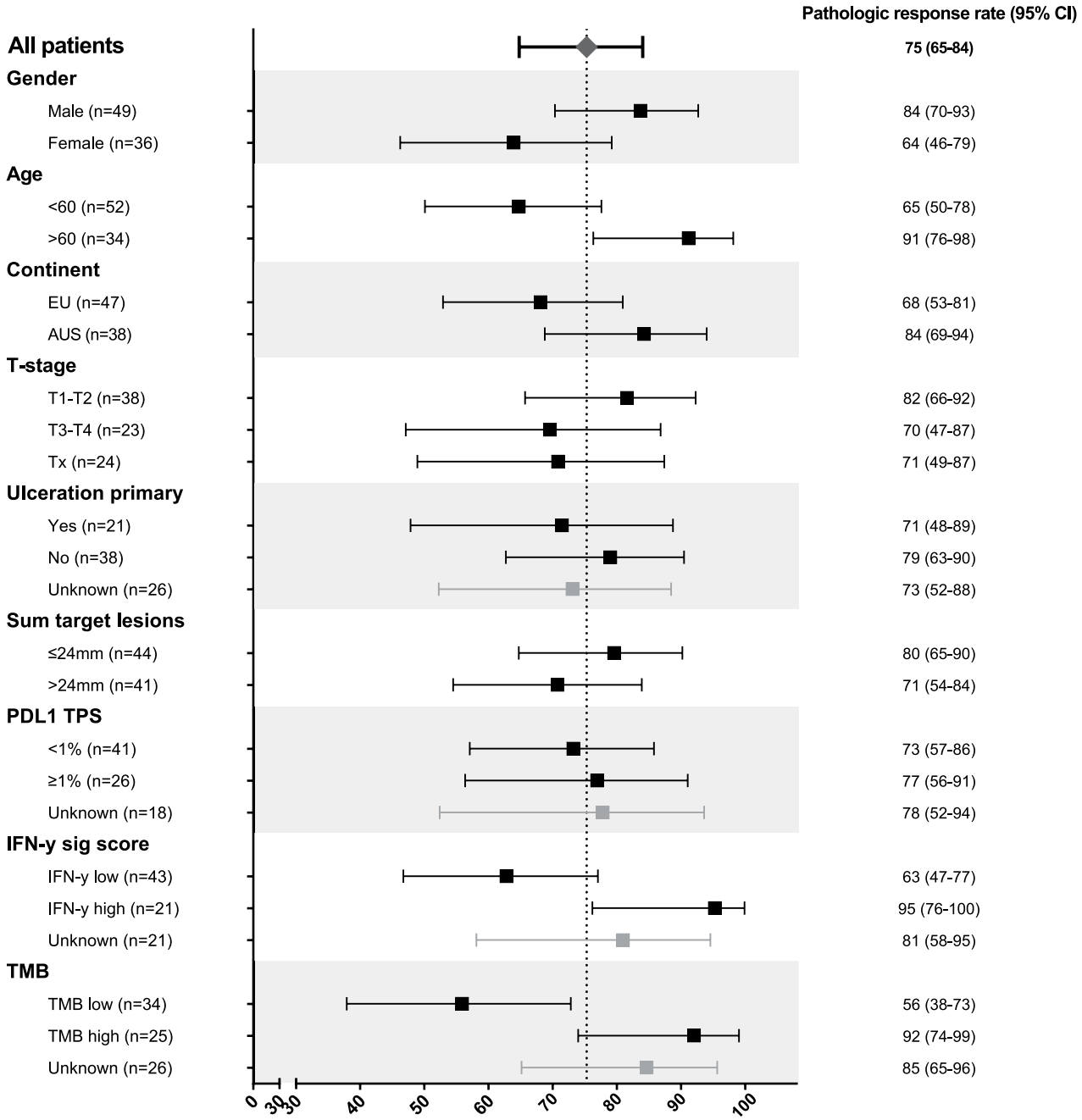
Extended Data Fig. 2 | Study design and flowchart OpACIN-neo. a, Study design of the OpACIN-neo study. Patients were randomized to receive 2 cycles of ipilimumab 3 mg/kg + nivolumab 1 mg/kg every 3 weeks (arm A, n = 30), 2 cycles of ipilimumab 1 mg/kg + nivolumab 3 mg/kg every 3 weeks (arm B, n = 30) or 2 cycles of ipilimumab 3 mg/kg every 3 weeks directly followed by 2 cycles nivolumab 3 mg/kg every 2 weeks (arm C, n = 26). Surgery was planned after 6 weeks. A biopsy was taken at screening and blood samples were taken at screening, baseline, week 6 and week 12. IPI; ipilimumab, NIVO; nivolumab, PBMC; peripheral blood mononuclear cells. **b**, Flowchart of the OpACIN-neo study showing the number of patients screened, allocated to a treatment arm, starting immunotherapy and undergoing surgery per treatment arm.



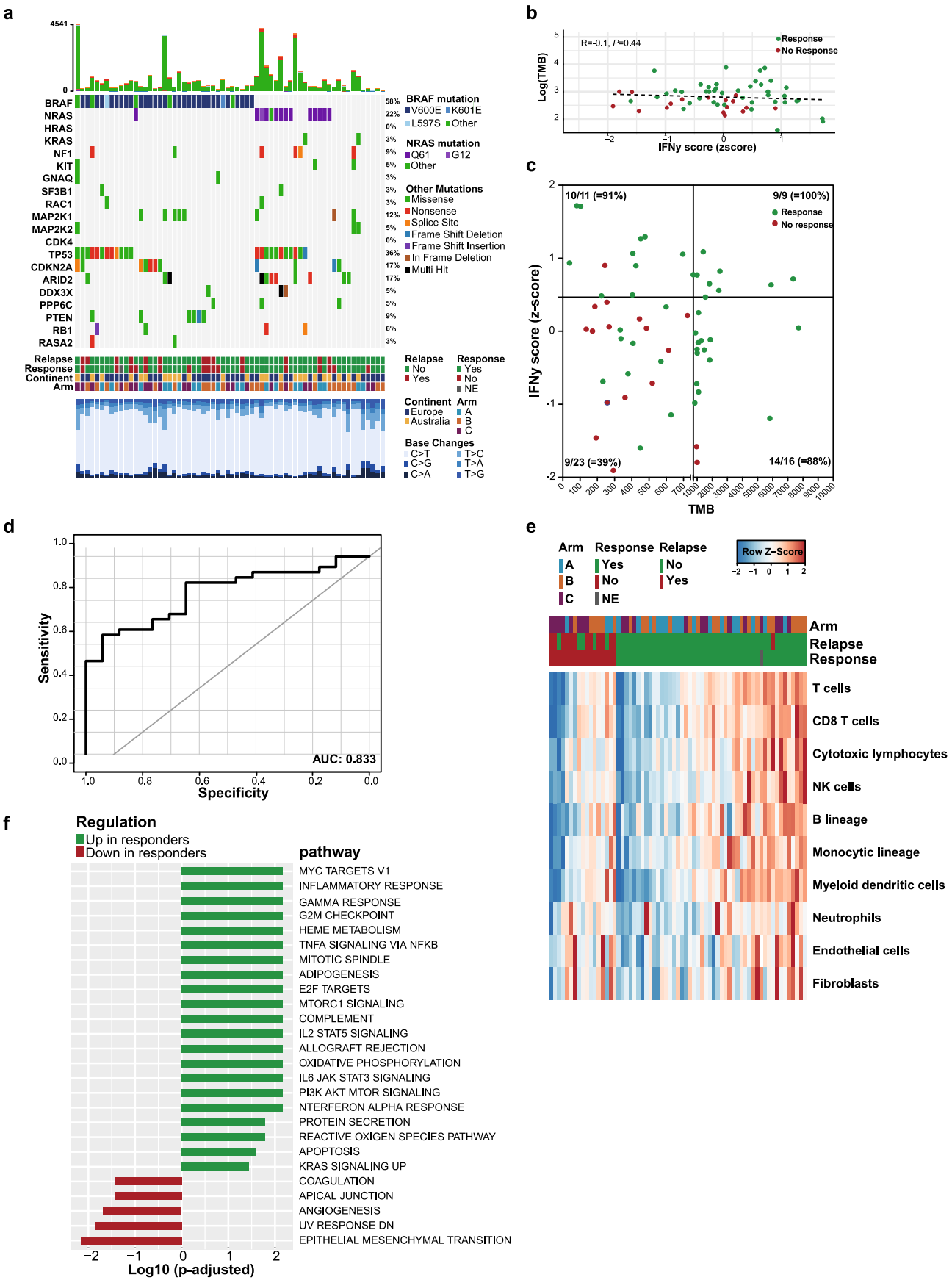
Extended Data Fig. 3 | Event-free survival and overall survival of OpACIN-neo. a, EFS for the total population of the OpACIN-neo study. A Kaplan-Meier curve for EFS of all patients (n = 86) was generated. The corresponding 95% CI is displayed and was computed using log transformation. **b**, EFS of the OpACIN-neo study by treatment arm including all patients from arm A (blue, n = 30), arm B (orange, n = 30) and arm C (purple, n = 26). *P* values were calculated using the log-rank test (two-sided). **c**, OS for the total population of the OpACIN-neo study. A Kaplan-Meier curve for OS of all patients (n = 86) was generated. **d**, OS of the OpACIN-neo study by treatment arm including all patients from arm A (blue, n = 30), arm B (orange, n = 30) and arm C (purple, n = 26). **a-b**, The asterisk denotes the patient who died due to irAEs.



Extended Data Fig. 4 | Ongoing surgery-related adverse events of OpACIN-neo. Frequency of maximum grade and ongoing surgery-related adverse events (AEs) of the OpACIN-neo study. Frequencies of maximum grade AEs are displayed in light blue (grade 1–2) and dark blue (grade 3–5), and frequencies of ongoing AEs in orange (grade 1–2) and red (grade 3–5). AEs that were reported at a frequency of >5% and all grade 3–5 AEs were included. All patients (n = 86) were included in the analysis of maximum grade AEs; for ongoing AEs only patients alive at time of data cutoff (n = 81) were included.

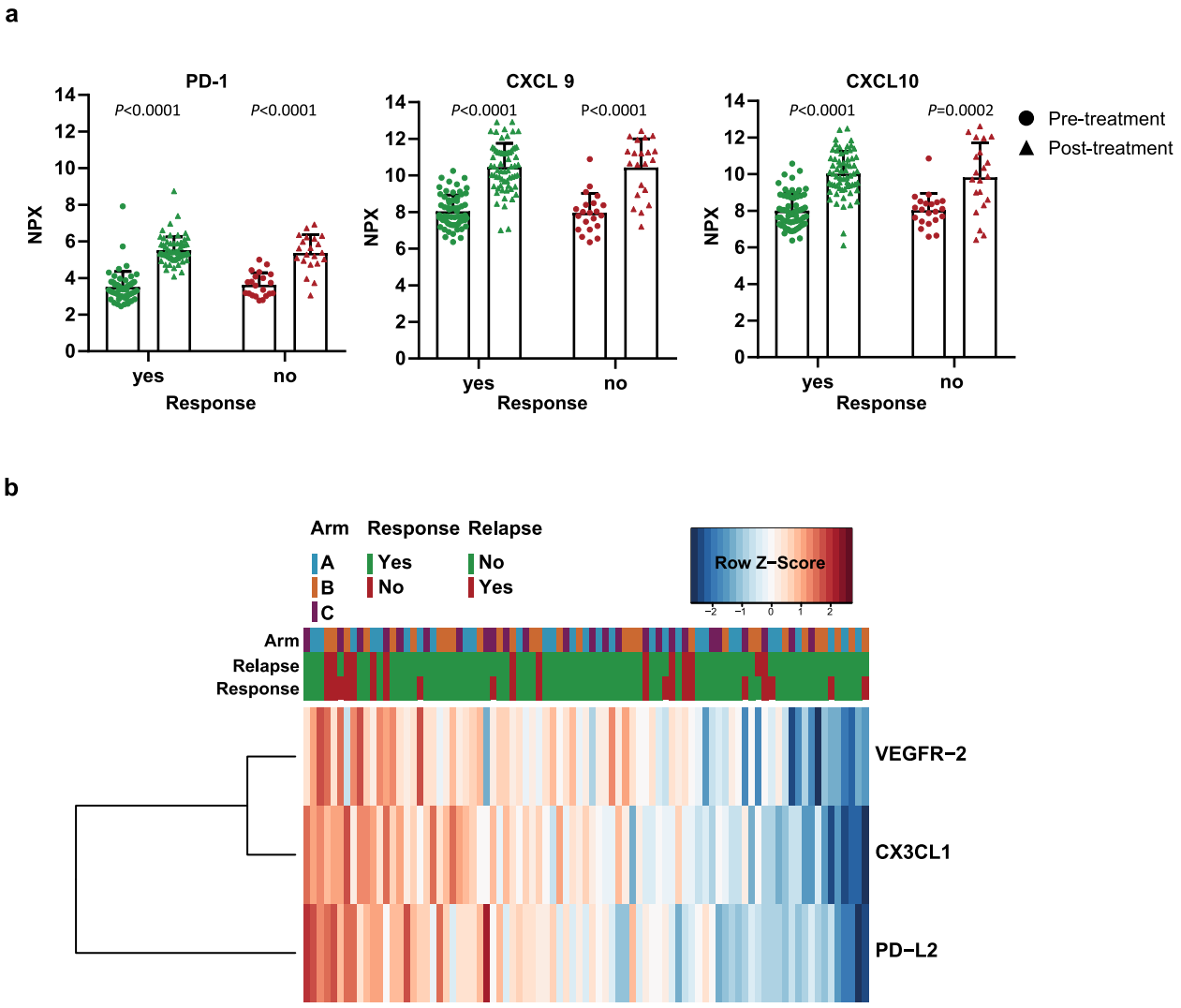


Extended Data Fig. 5 | Pathologic response rates according to subgroups. Forest plot of data for all patients who underwent surgery (n = 85). pRRs according to demographic, clinical and tumor characteristics are displayed. The 95% CIs were calculated using the Clopper-Pearson method. PD-L1 expression on pretreatment tumor biopsies was assessed centrally with an automated lab-validated immunohistochemistry assay, using the 22C3 antibody on a Ventana platform. PD-L1 expression was determined by the Tumor Proportion Score (TPS; the percentage of tumor cells with complete or partial membranous staining at any intensity).



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Whole-exome sequencing and RNA sequencing analysis of pretreatment tumor biopsies. a, Mutational load and mutational patterns of recurrent mutated cutaneous melanoma genes found by whole-exome sequencing. The frequency, mutation type and base changes are indicated. Each column represents one patient (n = 60 patients). **b**, Correlation between the IFN-γ score (values displayed as the average *z*-score of all the genes within the IFN-γ signature[14]) and TMB (displayed in log scale) for patients with pathologic response (n = 42, green) and no pathologic response (n = 17, red). The correlation coefficient and *P* value were calculated using the Pearson's correlation method. **c**, TMB and baseline average expression of IFN-γ score of patients with a response (green dots) and patients without a response (red dots). The quadrants are determined by the optimal cutoff for each of the biomarkers as defined by the sROC curves. Each quadrant indicates the number of responding patients and total number of patients in the corresponding quadrant. Data were available for 59 patients. **d**, sROC curve showing the AUC for the combination of the IFN-γ score and TMB (0.83) (n = 59). **e**, Heatmap of the MCP counter[15] RNA gene signature ordered according to average signature expression per response category of baseline tumor biopsies (n = 65). The MCP counter signature expresses the abundance of eight immune and two stromal cell populations. Each cell type is represented by the averaged *z*-score of the genes that it is consisted of, which were previously normalized by DESeq2. The score was computed from the average expression of all the ten cell types that form the MCP counter signature. Columns represent patients (green: pathological response/no relapse; red: no pathological response/relapse; grey: not evaluable (NE); blue: treatment arm A; orange: treatment arm B; purple: treatment arm C) and rows represent genes. Positive values (red) indicate higher expression and negative (blue) indicate lower expression. **f**, Gene set enrichment analysis displaying hallmark gene sets that are significantly enriched in responders (green) or nonresponders (red). Pathways are ordered according to the FDR. FDRs were computed as previously described[37].



Extended Data Fig. 7 | Extended plasma analysis using Olink proteomic assay. a, PDCD1, CXCL9, CXCL10 normalized protein expression (NPX) in plasma of patients (n = 85) before treatment (pretreatment; round dots) and after treatment (post-treatment; triangle dots) measured with Olink immunoassay (green: patients with pathologic response, n = 64; red: patients without pathologic response, n = 21). Data for PDCD1 are missing for pretreatment samples of 9 patients (all responders) because values were below detection limit. The mean and SD are shown. *P* values were calculated using the paired Student's t-test (two-sided). **b**, Heatmap of VEGFR2, CX3CL1 and PD-L2 NPX in plasma of patients before start of treatment. The heatmap depicts the ordered mean expression of these three genes of the 86 patients included in the OpACIN-neo cohort. The score of each patient expresses the baseline averaged *z*-score of the three mentioned genes mentioned beforehand which was previously normalized by DESeq2. Each column represents a different patient (green: response/no relapse; red: no response/relapse; blue: arm A; orange: arm B; purple: arm C) and rows indicate protein expression. Positive values (red) indicate higher expression and negative values (blue) indicate lower expression.

nature research

Corresponding author(s): Prof. Dr. C.U. Blank

Last updated by author(s): December 3, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The Clinical data was collected and processed in electronic case report forms (eCRF) using Tenalea at the Netherlands Cancer Institute, Melanoma Institute Australia and Karolinska Institute. Sequecing data (DNA and RNA) were collected as Fastq files; derived features (e.g. interferon-gamma signature, tumor mutational burden) were collected in .pdf and .tsv files. The plasma proteomics assays using proximity extension assay (PEA) technology were collected in .xlsx files.
Data analysis	Statisticall analysis: R version 3.6.3 and R studio version 1.2.1335 DNA (whole exome) sequencing: Illumina bcl2fastq (v2.20), Skewer (v0.2.2), FastQC (v0.11.5-cegat), Burrows-Wheeler aligner (v0.7.12), Picard MarkDuplicates (v1.140), VariantAnnotation (v1.30.1), Ensemble-vep release 100 Samtools (v1.10) GATK (v4.0.6.0), TVTB (v1.10.0), R packages: tidyverse (v1.3.0), maftools (v2.0.16), seqinr (v3.6.1), snowfall (v1.84.6.1). RNA sequencing: Star (v2.7), Samtools (v1.10) R packages: DESeq2 (v1.24.0), HTseq (v0.11.1) tidyverse (v1.3.0), fgsea (v1.10.1), MCPcounter (v1.1.0), biomaRt (v2.40.5), RColorBrewer (v1.1-2), heatmap.3 (github/lwaldron/LeviRmisc/), ggplot2 (v3.2.1), pROC (v1.16.2), cutpointr (v1.0.32). Plasma proteomics assay: R packages: heatmap.plus (v1.3), RColorBrewer (v1.1-2), devtools (v2.3.1), ggplot2 (v3.3.2). Survival analysis: R packages: survival (v2.44-1.1), ggplot2 (v3.2.1), survminer (v0.4.5), stats (v3.6.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

DNA and RNA sequencing data will be deposited in the European Genome-phenome Archive under the accession codes EGAS00001004832 (DNA) and EGAS00001004833 (RNA), and will be made available on reasonable request for academic use and within the limitations of the provided informed consent by the corresponding author upon acceptance. Every request will be reviewed by the institutional review board of the NKI; the researcher will need to sign a data access agreement with the NKI after approval. Sequencing data is shown in Fig. 2, Fig. 4 and Supplemental Fig. 8.
Plasma proteomic assay data are available from the corresponding author on reasonable request. This data was used in Fig. 3 and Supplemental Fig. 9.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	OpACIN study: The primary endpoint of the study is in particular the safety and feasibility of intermittent surgery during immunotherapy with nivolumab plus ipilimumab (neo-adjuvant arm). This needs to be contrasted to a therapy where the combination is given as adjuvant therapy (also experimental). Therefore, 20 patients were randomized to either receiving the combination of ipilimumab + nivolumab adjuvant, or to split neo-adjuvant and adjuvant with surgery in between (10 patients per arm). The study was defined as not safe and feasible, if 2 out of the first 5 patients (point estimate 0.4 (95%CI 0.05-0.85)) or 4 out of the 10 (point estimate 0.4 (95%CI 0.12-0.74)) patients in the neo-adjuvant arm would have experienced immune-related adverse events leading to delayed surgery (not performed during week 6) or experience grade 3/4 SUSARs after surgery, that are attributed to the pre-surgery immunotherapy. The investigators realized that numbers of immune-related adverse events smaller than respectively 2 and 4 still bare a substantial chance of error of taking the wrong conclusion about safety. The number of ten patients in each arm is chosen with the focus on producing relevant numbers of T cell responses that can be analyzed (immune-activating capacity). OpACIN-neo study: the primary objectives of this trial were to compare the frequencies of grade 3-4 immune-related adverse events (irAEs) and to compare efficacy (response rate) of three different neoadjuvant schemes of ipilimumab plus nivolumab. Based on the previous OpACIN study, a true grade 3-4 irAE rate of 89% was assumed, and any grade toxicity rate of 100% and a pathologic response rate of 80% for patients in Arm A. Assuming 89% grade 3/4 toxicity of the current schedule, a Fisher's exact test with two groups of 30 patients each will have 89% power at a 0.05 two-sided significance level to detect a difference in toxicity to a proportion grade 3/4 toxicity of 50% or lower. Assuming a pathological response rate of 80% with the current schedule, the lower 95% confidence interval of the response rate in a group of 30 patients will extend to 61% (based on an exact binomial test). A response rate lower than that suggests that the alternative schedule is less effective. In this setting, arm B will be compared to arm A and arm C will be compared to arm A.
Data exclusions	OpACIN-neo: 89 patients were randomized and 3 patients were excluded after randomization before receiving treatment because they did not meet in- and exclusion criteria (1 patient had intransit metastases, 1 patient had only lesions that could not be biopsied and 1 patient turned out to have positive serology for hepatitis C). Inclusion and exlcusion criteria were prespecified in the study protocol. Samples that were obtained for biomarker analyses were excluded if they did not pass quality control (eg. tumor cell percentage <30% or too low amount of DNA or RNA)
Replication	Experimental replicates were not attempted. Due to the scarce patient material replication of experiments was not possible.
Randomization	OpACIN: Patients were randomized by the independent trial office of the NKI. OpACIN-neo: Patients fulfilling the eligibility criteria were enrolled in the study. Patients were randomized using ALEA randomisation software, which implements a minimisation technique described by Pocock and Simon. Patients were stratified according to study centre.
Blinding	Both the OpACIN and OpACIN-neo were open-label studies; the investigators, site staff, and patients were aware of the treatment assignment during the study participation. As different dosing schemes were tested for the first time in this patient population for safety and efficacy, it would be essential for physicians to know the treatment regimen. In that way they could notify if they would notice signals of higher toxicity or lower efficacy in any of the arms. Pathologists were blinded for clinical outcome and treatment arm.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: Plasma proteomics analysis was performed using the Olink® Immuno-Oncology panel. This was a paid service by Olink proteomics. Specific antibody clones were not disclosed.

Validation: Additional details about the 92 markers, detection range, data normalization and standardization are available at https://www.olink.com/resources-support/document-download-center/.

Human research participants

Policy information about studies involving human research participants

Population characteristics: Resectable stage III melanoma patients with one or more measurable lymph node metastases (according to RECIST v1.1) that can be biopsied, no history of in-transit metastases within the last 6 months, naïve for CTLA-4/PD-1/PD-L1 immunotherapy, and more than 18 years old.
Of all the patients included the median age was 57.5 years and 57% was male, 99% had an ECOG performance status of 0 and only 1% had an elevated LDH level at baseline.

Recruitment: Patients were recruited by either surgical oncologists or medical oncologist from the melanoma cancer clinics in the Netherlands Cancer Institute (OpACIN and OpACIN-neo), Melanoma Institute Australia (OpACIN-neo) and Karolinska Institute (OpACIN-neo). Patients were generally referred to the participating study centers by outside hospitals. No specific bias in recruitment was identified.

Ethics oversight: Medical ethics review comittee of the Netherlands Cancer Institute and ethical committees at Melanoma Institute Australia and Karolinska Institute approved the trial.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration: OpACIN: NCT02437279
OpACIN-neo: NCT02977052

Study protocol: The full trial protocol of both trials can be found in the supplementary appendix.

Data collection: For OpACIN, patients were enrolled between August 2015 and October 2016 in the Netherlands Cancer Institute.
For OpACIN-neo, patients were enrolled between November 2016 and June 2018 in the Netherlands Cancer Institute, Melanoma Institute Australia and Karolinsta Institute.

The data cut-offs in the current manuscript were 8 MAY 2020 (OpACIN) and 6 FEB 2020 (OpACIN-neo).
Clinical data was collected through an eCRF by the clinical trial department of the Netherlands Cancer Institute. Clinical data was analyzed by the department of biostatistics at the Netherlands Cancer Institute.

Outcomes: OpACIN
The primary endpoint of the study is the safety and feasibility of intermittent surgery during immunotherapy with nivolumab plus ipilimumab (neo-adjuvant arm). The study was defined as not safe and feasible, if 2 out of the first 5 patients (point estimate 0.4 (95%CI 0.05-0.85)) or 4 out of the 10 (point estimate 0.4 (95%CI 0.12-0.74)) patients in the neoadjuvant arm would have experienced immune-related adverse events leading to delayed surgery (not performed during week 6) or experience grade 3/4 SUSARs after surgery, that were attributed to the presurgical immunotherapy.
Secondary outcomes included:
- RFS, determined according to RECIST 1.1 criteria.
- Rate and type of adverse events and late adverse events

OpACIN-neo
Primary outcomes:

- Safety as measured by the frequency of grade 3/4 immune-related adverse events (during the first 12 weeks).
- Response rate according to RECIST 1.1 at week 6
- Pathologic response according to central pathologic revision

Secondary outcomes:
- Recurrence-free survivial
- Description of late adverse event (irAE, up to 3 years after treatment initiation, until new treatment) according to CTCAE v4.03
- Description of associations of mutational load and RNA tumor signatures with tumor immune infiltrates and response
- Alteration in magnitude or breadth of the neo-antigen specific T cell responses in peripheral blood from baseline to surgery at week 6 in each 10 randomly chosen patients per arm.

Definitions of pathologic response and survival outcomes are described in supplemental table 9.