Systemic immune dynamics in cancer

Bakker, E.A.M.

CHAPTER 6

Neoadjuvant nivolumab or nivolumab plus ipilimumab in early-stage triple negative breast cancer: a phase 2 adaptive trial

Iris Nederlof[1,*], Olga I. Isaeva[1,*], Manon de Graaf[1,#], Robbert C.A.M. Gielen[1,#], Noor A. M. Bakker[1,2,#], Adrianne L. Rolfes[1], Hannah Garner[1,2], Bram Boeckx[3,4], Joleen J.H. Traets[1], Ingrid A.M. Mandjes[5], Michiel de Maaker[6], Thomas van Brussel[3,4], Maksim Chelushkin[1], Elisa Champhanet[1], Marta Lopez-Yurda[5], Koen van de Vijver[7,8], José G. van den Berg[8], Ingrid Hofland[6], Natasja Kliouеva[9], Ritse M. Mann[10], Claudette E. Loo[10], Frederieke H. van Duijnhoven[11], Victoria Skinner[11], Sylvia Luykx[12], Emile Kerver[13], Ekaterina Kalashnikova[14], Marloes G. J. van Dongen[15], Gabe S. Sonke[15], Sabine C. Linn[15], Christian U. Blank[15,16], Karin E. de Visser[1,2,17], Roberto Salgado[18,19], Lodewyk F.A. Wessels[2,20], Caroline A. Drukker[11], Ton N. Schumacher[2,16,21], Hugo M. Horlings[8], Diether Lambrechts[3,4], Marleen Kok[1,15]

* and # These authors contributed equally
Corresponding author: Marleen Kok, m.kok@nki.nl

Nature Medicine. 2024 August 14; 30(3223–3235). DOI: 10.1038/s41591-024-03249-3

1 Division of Tumor Biology and Immunology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
2 Oncode Institute, Utrecht, the Netherlands
3 Laboratory for Translational Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium
4 VIB Center for Cancer Biology, Leuven, Belgium
5 Biometrics Department, The Netherlands Cancer Institute, Amsterdam, the Netherlands
6 Core Facility Molecular Pathology & Biobanking, The Netherlands Cancer Institute, Amsterdam, the Netherlands
7 Department of Pathology, UZ Gent - Universitair Ziekenhuis Gent, Gent, Belgium
8 Department of Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
9 Department of Pathology, OLVG Hospital, Amsterdam, Netherlands
10 Department of Radiology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
11 Department of Surgical Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
12 Medical Oncology department, Tergooi Hospital - locatie Hilversum, Hilversum, the Netherlands
13 Department of Oncology, OLVG Hospital, Amsterdam, Netherlands
14 Natera, Natera Inc., San Carlos, CA, United States of America
15 Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
16 Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
17 Department of Immunology, Leiden University Medical Center, Leiden, the Netherlands
18 Department of Pathology, ZAS hospitals, Antwerp, Belgium
19 Division of Research, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia
20 Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, the Netherlands
21 Department of Hematology, Leiden University Medical Center, Leiden, the Netherlands

Abstract

Immune checkpoint inhibition (ICI) with chemotherapy is now standard of care for stage II-III triple negative breast cancer (TNBC). However, it is largely unknown for which patients ICI without chemotherapy could be an option and what the benefit of combination ICI could be. The adaptive BELLINI trial explored whether short combination ICI induces immune activation (primary endpoint: two-fold increase in CD8+ T cells or IFNG), providing rationale for neoadjuvant ICI without chemotherapy. In window of opportunity cohorts A (4 weeks anti-PD1) and B (4 weeks anti-PD1 + anti-CTLA4), we observed immune activation in 53% (8/15) and 60% (9/15) of patients, respectively. High tumor-infiltrating lymphocytes (TILs) correlated with response. Single-cell RNA sequencing revealed that higher pretreatment tumor-reactive CD8+ T cells, follicular helper T cells and shorter distances between tumor and CD8+ T cells correlated with response. Higher levels of regulatory T cells post-treatment associated with non-response. Based on these data, we opened cohort C for patients with high TILs (≥50%) who received 6 weeks neoadjuvant anti-PD1 + anti-CTLA4 followed by surgery (primary endpoint: pathological complete response, pCR). 53% (8/15) of patients had major pathological response (< 10% viable tumor) at resection, with 33% (5/15) having pCR. All cohorts met Simon's two-stage threshold for expansion to stage II. We observed grade ≥3 adverse events for 17% of patients, and a high rate (57%) of immune-mediated endocrinopathies. In conclusion, neoadjuvant immunotherapy without chemotherapy demonstrates potential efficacy and warrants further investigation in patients with early TNBC. ClinicalTrials.gov Identifier: NCT03815890.

Introduction

The addition of programmed death (ligand) 1 blockade (anti-PD(L)1) to neoadjuvant chemotherapy has changed the treatment landscape for patients with early (stage II-III) triple negative breast cancer (TNBC)1. However, all trials evaluating the efficacy of anti-PD(L)1 in TNBC combined it with chemotherapy[1–4]. This chemotherapy backbone inevitably results in a high rate of adverse events, significantly affects quality of life and could diminish T cell activity[5,6].

So far, no biomarkers have been established to predict which patients with early stage TNBC will benefit from anti-PD1. Therapy is currently given for a total duration of one year, while data in other tumor types have shown that a pCR can be reached after only a few weeks of treatment with immune checkpoint inhibitors (ICI)[7–11]. Overtreatment prevention is an increasingly important consideration due to the high number of patients needed to treat to prevent one recurrence and increasing toxicity with more intense and longer treatments. Therefore, there is an urgent clinical need to optimize treatment schedules and improve patient selection for specific treatments[12].

While numerous studies have integrated anti-PD(L)1 therapy with chemotherapy in early stage TNBC[1–3,13], data on combination ICIs are limited. ICIs targeting CTLA4 have revolutionized treatment for non-small cell lung cancer (NSCLC)8 and melanoma[14–16]. Additionally, neoadjuvant trials across various tumor types have shown impressive major pathological response rates when combining anti-PD(L)1 with low-dose anti-CTLA4[7,8,10,17]. A trial in metastatic breast cancer revealed long-lasting responses after combining low-dose anti-CTLA4 with anti-PD1[18], which are infrequently observed with anti-PD(L)1 alone. These findings provide a rationale to test low-dose anti-CTLA4 in combination with anti-PD(L)1 in early TNBC.

Simultaneously with the advent of ICI, tumor-infiltrating lymphocytes (TILs) have emerged as a putative prognostic and predictive biomarker[19–22]. TNBC patients with high TIL levels have an excellent prognosis even without chemotherapy[19,23], suggesting that TILs reflect an endogenous antitumor T cell response. Moreover, in metastatic TNBC, high TIL levels are associated with response to ICI[24,25]. Collectively, these findings imply that TILs may serve as a tool for identifying TNBC patients who are more likely to benefit from ICI and have a favorable prognosis, paving the way for exploring chemotherapy de-escalation. The BELLINI trial is an adaptive platform trial exploring the effect of ICI without chemotherapy starting with window of opportunity cohorts with a biological endpoint followed by neoadjuvant cohorts with complete pathological response (pCR) endpoint. This adaptive platform trial consists of sequential, single-cohort, phase 2 studies, where new cohorts can be opened based on signals obtained in prior cohorts. The first two cohorts evaluated whether four weeks of nivolumab (anti-PD1, cohort A) or nivolumab and low-dose ipilimumab (anti-PD1 and anti-CTLA4, cohort B) can lead to immune activation (primary endpoint). This four-week therapy regimen was scheduled before the start of regular therapy, and therefore the effect of ICI could be assessed independently of chemotherapy. Promising results in cohorts A and B among patients with high TILs (≥50%) led to the initiation of cohort C. In cohort C, we used a neoadjuvant design with six weeks of nivolumab plus low-dose ipilimumab followed by surgery to assess the pCR rate[14,26].

This is the first trial combining anti-PD1 with anti-CTLA4 in early breast cancer, as well as the first trial exploring what pCR rate could be achieved with ICI-only approaches and using TIL levels as an entry criterion to enrich for inflamed tumors.

Methods

Patients

Patients in cohorts A and B were eligible for enrollment if they were at least 18 years of age and had stage I-III (clinical tumor stage T1c-3, nodal stage N0-3, according to the primary tumor regional lymph node staging criteria of the American Joint Committee on Cancer, 7th

edition) triple negative breast cancer with confirmation of estrogen receptor and HER2 negativity (ER < 10% and HER2 0, 1 or 2 in the absence of amplification as determined by in situ hybridization) on a biopsy from the primary tumor in the breast; newly diagnosed, previously untreated disease; a WHO performance status score[63] of 0 or 1 and adequate organ functions. The TILs percentage needed to be 5% or more. To ensure balanced enrollment based on TIL levels, each cohort included 5 patients with low (5-10%), 5 patients with intermediate (11-49%), and 5 patients with high (≥50%) TIL levels. Patients with concurrent ipsilateral, bilateral, or multifocal primary tumors were also eligible for enrollment. For cohort C, patients had to meet the same criteria, but the nodal stage had to be N0, tumor stage T1c-T2, and TILs had to be 50% or more. The intention for cohort C was to explore the potential feasibility of chemotherapy de-escalation in patients with high TILs. Since withholding adjuvant capecitabine for high-risk patients and/or escalating locoregional treatment for patients with more extensive disease was undesired, cohort C included only LN-negative patients.

Exclusion criteria included history of immunodeficiency, autoimmune disease or conditions requiring immunosuppression (>10 mg daily prednisone or equivalent); other immunosuppressive medications intake within 28 days of study drug administration; chronic or recurring infections; occult breast cancer; fertility preservation due to breast cancer diagnosis; active hepatitis B virus or hepatitis C virus infection; clinically significant cardiovascular disease; previous systemic anti-cancer treatment.

Trial design and treatments

The BELLINI trial (full title: Pre-operative Trial for Breast Cancer With Nivolumab in Combination With Novel IO; NCT03815890) is a single center, non-blinded, non-randomized, non-comparative phase II study designed to evaluate the feasibility and efficacy of checkpoint inhibition before regular neoadjuvant therapy or surgery in patients with primary breast cancer. Cohorts for prespecified breast cancer subgroups are opened in a sequential manner. Here we report the first three TNBC cohorts for patients who were treated with nivolumab (cohort A) or nivolumab + ipilimumab for four (cohort B) or six (cohort C) weeks. A: Nivolumab monotherapy, 240mg on D1 and D15. B: Nivolumab+ipilimumab 1 mg/kg on D1 and nivolumab 240mg on D15. C: Nivolumab+ ipilimumab 1 mg/kg on D1 and D21. Regular therapy, consisting of neoadjuvant chemotherapy or primary surgery, started on D29 and onwards. Given the poor prognosis of patients with low TIL levels and the hypothesis that these women will probably not be the super-responders to ICI, patients were only eligible with TILs≥5%. A threshold of 5% TILs was selected to exclude true immune-deserted tumors. Equal distribution of patients with different levels of tumor of infiltrating lymphocytes over the cohorts was ensured by inclusion of 5 patients with TILs-low (5-10%), 5 patients with TILs-intermediate (11-49%), and 5 patients with TILs-high (≥50%) scores per cohort.

After cohorts A (in the protocol defined as cohort 1B) and B (in the protocol defined as cohort 2B) the protocol was amended to open cohort C (in the protocol defined as cohort 3B). Cohort C had the same inclusion criteria as cohort A and B, except that only inclusion of patients with clinically node-negative disease and with TIL levels of 50% or higher was allowed. With the amendment to open cohort C, the window of opportunity design was changed into a true neoadjuvant design with all patients going to surgery after the immunotherapy. After completing the interim analysis of cohorts A and B an amendment was approved to use pathological complete response (pCR) as primary endpoint instead of immune activation for cohort C and subsequent cohorts (see details on Endpoints below).

Ethics statement

All patients provided written informed consent before enrollment. This investigator-initiated trial was designed by the Netherlands Cancer Institute (NKI).The trial was conducted in accordance with the protocol, Good Clinical Practice standards and the Declaration of Helsinki. The full protocol, amendments, and the informed consent form were approved by the medical ethical committee of the Netherlands Cancer Institute (NKI, Amsterdam).

Endpoints

Cohorts A and B:

The primary endpoint for cohorts A and B is immune activation following two cycles of neoadjuvant ICI, defined as a 2-fold increase in CD8+ T cells assessed via immunohistochemistry and/or an increase in *IFNG* gene expression. High-quality paired biopsies are necessary for the evaluability of this primary endpoint.

As a secondary endpoint for cohorts A and B, we evaluated the clinical response.

Clinical response was defined as:

Radiological signs of response: At least a 30% decrease on MRI (partial response (PR) according to RECIST 1.1, not confirmed). The target (or index) lesion is defined as the largest enhancing lesion. In case of multifocality or multicentricity the largest mass and/or non-mass enhancement was measured in the axial/sagittal or coronal plane and defined as target/index lesion. In these cases, the total area occupied by the tumor (including all masses and non-mass enhancement) was also measured. The total tumor area was used for the RECIST measurements.

AND/OR

Pathological signs of response: Pathological response could be studied in biopsies from 28 patients due to the window of opportunity design. Absence of viable tumor after four weeks of therapy in the post-treatment biopsy was classified as a clinical response. For patients proceeding to surgery this was defined as partial or complete pathological response, according to the European Society of Mastology (EUSOMA criteria).

Cohort C:

The primary endpoint for cohort C is pathological complete response (pCR), defined as no viable tumor remaining in the breast and lymph nodes (ypT0N0)[64]. Major pathologic response (MPR, secondary endpoint) is a frequently used surrogate endpoint for efficacy in neoadjuvant trials evaluating immune checkpoint blockade across cancer types[8,11,26]. MPR was defined as ≤10% of residual viable tumor in the surgical specimen[17,65,66] or no viable tumor in the breast but residual tumor cells in the lymph nodes.

All Cohorts (A, B, C):

Secondary endpoints included feasibility, safety, and radiological response. Feasibility was determined based on any treatment-related complications that led to a delay in chemotherapy or primary surgery beyond six weeks from the start of therapy. All patients were closely monitored for adverse events (AEs) for 100 days after the administration of the last study treatment, following the Common Terminology Criteria for Adverse Events (CTCAE) v.5[67]. In addition, we reported all immune-related adverse events in the first year of follow-up. Radiological response was assessed according to the RECIST 1.1 guidelines, but not confirmed.

Statistical analysis

For this exploratory, hypothesis-generating study, no formal sample size calculation was performed for efficacy because there was no data on the efficacy of neoadjuvant immunotherapy in breast cancer at the time of the design of this study. For cohorts A and B, the null hypothesis of a true immune activation in ≤30% of patients was tested against a one-sided alternative. For cohort C, design was identical with the exception of null hypothesis being pCR in ≤30% of patients tested against a one-sided alternative. For 80% power, at a one-sided significance level of 0.05, 15 patients were accrued per cohort to be evaluated in the first stage. If there were 5 or less responses among these 15 patients, the cohort was closed for futility. Otherwise, the cohort could be expanded with 31 additional patients, reaching a total of 46. We decided to publish after stage I, which was allowed by protocol, due to the observation that very early responses to ICI without chemotherapy are possible in TNBC, which warrants efforts to de-escalate therapy for a subset of patients, in contrast to the current therapy escalation for all TNBC patients. Median follow-up time was obtained using the reverse Kaplan-Meier method. Analyses were performed using R[68] v.4.2.1.

Pathology assessments and IHC analyses

All patients underwent baseline tumor staging, consisting of ultrasound of the breast, axilla and periclavicular region and MRI imaging of the breast. PET-CT imaging was performed in all participants to confirm the clinical stage. Pretreatment tumor histological biopsies (4

core biopsies, 14G needle) were taken for all patients, and post-treatment tissue was either obtained through a biopsy (3 core biopsies, 14G needle) for patients continuing neoadjuvant chemotherapy (n=28) and the surgical specimen was used for those undergoing surgery right after the ICI study treatment (n=3). Histopathological examination of biopsies and resection specimens was carried out by five experienced breast cancer pathologists (HMH, RS, KvdV, JvdB, NK). Resected tumors were examined in their entirety and regression of resected tumors was assessed by estimating the percentage of residual viable tumor of the macroscopically identifiable tumor bed, as identified on routine hematoxylin and eosin (H&E) staining. Formalin-fixed paraffin-embedded tissue sections were used for H&E stainings and for immunohistochemical analysis of CD8 (C8/144B, DAKO), PDL1 (22C3, DAKO) and PD1 (NAT105, Roche Diagnostics). The percentage of tumor cells and TILs was assessed by pathologists trained for TILs assessment on H&E-stained slides according to the international standard from the International Immuno-Oncology Biomarker Working Group[22] (see www.tilsinbreastcancer.org for all guidelines on TILs assessment in solid tumors). After a pathologist provided an initial TILs score, an 'expert TILs score' was generated as a consensus score from at least 2 out of 4 trained pathologists using slidescore.com for online scoring[69]. TILs scores for inclusion were scored on the diagnostic biopsy of the patient to allow for stratification of patients (low ≥5-10%, intermediate 11-49%, high ≥50%).

Immunohistochemistry

Immunohistochemistry of the FFPE tumor samples was performed on a BenchMark Ultra autostainer (Ventana Medical Systems). The double stain was performed on a Discovery Ultra autostainer. Briefly, paraffin sections were cut at 3 um, heated at 75°C for 28 minutes and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1, Ventana Medical Systems) for 48 minutes at 95°C (PDL1) or 64 minutes at 95°C. (PD1/CD8 double). PDL1 was detected using clone 22C3 (1/40 dilution, 1 hour at RT, Agilent/DAKO, Lot11654144). Bound antibody was detected using the OptiView DAB Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems).

For the double staining PD1 (Yellow) followed by CD8 (Purple) the PD1 was detected in the first sequence using clone NAT5 (Ready-to-Use, 32 minutes at 370C, Roche Diagnostics, Lot11654144). The PD1-bound antibody was visualized using Anti-Mouse NP (Ventana Medical systems, Ready to Use dispenser, LotK09956) for 12 minutes at 37C followed by Anti-NP AP (Ventana Medical systems, Ready to Use dispenser, LotJ23971) for 12 minutes at 37C, followed by the Discovery Yellow detection kit (Ventana Medical Systems). In the second sequence of the double staining procedure, CD8 was detected using clone C8/144B (1/200 dilution, 32 minutes at 37C, Agilent, Lot41527763). CD8 was visualized using Anti-Mouse HQ (Ventana Medical systems, Ready to Use dispenser, LotK20711) for 12 minutes

at 370C followed by Anti-HQ HRP (Ventana Medical systems, Ready to Use dispenser, LotK22062) for 12 minutes at 37C, followed by the Discovery Purple Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). A PANNORAMIC® 1000 scanner from 3DHISTECH was used to scan the slides at a 40x magnification.

Distance analysis between tumor and CD8+ T cells

Spatial analysis was performed on the pretreatment biopsies of all included patients. The stained slides were scanned, and image analysis was performed with the HALO image analysis software from Indica Labs, v3.4.2986.185 (cohorts A and B) and v.3.6.4134 (cohort C). Within HALO, the multiplex IHC module was used to phenotype and quantify CD8-positive cells. Cell segmentation was performed by the detection of hematoxylin (detection weight = 1) and PD1 (detection weights 0.045 for cohorts A&B; 0.5 for cohort C) and CD8 for cohort C (detection weight = 0.5) staining, utilizing a nuclear segmentation aggressiveness of 0.045. Minimal intensity thresholds to consider a cell positive for a marker were set for hematoxylin (0), PD1 (0.25 for cohorts A&B, 0.1 for cohort C), and CD8 (0.1) separately. Biopsies were analyzed in total, while for resection specimens the analysis was restricted to representative tumor beds as annotated by a breast cancer pathologist. The quantified levels of CD8+ and PD1+CD8+ cells were corrected for the analyzed tissue area (cells /μm^2).

Artificial intelligence tumor classifiers (Object Phenotyper, HALO AI) were developed to discriminate between tumor and non-tumor cells in cohorts A&B and in cohort C. Individual cells were segmented (nuclei seg BF v.1.0.0), and the classifiers were trained by annotating single cells as tumor or non-tumor. The annotations were guided by marked tumor regions on H&E-stained slides by a trained BC pathologist. The classifiers were finalized with 20.000 iterations and a cross-entropy of 0.009 (cohort A&B) and >10.000 iterations and cross-entropy of 0.021 (cohort C).

Merging the results of the multiplex IHC and tumor classifier enabled the visualization of the spatial distribution of tumor and CD8+ cells (**ED Fig.1B-F**). Using the nearest neighborhood analysis, the average distance between the tumor and immune cells was quantified by taking the mean of the distances between every tumor cell and its nearest cell of the above-mentioned immune phenotypes in the pretreatment biopsies (**ED Fig.1F**). Distances from tumor cells to the nearest CD8+ T cells were taken as a measure of proximity of CD8+ T cells to the tumor.

DNA and RNA isolation

DNA and RNA were extracted from fresh-frozen, pre- and post-treatment tumor material using the AllPrep DNA/RNA Kit (QIAGEN) for frozen material, following the manufacturer's protocol, in a QIAcube (QIAGEN). Germline DNA was isolated from patient peripheral blood mononuclear cells using the DNeasy Blood & Tissue Kit (QIAGEN).

Bulk RNA sequencing

Total RNA Quality Control

Quality and quantity of the total RNA was assessed by the 2100 Bioanalyzer using a Nano chip (Agilent, Santa Clara, CA). Total RNA samples having RIN>8 were subjected to library generation.

TruSeq Stranded mRNA library generation

Strand-specific libraries were generated using the TruSeq Stranded mRNA sample preparation kit (Illumina Inc., San Diego, RS-122-2101/2) according to the manufacturer's instructions (Illumina, Document # 1000000040498 v00). Briefly, polyadenylated RNA from intact total RNA was purified using oligo-dT beads. Following purification, the RNA was fragmented, random primed and reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen, part # 18064-014) with the addition of Actinomycin D. Second strand synthesis was performed using Polymerase I and RNaseH with replacement of dTTP for dUTP. The generated cDNA fragments were 3' end adenylated and ligated to IDT xGen UDI(10bp)-UMI(9bp) paired-end sequencing adapters (Integrated DNA Technologies, Inc., Coralville) and subsequently amplified by 12 cycles of PCR. The libraries were analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent, Santa Clara, CA), diluted and pooled equimolar into a multiplex sequencing pool.

Sequencing

The libraries were sequenced with 54 paired-end reads on a NovaSeq6000 using a S1 Reagent Kit v1.5 (100cycles) (Illumina Inc., San Diego).

Data analysis

RNA sequencing data were aligned to GRCh38 with STAR[70] 2.7.1a, with the twopassMode='Basic'. FPKM were obtained with RSeQC[71] 4.0.0 FPKM_count.py and subsequently normalized to transcripts per million. Data quality was assessed with FastQC[72] 0.11.5, FastQ Screen[73] 0.14.0, the Picard CollectRnaSeqMetrics[74,75] and RSeQC[71] 4.0.0 read_distribution.py and read_duplication.py and were found to be suitable for the downstream analysis. TNBCtype[76] was used for the Lehmann subtype classification[77]. The Gseapy[78] 1.0.3 ssgsea tool with the sample_norm_method='rank' was used for gene set signature scoring. For the signature analysis, p-values were significant after FDR correction (Benjamini–Hochberg) at 10% significance level. Data were analyzed with Python[79] 3.10.5. Pandas[80,81] 2.0.0 and numpy[82] 1.22.4 were used for data handling. Matplotlib[74] 3.5.2, seaborn[83] 0.12.2 and statannotations[84] 0.5.0 were used for plotting.

Whole exome sequencing

For each sample the amount of double stranded DNA was quantified by using the Qubit® dsDNA HS Assay Kit (Invitrogen, cat no Q32851). A maximum amount of 2 µg of double stranded genomic DNA was fragmented by covaris AFA technology to obtain fragment sizes of 200-300 bp. Samples were purified using Agencourt AMPure XP Reagent (Beckman Coulter, cat no A63881) in a 2x reaction volume settings according to manufacturer's instructions. The fragmented DNA was quantified and qualified on a BioAnalyzer system using the DNA7500 assay kit (Agilent Technologies cat no. 5067- 1506). With a maximum input amount of 1 µg fragmented DNA, NGS library preparation for Illumina sequencing was performed using the KAPA HTP Prep Kit (KAPA Biosystems, KK8234) in combination with xGen UDI-UMI Adapters of IDT (Integrated DNA Technologies). During the library amplification step, 4 cycles of PCR were performed to obtain enough yield for the exome enrichment assay. All DNA libraries were quantified on a BioAnalyzer system using the DNA7500 assay kit. Exome enrichment was performed on library pools of 6 unique dual indexed libraries, 500 ng each, using the xGen™ Exome Hyb Panel v2 (IDT, cat no 10005152) and xGen™ Hybridization Capture Core Reagents according to manufacturer's protocol, with hybridization time adjusted to 16 hours and 10 cycles of PCR performed during post-capture PCR. All exome enriched library pools were quantified on a BioAnalyzer system using the DNA7500 assay kit, pooled equimolar to a final concentration of 10nM and subjected to paired-end 100 bp sequencing on an Illumina Novaseq 6000 instrument using a NovaSeq 6000 S4 Reagent Kit v1.5 (Illumina, 20028313), according to manufacturer's instructions.

Data analysis

Sequencing reads were aligned to the human reference GRCh38 (Ensemble, v. 105) using BWA[85] 0.7.17. Duplicated reads were marked using Picard[75] MarkDuplicates 2.25.0, after which quality scores were recalibrated using GATK4[86] BaseRecalibrator 4.2.2.0. Single-nucleotide variants (SNVs) and short insertions and deletions (indels), were called using GATK4[86] Mutect2 4.2.2.0 on the tumor samples matched with germline samples. Subsequently, variants were filtered by the PASS filter, and annotated using Ensembl Variant Effect Predictor 105. maftools[87] 2.10.5 package was used for the analysis. Tumor mutational burden (TMB) was calculated by summarizing the total number of non-synonymous somatic mutations with a minimal variant allele frequency of 20%. Data were analyzed with Python[79] 3.10.5 and R[88] 4.1.3. Pandas[80,81] 2.0.0 was used for data handling. maftools[87] 2.10.5, Matplotlib[74] 3.5.2, seaborn[83] 0.12.2, and statannotations[84] 0.5.0 were used for plotting.

Single cell RNA sequencing and TCR sequencing

Preparation of the single cell suspension

Following biopsy or obtaining resection specimens, samples were rapidly processed for single-cell RNA sequencing (scRNA-seq). Samples from cohort A were minced on ice and frozen in

10% DMSO FCS in the -80˚C degrees. Within 4 weeks after freezing, samples were defrosted in 37˚C degrees medium. Samples from cohort B were minced on ice and immediately processed for single cell sequencing (not frozen), which did not result in a batch effect.

Samples were transferred to a tube containing 1mL digestion medium containing collagenase P (2 mg/ml, ThermoFisher Scientific) and DNAse 1 (10 U/µL, Sigma) in RPMI (ThermoFisher Scientific). Samples were incubated for 20 min at 37˚C degrees and were pipetted up and down every 5 minutes for 30 seconds. Next, samples were filtered on a 40 micron nylon mesh (ThermoFisher Scientific) and directly after the same volume of ice cold PBS containing 0.04% BSA was added. Following centrifugation at 300 g and 4˚C degrees for 5 min, the supernatant was removed and discarded, and the cell pellet was resuspended in red cell blood lysis buffer for 5 minutes at room temperature and then centrifuged again at 300 g at 4 ˚C degrees for 5 min. The supernatant was removed and discarded and the pellet was resuspended in PBS containing 0.04% BSA. Next, 10 µl of this cell suspension was counted using an automated cell counter (ChemoMetec NucleoCounter NC-200) to determine the concentration of live cells. The entire procedure was usually completed within 1h and 15 minutes.

Single cell RNA-seq data acquisition and preprocessing

Libraries for scRNA-seq were generated using the Chromium Single Cell 5' library and Gel Bead & Multiplex Kit from 10x Genomics. We aimed to profile 10 000 cells per library if a sufficient number of cells was retained during dissociation. All libraries were sequenced on a HiSeq4000 or NovaSeq6000 until sufficient saturation was reached.

Data analysis

After quality control, raw sequencing reads were aligned to the human reference genome GRCh38 and processed to a matrix representing the UMI's per cell barcode per gene using CellRanger (10x Genomics, v2.0). The data were analyzed with scanpy[89] 1.9.3 and Seurat[90] v3. Cellbender[91] 0.3.0 was used for eliminating technical artifacts, and cells above the quality cutoff of 0.5 were filtered out. Cells with mitochondrial RNA content >0.25, the number of genes <200 or >6000 and <400 counts were filtered out. After normalization, regression for the number of UMIs, percentage mt-RNA, sample ID, cell cycle, hypoxia, interferon content and cell stress was performed on the 2000 most variable genes followed by a principal component analysis. Next a UMAP was generated and clustering was performed at resolution 0.2 using the 30 most informative components. Major cell types were identified based on canonical marker genes.

For the T cell subclustering, the T cells were selected from the full Seurat object and the analysis described above was repeated with 10 principal components based on the elbow plot and clusters were identified at a resolution of 0.6 and were annotated based on breast

cancer tissue-specific marker genes[92]. Cells expressing markers of other cell types (immunoglobulins, hemoglobin) were filtered out. PCA was calculated on highly variable genes with k=30. Clustering was performed with Phenograph[93] with k=30. Cluster identification was performed based on canonical marker genes. Signature scores were calculated with sc.tl.score_genes. Groups were compared with sc.tl.rank_genes_groups, with method='wilcoxon' and use_raw=True. EnrichR[94,95] was used for the pathway enrichment analysis. Activated Tregs were defined based on the level of *CD137* gene expression >0.5 in the Treg cell population. PD1+Ki67+CD4+ cells were defined based on the level of *MKI67* gene expression >0 in the Tfh cell population. Scirpy[96] 0.11.2 was used for the TCR analysis. Clonotypes were defined based on the amino acid structure. Clonality was calculated as (1 - normalized Shannon entropy). Data were analyzed with Python[79] 3.10.5. Pandas[80,81] 2.0.0 and numpy[82] 1.22.4 were used for data handling. Matplotlib[74] 3.5.2, seaborn[83] 0.12.2, sc-toolbox[97] 0.12.3 and statannotations[84] 0.5.0 were used for plotting.

ctDNA analysis

A proprietary bioinformatics tissue variant calling pipeline was used to select a set of 16 high-ranked, patient-specific, somatic, clonal single nucleotide variants (SNVs) from WES. The Signatera amplicon design pipeline was used to generate mPCR primer pairs for the given set of 16 variants. For cfDNA library preparation, up to 20 000 genome equivalents of cfDNA from each plasma sample were used. The cfDNA was end-repaired, A-tailed, and ligated with custom adapters, followed by amplification (20 cycles) and purified using Ampure XP beads (Agencourt/Beckman Coulter). A proprietary multiplex PCR (mPCR) methodology was used to run patient-specific assays. Sequencing was performed on these mPCR products on an Illumina HiSeq 2500 Rapid Run (50 cycles) using the Illumina Paired End v2 kit with an average read depth of >100000X per amplicon. All paired-end reads were merged using Pear 0.9.8 software and mapped to the hg19 reference genome with Novoalign version 2.3.4 (http://www.novocraft.com/). Plasma samples with at least 2 variants with a confidence score above a predefined algorithm threshold were defined as ctDNA-positive.

Flow cytometry of fresh blood

The flow cytometry was performed as previously described[98]. In short, fresh blood samples were processed and analyzed within 24 hours after blood draw. Peripheral blood was collected in EDTA vacutainers (BD) and subjected to red blood cell lysis (lysis buffer: dH2O, NH4Cl, NaHCCO3, EDTA). Cells were suspended in PBS containing 0.5% BSA and 2mM EDTA and counted using the NucleoCounter NC-200 (Chemometec) automated cell counter. To obtain absolute white blood cell (WBC) counts per mL of human blood, the total amount of post-lysis cells was divided by the volume (mL) of blood obtained from the patient. For surface antigen staining, cells were first incubated with human FcR Blocking Reagent (1:100

Miltenyi) for 15 min at 4°C and then incubated with fluorochrome-conjugated antibodies for 30 min at 4°C. For intracellular antigen staining, cells were fixed with Fixation/Permeabilization solution 1X (Foxp3/Transcription Factor Staining Buffer Set, eBioscience) for 30 min at 4°C and stained with fluorochrome-conjugated antibodies in Permeabilization buffer 1X (eBioscience) for 30 min at room temperature. Viability was assessed by staining with either 7AAD staining solution (1:10; eBioscience) or Zombie Red Fixable Viability Kit (1:800, BioLegend). Data acquisition was performed on an LSRII SORP flow cytometer (BD Biosciences) using Diva software and data analysis was performed using FlowJo 10.6.2. Gating strategy is displayed in **ED Fig.5A**.

Data availability

DNA and RNA-seq data are stored in the European Genome–Phenome Archive EGAS50000000567 (RNA-Seq) and EGAS50000000568 (WES)). Sequencing data and source data supporting the findings of this study will be made available from the corresponding author (m.kok@nki.nl) for academic use, within the limitations of the provided informed consent. Data will not be made available for commercial use. A first response to the request will be sent in <4 weeks. Data requests will be reviewed by the corresponding author and Institutional Review

Board of the NKI and after approval, applying researchers will have to sign a data transfer agreement with the NKI.

Code availability

No custom developed code was used for the analysis of the study data.

Results

Design and patient characteristics

The BELLINI trial (NCT03815890; Fig.1A,G, ED Fig.1A) is a pre-operative, window of opportunity (WOO), phase II, multiple-cohort non-randomized study in early (stage I-III) breast cancer utilizing an adaptive Simon's two-stage design[27]. Here, we report the initial results from the first two WOO cohorts exploring the immune-activating capacity of short-term neoadjuvant nivolumab ± ipilimumab (cohorts A and B, n=31) in patients with ≥5% TILs as well as the initial results of cohort C that was opened based on the results of cohorts A and B. The first patient was included on 19 September 2019 and the last patient on 24 January 2023.

Cohort A (n=15) received two cycles of nivolumab (240 mg) on days 1 and 15. Cohort B (n=15) received two cycles of nivolumab (240 mg) on days 1 and 15, plus one cycle of ipilimumab (1 mg/kg) on day 1. To exclude patients with a poor prognosis, less likely to

respond to ICI and not suitable for chemotherapy de-escalation, we enrolled patients with ≥5% TILs in cohorts A and B. Baseline characteristics were similar between cohorts A and B, except for a higher proportion of patients with positive lymph nodes in cohort B (Table 1).

The primary endpoint for cohorts A and B was immune activation, defined as at least a two-fold increase in CD8+ T cells (measured by immunohistochemistry (IHC), ED Fig. 1B-F) and/or increased interferon gamma (IFNG) gene expression. This endpoint was based on the observation that significant increases in intratumoral CD8+ T cells[25,28] and higher IFNG signature scores[17,29] in serially biopsied tumors are correlated with responses to anti-PD(L)1.

Clinical response (secondary endpoint) in cohorts A and B was defined as PR/CR on MRI (RECIST1.1) or no viable tumor in post-treatment biopsy for patients proceeding to neoadjuvant chemotherapy. For patients directly proceeding to surgery this was defined as partial or complete pathological response (EUSOMA). Other secondary endpoints included safety and translational analyses. MRI scans and biopsies were collected at baseline and after two ICI cycles.

Efficacy of short-term nivolumab and nivolumab+ipilimumab in early TNBC (window of opportunity)

Immune activation was achieved in 8 tumors (53.3%) in the nivolumab cohort (A) and 9 (60%) in the nivo-ipi cohort (B) (Fig.1B). Therefore, both cohorts met the Simon's two-stage27 threshold for expansion to stage II. After four weeks, patients proceeded to standard neoadjuvant chemotherapy followed by surgery (n=28) or surgery without neoadjuvant chemotherapy (n=3). Clinical response was observed in 12/31 patients (38.7%, 95% CI 23.7%-56.2%) with 7/31 patients (22.6%, 95% CI 11.4%-39.8%) having a partial response (PR) according to RECIST 1.130 (Fig.1C,D). 10/31 patients had no viable tumor in the biopsy and in the three patients who underwent surgery directly after ICI, two partial and one complete pathological response was seen. Despite these clear pathological responses, MRI showed modest downsizing, indicating MRI underestimates early ICI response (ED Fig. 1H), consistent with findings in early-stage melanoma31, colorectal and gastroesophageal cancers17,32. Strikingly, clinical response was only observed for patients with TILs ≥30% (Fig.1E) and a CPS PDL1 ≥20% (Fig.1F). Patients with lower pretreatment CD8+ T cell levels were more likely to achieve immune activation (ED Fig.1G), likely due to either less possibility for value doubling or to a very early immune response in highly inflamed tumors.

Short-term neoadjuvant nivolumab + ipilimumab can induce pathological responses in patients with high TILs

Both cohorts A and B met the predefined thresholds of the Simon's two-stage design[27], allowing expansion to stage II. However, given the promising clinical responses observed in cohorts A and B and the approval of neoadjuvant pembrolizumab plus chemotherapy[4],

Table 1. *Baseline patient characteristics*

Characteristic	A: Nivo *(n=16)*	B: Nivo-ipi 4 wks *(n=15)*	C: Nivo-ipi 6 wks *(n=15)*
Median age, years (IQR range)	48 (39.8-53.2)	50 (42.5-57.5)	51 (36.0-56.5))
WHO PS[a], *n (%)*			
0	16 (100)	14 (93.3)	15 (100)
1	0 (0.0)	1 (6.7)	(0.0)
Histological subtype, *n (%)*			
NST	16 (100)	13 (86.7)	14 (93.3)
Metaplastic	0 (0.0)	1 (6.7)	0 (0.0)
Lobular pleiomorphic	0 (0.0)	1 (6.7)	1 (6.7)
Tumor stage, *n (%)*			
T1	5 (31.3)	5 (33.3%)	2 (13.3)
T2	10 (62.5)	9 (60.0)	13 (86.7)
T3	1 (6.2)	1 (6.7)	0 (0.0)
Nodal status, *n (%)*			
N0	13 (81.3)	5 (33.3)	15(100)[d]
N1	2 (12.5)	9 (60.0)	0 (0.0)
N3	1 (6.3)	1 (6.7)	0 (0.0)
Tumor grade[b], *n (%)*			
2	1 (6.3)	4 (26.7)	0 (0.0)
3	15 (93.8)	11 (73.3)	15 (100)
Germline BRCA1/2 mutation, *n (%)*			
Yes	3 (18.8)	3 (20.0)	4 (26.7)
No	12 (75.0)	10 (66.7)	11 (0.0)
Unknown	1 (6.3)	2 (13.3)	0 (0.0)
TILs[c], *(%)*			
Median (IQR)	40.8 (6.2, 60.3)	37.5 (23.8, 61.4)	52.5 (45.3, 73.8)

[a]WHO performance status. [b]Tumor grade according to Bloom Richardson. [c]TILs were averaged between the diagnostic TILs score and the study pretreatment TILs score. sTILs were scored according to international guidelines[22] as a numerical variable. All samples were evaluated by at least two breast cancer pathologists and their score for each sample was averaged. [d]Cohort C only allowed inclusion of N0 patients. Abbreviations: NST, no special type; TILs, stromal tumor-infiltrating lymphocytes.

the study team decided not to proceed to stage II with the WOO design but to open cohort C with a true neoadjuvant design (n=15, Fig.1G). Since all patients with a clinical response in cohorts A and B had high TILs, cohort C was opened for patients with ≥50% TILs and allowed only patients with node-negative disease since for this patient population chemotherapy de-escalation could be an option in the future. The treatment schedule with combination ICI for cohort C was based on our data obtained in cohorts A and B as well as on the well-established effective and tolerable combination ICI schedule in melanoma[14,26].

Patients in cohort C underwent a 6-week treatment regimen of nivolumab and ipilimumab (administered on days 1 and 21), followed by surgery (Fig.1G). Five patients had a pCR, (33.3%, 95% CI 15.2%-58.3%, Fig.1H) with confirmed tumor-negative lymph nodes (ypT0N0). Less than 10% viable tumor remaining was seen in 3/15 patients (20%, 95% CI 4%-48%, Fig.1H), making major pathological response rate (MPR) 8/15=53% (95% CI 27%-

Figure 1

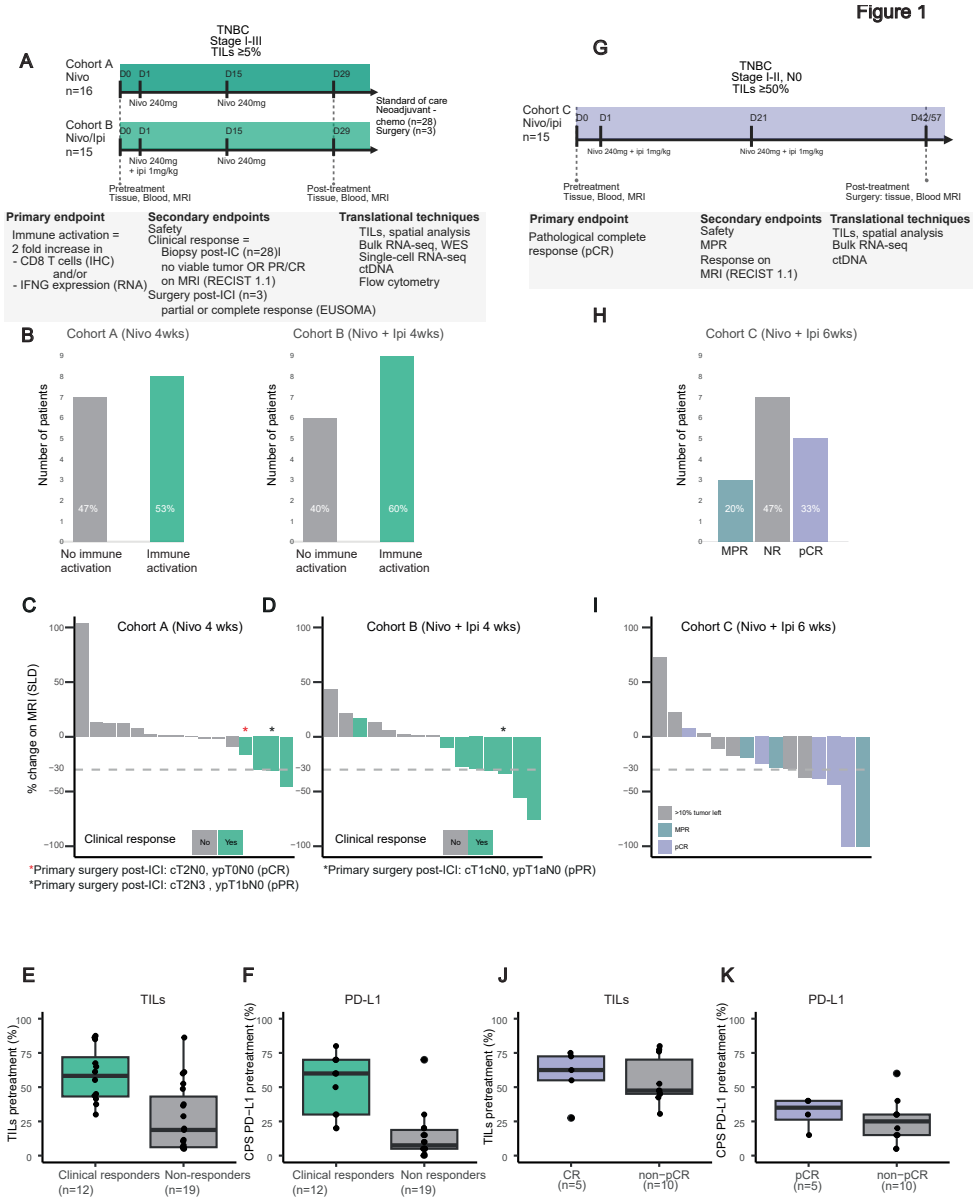


***Figure 1: BELLINI trial design, efficacy data and baseline biomarkers**. **A.** Trial design for cohorts A and B. Cohort A received 2 cycles of nivolumab (anti-PD1). Cohort B received 2 cycles of nivolumab (anti-PD1) and one cycle of ipilimumab (anti-CTLA4). Biopsies and blood were taken pretreatment and after 4 weeks of treatment after which patients proceeded to standard of care: neoadjuvant chemotherapy (n=28) or primary surgery (n=3). **B.** Numbers of patients reaching immune activation in cohorts A (n=15) and B (n=15). **C-D.** Changes in tumor size according to the MRI for cohort A (C) and cohort B (D). The gray dashed line at -30%: radiological PR. The green bars indicate clinical responses (radiological PR and/or pathological response). Asterisks (*) represent patients with resection after ICI only (n=3). pPR: pathological partial response according to EUSOMA. **E.** TILs in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n=31 patients. **F.** Combined positive PDL1 score (CPS) in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n=31 patients. **G.** BELLINI trial design for cohort C. Cohort C (n=15) received 2 cycles of nivolumab and ipilimumab on days 1 and 21. Biopsies and blood were taken pretreatment and after 6 weeks. Patients proceeded to primary surgery (n=15). **H.** pCR and MPR (<10% viable tumor left) rates in cohort C. **I.** Changes in tumor size according to the MRI in cohort C. The gray dashed line at -30%: radiological PR. Dark blue bars; pCR. **J.** TILs in pretreatment biopsies of patients according to pCR status in cohort C. n=15 patients. **K.** CPS in pretreatment biopsies for patients according to pCR status in cohort C. n=15 patients. Figures **A, G** were created with BioRender.com. In **E, J** levels of TILs calculated as average from TIL levels at diagnostic- and pretreatment study. In **E-F, J-K** boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. P-values were derived using a two-sided Mann-Whitney test.*

79%). Notably, of the 5 patients with a pCR only one had a complete radiological response (Fig.1I). Because of high TILs, N0 status and pCR which are all very favorable prognostic features, all 5 patients with a pCR were offered the option of omitting adjuvant chemotherapy and all chose not to undergo adjuvant chemotherapy (shared decision). Patients without pCR were advised adjuvant chemotherapy.

Safety data and follow-up

Toxicity data are summarized in Table 2 (all events required steroids or persisted) and detailed in ED Table 1. Neither neoadjuvant nivolumab nor nivolumab-ipilimumab resulted in previously unreported toxicities. All patients were monitored for (immune-related; IR) toxicities until one year post ICI-therapy. Treatment-related adverse events (AEs) of any grade occurred in 41/46 patients (89%). A total of 8 (17%) patients developed grade ≥3 treatment-related AEs, of which 6 were treated in cohort C. Except for the endocrinopathies all adverse events resolved. Notably, 19/46 patients (41%) developed treatment-related hypothyroidism. All patients with hypothyroidism remain dependent on replacement therapy. Six patients (13%) developed adrenal insufficiency and require ongoing corticoid replacement therapy. One patient developed a diabetic ketoacidosis and remains insulin-dependent.

All patients proceeded with tumor resection or neoadjuvant chemotherapy as scheduled. 44 patients received both ICI doses, and two patients only received one dose due to suspected immunotoxicity.

With a median follow-up duration of 32.5 months in cohorts A and B (interquartile range 28.1-40.3 months), one patient in cohort A (cT2N0; intermediate TILs) developed a second primary tumor, and one patient in cohort B (cT2N1; intermediate TILs) died from metastatic TNBC despite receiving standard of care (neo)adjuvant chemotherapy. Median follow-up for cohort C was 17.6 months (interquartile range 18.8-22.1 months). One patient (no response to ICI) refused adjuvant chemotherapy and radiotherapy and developed recurrent TNBC (pT1cNx, 80% TILs).

Table 2. *Summary of adverse events*

	A: Nivolumab (N=16)		B: Nivo+Ipi 4 wks (N=15)		C: Nivo+Ipi 6 wks (N=15)	
	Number of patients (percent)					
Immune-mediated adverse events	Any grade	Grade ≥3	Any grade	Grade ≥3	Any grade	Grade ≥3
Hypothyroidism^	6 (38%)	0 (0%)	7 (47%)	0 (0%)	6 (40%)	0 (0%)
Adrenal insufficiency*	1 (6%)	0 (0%)	2 (13%)	1 (7%)	3 (20%)	1 (7%)
Diabetes Mellitus	0 (0%)	0 (0%)	1 (7%)	1 (7%)	0 (0%)	0 (0%)
Colitis	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)
Hepatitis**	0 (0%)	0 (0%)	2 (13%)	0 (0%)	3 (20%)	3 (20%)
Polymyalgia rheumatica	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)
Pneumonitis	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (13%)	1 (7%)

This table sums all immune-mediated adverse events that required treatment with steroids or didn't resolve (endocrinopathies). A detailed list of all adverse events according to CTCAE criteria can be found in Table S1. * All patients were classified as having secondary adrenal insufficiencies and all patients remain dependent on corticosteroid replacement. ** We have included all patients requiring steroids and one patient with grade 3 IR hepatitis that did not receive steroid treatment ^ All patients are still dependent on hormone replacement therapy.

Pretreatment composition of the tumor microenvironment is associated with ICI response

Due to limited sample size, we compared clinical responders versus non-responders from both cohorts (A+B) combined and not for the cohorts separately. Clinical responders in cohorts A and B had significantly higher pretreatment TILs (p=0.0014, Fig.1E) and PDL1 scores (p=8.6e-05, Fig.1F) compared to non-responders. CD8+ T cell density was not significantly associated with clinical response (Fig.2A, ED Fig.1B-F). Spatial analysis showed that responders had significantly shorter distances from tumor cells to the nearest CD8+ T cells (p=0.00001, Fig.2B). Responders also exhibited a larger density of double-positive CD8+PD1+ cells (p=0.02, ED Fig.2A) and PD1+ cells (p=0.001, IHC, ED Fig.2B) pretreatment.

In cohort C, TILs were not different between responders and non-responders, probably due to the more homogeneous patient population with only patients with ≥50% TILs (Fig.1J). In line with this, patients with pCR had similar PDL1 scores , CD8+ T cell density (cells /µm2) and distances from tumor to nearest CD8+ T cells as patients without pCR (Fig.1K, Fig. 2D-E). We found no association between tumor mutational burden and clinical response (ED Fig.2C-D). There were no statistically significant differences between clinical responders and non-responders in TNBC subtypes33 (ED Fig. 2E).

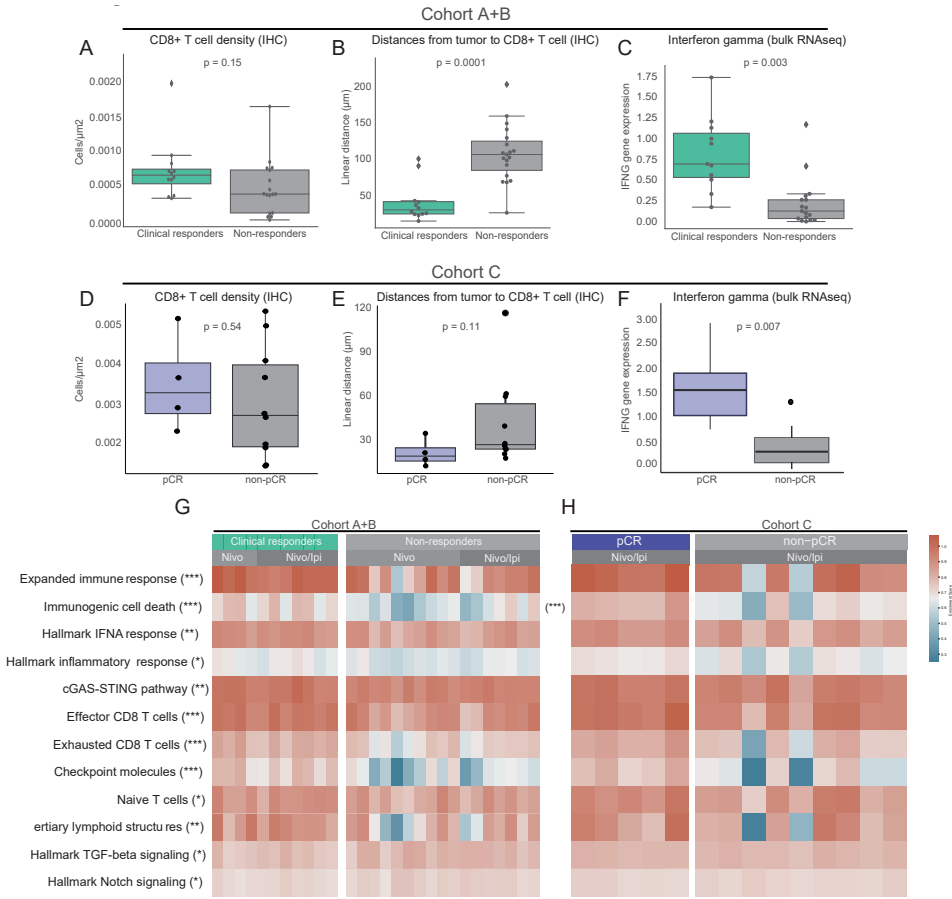
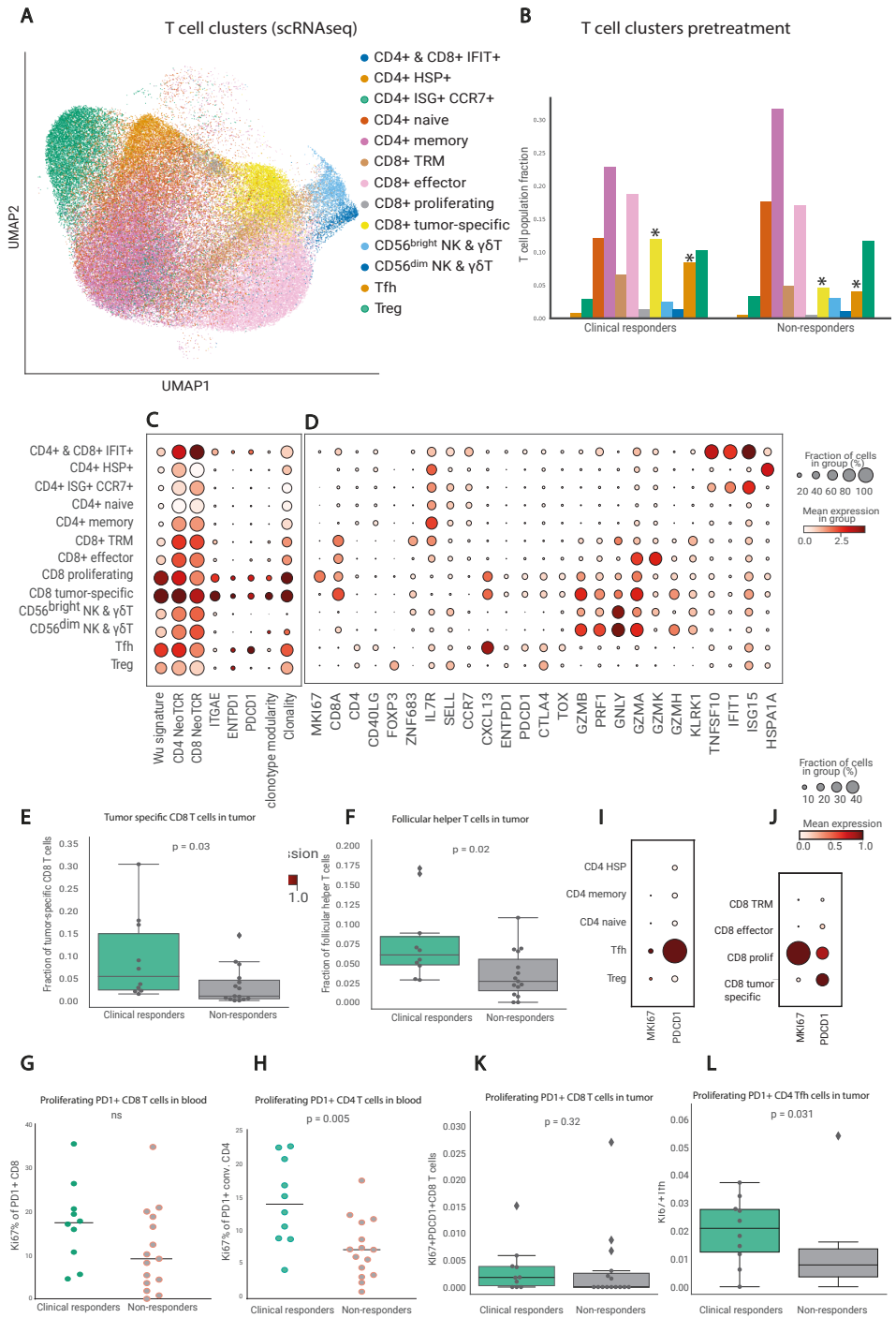


Figure 2: Pretreatment immune activation associated with clinical response. A. *CD8+ density (IHC) in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n=31 patients.* ***B.*** *Median distances (µm) from tumor cells to the nearest CD8+ T cells in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n=31 patients.* ***C.*** *IFNG gene expression scores in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n=28 patients.* ***D.*** *CD8+ density (IHC) in pretreatment biopsies of patients with and without pCR in cohort C. n=14 patients.* ***E.*** *Median distances from tumor cells to the nearest CD8+ T cells in pretreatment biopsies of patients with and without pCR in cohort C. n=14 patients.* ***F.*** *IFNG gene expression scores in pretreatment biopsies of patients with and without pCR in cohort C. n=14 patients.* ***G-H.*** *Gene set enrichment expression scores in pretreatment biopsies of patients with and without clinical response in cohorts A and B (G, n=28 patients) or pCR (H, n=14 patients) in cohort C. Heatmaps include: Expanded immune signature57, Immunogenic cell death signature58, Hallmark IFNA response gene set, Hallmark inflammatory response gene set, cGAS-STING pathway gene set59, Effector CD8+ T cell gene set60, Exhausted T cell gene set60, Checkpoint molecules gene set60, Naive T cell gene set61, Tertiary lymphoid structures gene set62, Hallmark TGF-beta signaling gene set, Hallmark Notch signaling. Asterisks represent the p-value levels: "*": p≤0.05, "**": p≤0.01, "***": p≤0.001. Reported p-values were significant after Benjamini-Hochberg (FDR) correction at 10% significance level. In* ***A-F****, boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. P-values were derived using a two-sided Mann-Whitney test.*



◀◀◀ Figure 3: Pretreatment T cell profiles of the tumor microenvironment and peripheral blood associated with clinical response in cohorts A and B. A. *UMAP representation of the T cell clusters in the single-cell RNA-Seq dataset (cohorts A and B). n=52 samples from 29 patients, 80 000 cells.* ***B.*** *Fractions of different T cell populations relative to all T cells in the pretreatment biopsies from clinical responders (left) and non-responders (right) in cohorts A and B.* ***C.*** *Dotplot illustrating markers of different T cell clusters based on single-cell RNA-Seq data (cohorts A and B).* ***D.*** *Dotplot illustrating differences in tumor reactivity markers in different T cell clusters based on single-cell RNA-Seq data (cohorts A and B). Wu_signature - CD8+ T cell tumor specificity signature34; CD4_NeoTCR - CD4+ T cell tumor specificity signature35; CD8_ NeoTCR - CD8+ T cell tumor specificity signature35.* ***E.*** *Tumor-specific CD8+ T cell fractions relative to all T cells in pretreatment biopsies of patients with and without clinical response (cohorts A and B). n = 25 patients.* ***F.*** *Tfh fractions relative to all T cells in pretreatment biopsies of patients with and without clinical response (cohorts A and B). n=25 patients.* ***G-H.*** *Ki-67 expression on (G) PD1+ CD8+ T cells and (H) conventional CD4+ T cells pretreatment in peripheral blood of patients with and without clinical response in cohorts A and B. n= 25 patients* ***I.*** *Dotplot for PDCD1 and MKi67 expression in CD4+ T cell clusters (tumoral, scRNA-Seq, cohorts A and B).* ***J.*** *Dotplot for PDCD1 and MKi67 expression in CD8+ T cell clusters (tumoral, scRNA-Seq, cohorts A and B).* ***K.*** *Fraction of Ki-67+ Tfh cells relative to all T cells in pretreatment biopsies of patients with and without clinical response (cohorts A and B). n=25 patients.* ***L.*** *Fraction of proliferating PD1+ CD8+ T cells relative to all T cells in pretreatment biopsies of patients with and without clinical response based on single-cell RNA-seq data (cohorts A and B). n=25 patients. In* ***E–F, K-L*** *boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. P-values were derived using a two-sided Mann-Whitney test.*

Tumors of clinical responders harbor preexisting inflammatory profiles and tumor-specific CD8+ T cells

We conducted in-depth analyses between clinical responders and non-responders using bulk RNA-Seq (all cohorts) and single-cell RNA-Seq and TCR sequencing (cohorts A and B) pre- and post-treatment. Bulk RNA-Seq revealed higher pretreatment levels of IFNG gene expression (p=0.0003, Fig.2C) and inflammatory gene signatures in clinical responders (p<0.05 for all, FDR 10%, Fig.2G, ED Fig.3A-E). Clinical responders also exhibited higher gene signatures associated with immune infiltration (p<0.05 for all, FDR 10%, Fig.2G, ED Fig.3F-J). Conversely, clinical non-responders displayed upregulation of TGF-beta and Notch signaling (p<0.05 for both, FDR 10%, Fig.2D, ED Fig. 3L-M). Though TIL levels and distances from tumor cells to CD8+ T cells were not different in responders versus non-responders in cohort C that included TIL high patients only, patients with pCR had significantly higher pretreatment IFNG gene expression (Fig. 2F) and higher scores of gene signatures related to immune response and T cell infiltration (Fig. 2H), consistent with our previous observations of a more inflammatory profile of the tumor microenvironment in clinical responders in cohorts A and B.

After single-cell RNA-Seq data preprocessing, we obtained 80 000 high quality T cells from 52 samples (29 patients). Following unsupervised clustering of the T cells, we identified various subpopulations (Fig.3A-D, ED Fig.4A-T), including CD8+ effector T cells, CD8+ tissue resident memory (CD8+ TRM) T cells, proliferating CD8+ T cells, naive CD4+ T cells, follicular B helper T cells (Tfh), memory CD4+ T cells, regulatory T cells (Tregs), CD56bright and CD56dim NK cells. Intriguingly, we identified a cluster of CD8+ T cells with features of tumor-

specific T cells. This cluster was characterized by the highest clonality and highest expression of tumor recognition signatures derived using functional tumor recognition experiments34,35 (Fig.3C-D). This CD8+ tumor-specific cluster was marked by high expression of tumor-reactive markers (CD39, CD103, PDCD1), IFNG, effector molecules (GZMB, NKG7, PRF1, GNLY), chemokines (CCL5, CCL4, CXCL13, CCL3) and exhaustion markers (LAG3, HAVCR2, TIGIT, TOX, CTLA4, Fig.3C,D). Clinical responders exhibited higher fractions of pretreatment CD8+ tumor-specific T cells (Fig.3E). This is a first report of tumor-specific T cell population identified using single-cell RNA-Seq in clinical trial data showing an association with response. Clinical responders also had higher fractions of CD4+ Tfh cells (Fig.3F). Presence of tumor-specific CD8+ T cells and Tfh in pretreatment biopsies was correlated with tumor decrease on MRI, indicating a continuous association between the abundances of these cells pretreatment and the depth of the tumor response (ED Fig.4U, V). Patients with different TIL levels had similar T cell subtypes pretreatment (EDFig.4W).

Flow cytometry of blood samples (19 markers, ED Table 2, ED Fig.5A) revealed increased Ki-67-positive cells within the PD1+ conventional CD4+ T cell population in clinical responders (p=0.005, Fig.3G). A similar trend was observed for CD8+ T cells (Fig.3H). The increased proliferation of PD1+CD4+ T cells observed in the blood could also be traced back to the tumor, with responders having higher levels of Ki67+ Tfh which was the CD4+ T cell cluster with the highest PDCD1 gene expression in the tumor single-cell RNA-Seq data (Fig.3I, L). In line with the blood data, the levels of PD1+ proliferating CD8+ T cells were not significantly different between clinical responders and non-responders (Fig.3J-K), suggesting a specific role for proliferating CD4+ T cells systemically as well as in the TME.

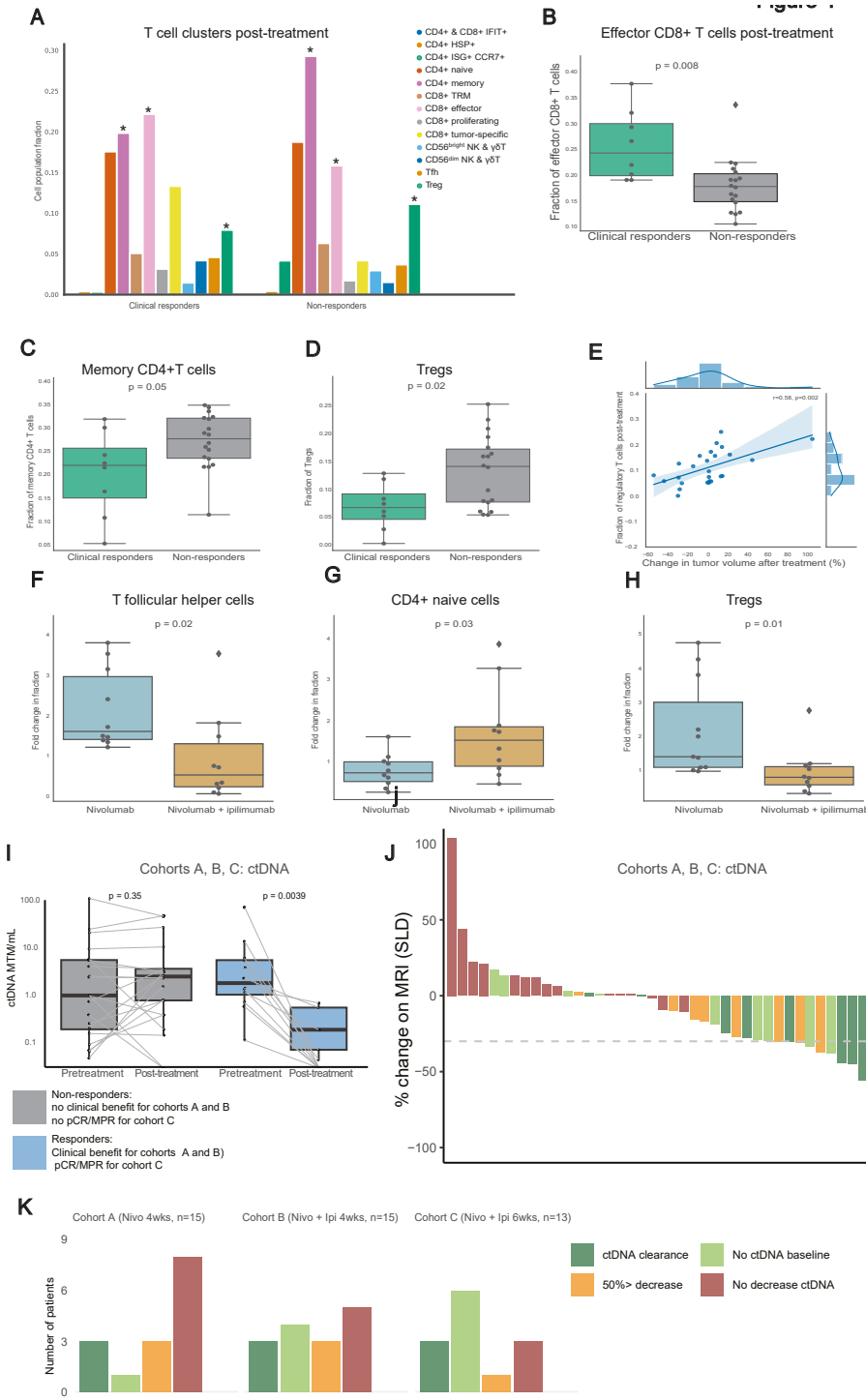


Figure 4: Effects of anti-PD1 ± anti-CTLA4 on the T cell profiles in the tumor microenvironment after treatment in cohorts A and B. ctDNA data for all cohorts. A. *Fractions of different T cell clusters relative to all T cells in post-treatment biopsies of patients who did (left) and did not (right) experience clinical response based on single-cell RNA-Seq data.* ***B.*** *Effector CD8+ T cell fractions relative to all T cells in post-treatment biopsies versus response (cohorts A, B). n=26 patients.* ***C.*** *Memory CD4+ T cell fractions relative to all T cells in post-treatment biopsies versus response (cohorts A and B). n=26 patients.* ***D.*** *Regulatory T cell fractions relative to all T cells in post-treatment biopsies versus response (cohorts A, B). n=26 patients.* ***E.*** *Fractions of regulatory T cells relative to all T cells in post-treatment biopsies of patients (cohorts A and B) in relation to the change in tumor volume after treatment assessed using MRI (RECIST 1.1). n=26 patients.* ***F-H.*** *Fold changes in fractions of T cell populations relative to all T cells in cohort A and cohort B. n=22 patients.* ***F.*** *Follicular B helper T cells.* ***G.*** *Naive CD4+ T cells.* ***H.*** *Regulatory T cells.* ***I.*** *Changes in circulating tumor DNA (ctDNA) levels of responding and non-responding patients upon treatment. Patients from all cohorts (A, B, C) were included.* ***J.*** *Waterfall plot of all patients (n=46, all cohorts) colored according to the fold change in ctDNA levels in blood upon treatment. The groups represent: ctDNA clearance; post-therapy decrease in ctDNA levels of 50% or more post therapy; no ctDNA at baseline; no decrease in ctDNA. The gray dashed line at -30%: radiological PR.* ***K.*** *Barplots summarizing the number of patients for each ctDNA response category in each cohort (A, B, C). ctDNA at baseline was available for 43/46 patients. In* ***B-D, F-I*** *boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. P-values in* ***B-D, F-H*** *were derived using a two-sided Mann-Whitney test. P-values in I were derived using paired Wilcoxon test.*

Dynamics and post-treatment composition of the tumor microenvironment are distinct in clinical responders and non-responders

Single-cell RNA-Seq analysis revealed that though the clinical responders had higher proportions of tumor-specific CD8+ T cells pretreatment, post-treatment their tumors included higher levels of effector CD8+ T cells compared to non-responders (p=0.008, Fig.4A, B). This suggests that effector CD8+ T cells contribute to ICI-induced tumor regression and underscore the ongoing antitumor CD8+ T cell response even four weeks after treatment initiation.

Conversely, non-responders had elevated memory CD4+ T cells (p=0.05, Fig.4A, C) and Tregs (p=0.02, Fig.4A, D) post-treatment, potentially suggesting the involvement of Tregs in mediating resistance to ICI, consistent with prior studies[36]. Intriguingly, we observed an association between the fraction of Tregs after treatment and the lack of response or in some patients even increase in tumor volume on MRI (Fig.4E). This correlation was specifically mediated by activated (CD137+) Tregs, rather than non-activated Tregs (ED Fig.5B-C).

We also investigated whether the addition of anti-CTLA4 led to differential alterations in the TME compared to nivolumab monotherapy, though the study was not powered for cohort comparisons. Patients receiving nivolumab plus ipilimumab showed a reduced fold change in Tfh cells (p=0.02, Fig.4F), but an increased fold change in naive CD4+ T cells (p=0.03, Fig.4G). Additionally, the combination ICI resulted in a decreased fold change in Tregs (p=0.01, Fig.4H) compared to monotherapy, including both activated and non-activated Tregs (ED Fig.5D-E).

ctDNA dynamics during early response to ICI

To assess the impact of short-term ICI on circulating tumor DNA (ctDNA), we conducted ctDNA analysis pretreatment and after four weeks (cohorts A and B) or six weeks (cohort C) of ICI using a tumor-informed ctDNA assay (Signatera). Despite the early tumor stages included (mostly I-II), pretreatment ctDNA was detected in 32/43 (74%) patients. After treatment, 9 (21%) patients had complete ctDNA clearance, while additional seven patients had a reduction of ≥50% in ctDNA load (MTM/mL, Fig.4I-J). All clinical responders in cohorts A+B and pCR/MPR patients (n=8) in cohort C demonstrated at least a 50% drop in ctDNA or were negative for ctDNA at baseline (Fig.4I-K).

Discussion

In this study, we demonstrate that neoadjuvant nivolumab, with or without ipilimumab, is a feasible chemotherapy-free regimen for patients with early stage TNBC. We show that nivolumab ± ipilimumab induces immune activation in the majority of patients and can result in complete pathological responses and ctDNA clearance. Pre-existing inflammatory features such as higher TILs, shorter distances from CD8+ T cells to the tumor and higher baseline fractions of tumor-specific CD8+ T cells were associated with response. In contrast, higher fractions of Tregs post-treatment were associated with lack of response. While standard chemo-immunotherapy for TNBC with 4 chemotherapy agents plus anti-PD1 is a 5-month treatment regimen leading to a 63% pCR rate, our work suggests that with only six weeks of anti-PD1 plus low-dose anti-CTLA4 a 33% pCR rate may be obtained in TNBCs with high TILs. This suggests that for some patients a short-term immunotherapy-first approach may be an option if confirmed by future research in larger cohorts with more robust follow-up. However, a substantial group of patients still needs chemotherapy and/or longer treatment in order to obtain a pCR. Although we did not observe any unexpected toxicity, the rate of persisting endocrinopathies, in particular hypothyroidism, was high compared to reports in other tumor types or in breast cancer when anti-PD(L)1 is added to neoadjuvant chemotherapy. Although the 33% pCR rate would allow expansion of cohort C to stage II, with 40% grade 3-4 toxicity, 40% hypothyroidism and 20% adrenal gland insufficiencies, substantial toxicity is a serious concern, especially considering the relatively good prognosis of TNBC patients with high TILs.

To our knowledge, the BELLINI trial is the first to investigate the feasibility and potential efficacy of ICI without concurrent chemotherapy in early stage TNBC. Moreover, for the first time, the scoring of TILs is used as an inclusion criterion to select patients with a good prognosis for whom development of de-escalated treatment regimens is most promising. Larger clinical trials also using TILs according to this workflow when including patients have recently started (NCT05929768). In addition, ETNA trial (NCT06078384) will explore whether stage I TNBC patients with high TILs can forgo (neo)adjuvant chemotherapy or be treated with immunotherapy alone. The larger international OPTImaL patient preference study (NCT06476119) will also allow the option of no chemotherapy for this patient population. In addition, other studies use TILs as inclusion criteria for immunotherapy-first approaches: Pop-Durva (NCT05215106) and pan-cancer NEOASIS trial (NCT06279130). Further studies that are sufficiently powered to assess long-term outcomes are needed on the use of TILs or other immune-based biomarkers as entry criteria for immunotherapy or de-escalation studies, especially since patients with lower stage TNBC and high TILs can have an excellent outcome with local treatment alone[19,37].

Immune-related endocrine disorders were the most common adverse events observed. Specifically, 41% of the patients developed hypothyroidism, which, though usually easy to

manage, is a permanent condition, and 13% developed adrenal insufficiency, a serious long-term toxicity. Comparable neoadjuvant ICI-only studies with nivolumab + low-dose ipilimumab in head and neck squamous carcinoma, colorectal cancer, urothelial carcinoma and melanoma reported hypothyroidism in 4-8% of patients[9–11,14] and adrenal insufficiency in 0-8% of patients[9–11,14]. However, the recent largest phase III trial (stage III melanoma, n=423) reports substantial higher rates of endocrinopathies with 23.6% hypothyroidism and 9.9% adrenal gland insufficiency[26]. Importantly, for cancer types with poor prognosis such as stage III melanoma, high toxicity rates might be acceptable, while this is different for patient populations with more favorable outcomes. The higher rates of hypothyroidism and adrenal insufficiencies in BELLINI compared to these studies could stem from different patient demographics. Patients with TNBC are typically female and relatively young, potentially contributing to different systemic immunity and adverse event incidence[38]. In BELLINI, we reported all immune-mediated adverse events during the first year of follow-up, with 4/6 patients developing adrenal insufficiency >100 days since inclusion. Trials with shorter reporting periods may miss these late events, leading to underreported delayed toxicity, especially in centers not specialized in evaluating ICI regimens. When focusing on patients with similar demographics and disease, we still observe a higher rate of endocrine adverse events in BELLINI compared to neoadjuvant trials for TNBC evaluating ICI plus chemotherapy. The KEYNOTE-522 trial reported thyroid dysfunction in 22% of patients treated with anti-PD1 plus chemotherapy[1]. Adrenal insufficiency/hypophysitis was reported for 4.5% of patients in the KEYNOTE-522 study. A recent study with an oncolytic virus without chemotherapy found that 3/6 breast cancer patients developed hypothyroidism[39], which is more in line with our observations. The lower hypothyroidism rate in the KEYNOTE-522 compared to the oncolytic virus study[39] and BELLINI could suggest that chemotherapy results in partial blunting of the immune response. Lastly, the preselection of patients with higher TILs in BELLINI may have resulted in patients that are more likely to develop immune-related adverse events due to a different systemic immunity. We also cannot rule out the influence of chemotherapy given after ICI, where steroids are used as antiemetics. Our cohort sizes are too small to compare toxicities induced by 4-week nivolumab versus 4-week nivo-ipi versus 6-week nivo-ipi. However, in the latter group, we observed more non-endocrinopathies such as colitis, hepatitis, and pneumonitis, while endocrinopathies were already remarkably high with nivolumab monotherapy. This potentially signifies that neoadjuvant ICI without chemotherapy could result in a higher rate of hypothyroidism in breast cancer patients. Of note, it was demonstrated that immunotherapy-related thyroid dysfunction and other immune-related adverse events are associated with improved survival in multiple cancer types[40,41,42,43]. Nevertheless, upfront prediction of risk of immunotherapy-related toxicity for individual patients is a large unmet clinical need and the burden of adverse events should be evaluated in light of the prognosis of each patient[44].

The advantage of WOO studies like BELLINI is the opportunity to evaluate promising drugs and drug combinations in an efficient manner and to analyze pre- and post-treatment tumor material that can provide insights into the therapy effects. Our primary endpoint, immune activation defined as doubling of CD8+ T cells and/or IFNG expression, was reached in 17/30 patients (57%). Although both cohorts reached the >30% immune activation rate, allowing cohort expansion, we observed more doubling of CD8+ T cells in patients with low pretreatment levels of these features. This could be due to the biopsy timing with deep responses at 4 weeks in tumors with high endogenous CD8+ T cells and/or a 'saturation' of CD8+ T cells in patients with high pretreatment values. In contrast to CD8+ T cells, IFNG counts may double even with high pretreatment values, however, they could also be impacted by decreased antigen availability in case of tumor regression. This suggests that different biomarker approaches could apply to inflamed and non-inflamed tumors. Recent insights from the developments of personalized neoadjuvant immunotherapy in melanoma indicate that patients with high pre-existing IFNG levels or a significant increase in IFNG signature upon treatment were most likely to benefit45. The disadvantage of WOO designs with short scheduled treatments is the non-guaranteed benefit for participating patients. Also, information on established endpoints such as pCR rate is needed before a novel treatment approach will be tested in larger trials. For this reason, the adaptive BELLINI trial allowed opening of new cohorts with established endpoints to bring therapies to the next step. Although allowed by the protocol and statistical analysis plan, reporting only stage I data of a Simon's 2-stage design comes with the risk of false-positive findings. Similarly to the cohorts A and B, cohort C also reached the threshold of sufficient responders to expand into stage II. However, given the relatively high rate of endocrinopathies, which are chronic, cohort C was not expanded to stage II. In this view, testing novel anti-CTLA4-targeting antibodies, such as botensilimab46, intentionally designed to overcome the limitations of conventional ICI such as persisting endocrinopathies could be interesting for breast cancer patients.

When analyzing pretreatment tumor characteristics in high-TIL tumors only (cohort C), we found that the inflammatory phenotype and markers were still discriminative between responders and non-responders and remarkably similar to the clinical responders and non-responders in cohorts A+B. In cohort C, pathological complete responders had higher inflammatory gene expression profiles pretreatment, including signatures for IFNG response, checkpoint molecules, exhausted CD8+ T cells and immunogenic cell death. This suggests that, even in patients with high TILs, the profiling of baseline inflammatory status may facilitate early identification of (non)responders and should be considered in addition to TILs.

The recent publication of the tumor-specific T cell signatures[34,35] enabled us to identify and follow tumor-specific CD8+ T cells in a clinical trial setting. Importantly, using these signatures as a proxy for the tumor reactivity, we demonstrate for the first time that the presence of tumor-specific CD8+ T cells pretreatment is linked to ICI response.

Additionally, we observed decreased fractions of Tregs in clinical responders compared to non-responders after treatment, in line with prior reports on the role of Tregs in resistance to ICI47. In a resistant mouse tumor model, anti-PDL1 therapy led to Treg activation, and Tregs were shown to be activated in the single-cell data of NSCLC and basal cell carcinoma patients not responding to anti-PD(L)1 ICI[36]. In this recent study, ICI treatment induced higher expression of genes involved in Treg-mediated immune suppression (PDCD1, CTLA4, CD38) and cell cycle (MKI67) in Tregs from the tumors of non-responders[36]. Together, these findings demonstrate that Treg cells might play a critical role in resistance to ICI.

To date, data on combining anti-PD(L)1 with low-dose anti-CTLA4 was lacking in early-stage breast cancer. Due to the non-comparative design and the small sample size, our data on the potential additive effect of ipilimumab should be considered exploratory. At the single-cell level, the addition of ipilimumab resulted in lower fold change in Tregs in the TME upon treatment. We also observed a correlation between higher levels of activated Tregs post-treatment and the lack of response or in some cases even slight increase in tumor volume on MRI. This suggests that activated Tregs play a role in resistance to immune checkpoint blockade and that depleting activated Tregs could be a promising strategy for TNBC patients unresponsive to anti-PD1-based treatments. Of note, we cannot exclude that the lack of response or the increase of tumor volume observed by imaging was in part due to pseudoprogression. A growing body of literature analyzing anti-CTLA4 using in-vivo models indicates that anti-CTLA4 can deplete Tregs48. However, whether anti-CTLA4 can deplete Tregs in human tumors remains a matter of debate49. A recent study by van der Leun et al. in head and neck squamous cell carcinoma also demonstrated an increase in transitional CD8+ T cells and a decrease in CD137+ Tregs in responders after treatment with anti-PD1 and anti-CTLA4 therapy[50], indicating that this might be a consistent pattern across multiple tumor types.

After the results of the landmark trials in early stage TNBC that added PD1 blockade to standard neoadjuvant chemotherapy[1,3,51,52], our current data provide a rationale to further explore the following observations. First, we observed complete and near-complete pathological responses after only six weeks of treatment with ICI in patients with high TILs. This suggests that a subgroup of TNBC could be treated with chemo-free regimens if further research powered for long-term outcome analysis will confirm our results. More research is needed on the optimal selection strategy and treatment regimen, especially in view of the observed high endocrinopathy rate. It is tempting to speculate whether extending the six-week treatment period could result in higher pCR rates and thereby reach responses similar to outcomes obtained with chemo+IO. This can only be done if accompanied toxicity would not increase. However, it remains unknown whether pCR after immunotherapy has the same prognostic value as pCR after chemotherapy. Therefore, larger trials are needed to validate the pCR rate after short-term ICI alone and to determine if this results in excellent

survival rates, as seen in other cancers[53,54]. Moreover, pCR might not be the optimal endpoint since KEYNOTE-522 and GeparNUEVO have indicated that the benefit of PD1-blockade is not exclusively seen in patients with pCR[51,55]. Second, our exploratory clinical and translational data suggest that combination ICI is feasible and could potentially enhance the effects of PD1 blockade. However, the benefit-risk ratio of such combinations should always be carefully monitored. Third, establishing the feasibility of patient inclusion based on TIL opens the door for more immune biomarker-driven trials, which is particularly important in diseases like TNBC that include both inflamed and non-inflamed subtypes. The potential integration of additional inflammation analyses, for example, using IFNG gene expression on top of TILs as suggested by our data, may optimize patient selection, increase pCR rates for ICI-only approaches and could help treatment personalization in the future. Lastly, a substantial fraction of patients achieved ctDNA clearance after short-term ICI. Given the strong prognostic value of early ctDNA decrease, as shown by the I-SPY trial[56], future studies are needed to investigate the feasibility and reliability of TILs-informed patient inclusion and the potential of ctDNA-informed therapy adjustments.

Acknowledgements

We are grateful to the participants and their families for participating in the trial. We thank all supporting clinical trial staff, in particular nurse specialists and the Departments of Medical Oncology, Surgery, Radiology and Pathology of the participating centers. We thank the NKI trial lab for handling incoming blood samples. We are grateful to Annegien Broeks and the Core Facility of Molecular Pathology and Biobanking for the storage and handling of human tumor material and to the Genomics Core Facility for DNA and RNA sequencing. We acknowledge the supporting staff of the clinical trials of the Departments of Medical Oncology and the Triallab. We thank Chris Klaver and Maxime Duijst for blood sample experiments and the Flow Cytometry Facility the AKL for support in these experiments. We thank Bauke Stegenga from Bristol Myers Squibb for trial support and arranging supply of nivolumab and ipilimumab. R.S. is supported by the Breast Cancer Research Foundation (grant number 17-194). Research at the NKI is supported by institutional grants of the Dutch Cancer Society and of the Dutch Ministry of Health, Welfare and Sport. Research in the laboratory of M.K. is funded by the Netherlands Organization for Scientific Research (VIDI 09150172010043) and Victoria's Secret Global Fund for Women's Cancers Rising Innovator Research Grant, in Partnership with Pelotonia & AACR (23-30-73-KOK). Trial costs were supported by Bristol Myers Squibb. The funders had no role in study design, data collection or analysis, decision to publish or preparation of the manuscript.

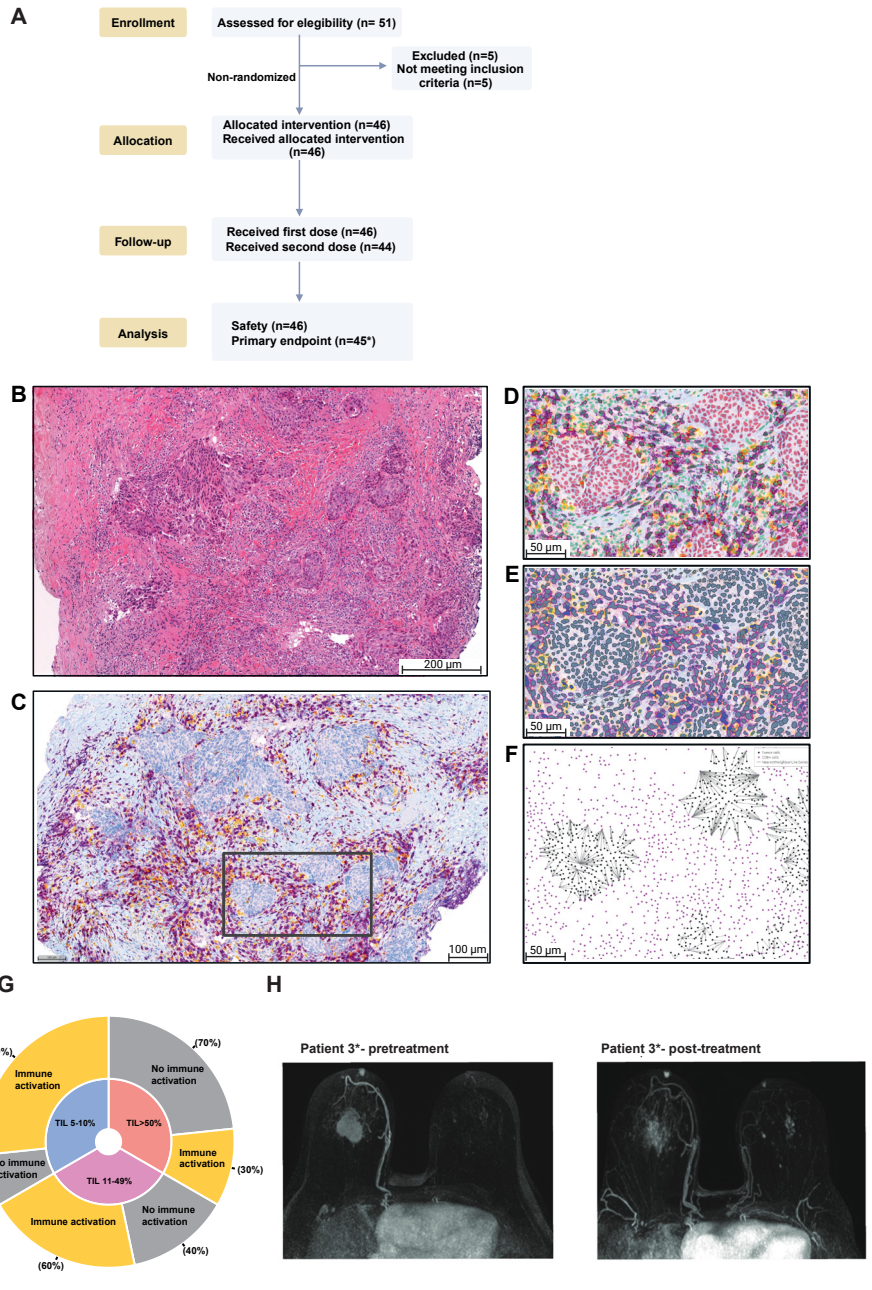
Author Contributions

I.N. and O.I.I. contributed equally to this work as shared co-first authors. M.d.G., R.C.A.M.G. and N.A.M.B. contributed equally to this work as shared co-second authors. I.N. wrote the study protocol, coordinated trial procedures, performed wet lab experiments for single-cell RNA-Seq analyses, analyzed and interpreted clinical and translational data of the trial. O.I.I. designed, performed and interpreted computational analyses of the DNA, bulk and single-cell RNA sequencing data, analyzed and interpreted translational data. I.N., O.I.I., and M.K. wrote the paper. N.A.M.B., E.C. and H.G. were responsible for blood sample processing and analysis, supervised by K.E.d.V. and M.K., and H.G. designed the flow cytometry panel. M.d.G. performed the spatial analyses and helped with collection and analysis of clinical data. R.C.A.M.G. and A.L.R. coordinated trial procedures and collected clinical data for cohort C. B.B., J.G.H.T and M.C. performed bioinformatics analyses and contributed to their design. I.A.M.M. was the clinical projects manager. M.d.M. processed FFPE for IHC and isolated DNA and RNA from tissue biopsies. T.v.B. performed wet lab experiments for single-cell RNA-Seq analyses. M.L.-Y. performed the statistical analysis of the trial data. J.G.v.d.B., N.K.., H.M.H., K.v.d.V. and R.S. performed the TILs and histological scoring of the pathology slides. I.H. developed and performed double CD8-PD1 staining. R.M.M. and C.E.L. revised MRI scans and, together with colleagues, were involved in taking biopsies. Ek.K. organized the ctDNA experiments. Em.K., F.H.v.D., V.S., S.L., C.A.D., M.G.J.v.D., G.S.S., S.C.L. and M.K. were the main treating physicians. I.N., I.A.M.M, G.S.S., T.N.S., C.U.B., S.C.L. and M.K. wrote the trial protocol. H.M.H supervised the computational pathology analyses. L.F.A.W. and D.L. supervised computational analyses. M.K. was the principal investigator of the trial, supervised all the analyses presented in the paper and acquired funding. All authors edited and approved the paper.

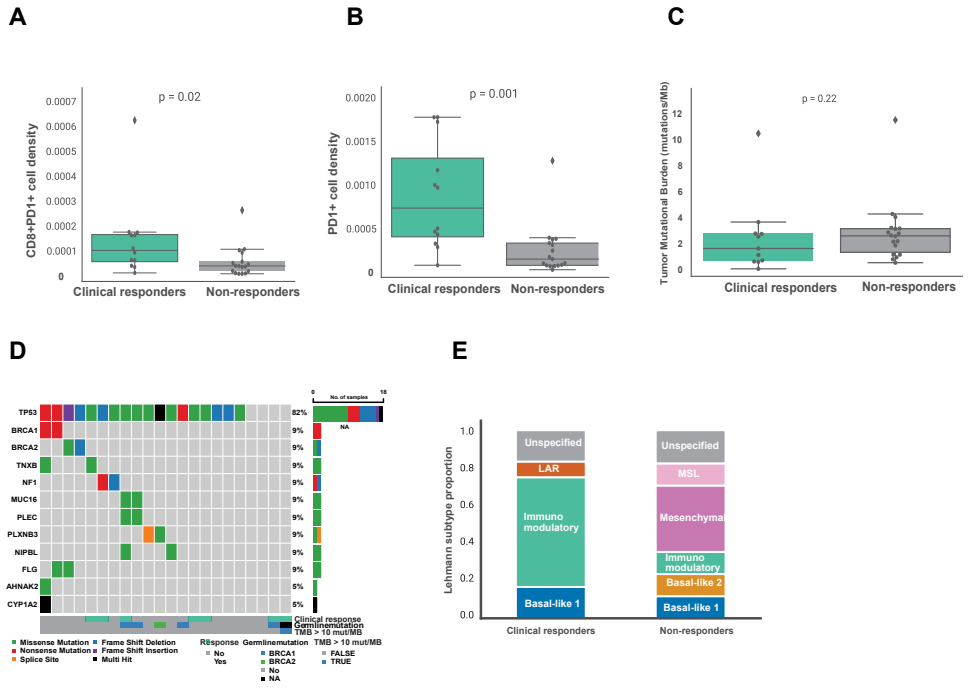
Competing interests

I.N., O.I.I., M.d.G, R.C.A.M.G., N.A.M.B., A.L.R., H.G., B.B., J.G.H.T., I.A.M.M., M.d.M, T.v.B., M.C., E.C., M.L.-Y., K.v.d.V., J.G.v.d.B., I.H., N.K., C.E.L., F.H.v.D., V.S., S.L., Em.K., C.A.D., M.G.J.v.D., H.M.H. and D.L. have no competing interests to declare. R.M.M. reports research grants from Siemens Healhtineers, Bayer Healthcare, Screenpoint Medical, Beckton & Dickinson, PA Imaging, Lunit and Koning, and is an advisory board member for Screenpoint, Bayer, Siemens and Guerbet, all outside the scope of this work. Ek.K. is an employee of Natera, Inc. G.S.S. reports research funding to the institute from Merck, Agendia, AstraZeneca, Roche and Novartis and a consulting role for Novartis, Seattle Genetics and Biovica, outside the submitted work. S.C.L. reports research funding to the institute from Roche/Genentech, AstraZeneca, BMS, Tesaro, Merck, Immunomedics, Eurocept Pharmaceuticals, Agendia and Novartis and a consulting role and travel grant from Daiichi Sankyo, outside this work. C.U.B. has received research grants from Novartis, BMS and NanoString, is a paid advisory board member for BMS, MSD, Roche, Novartis, GlaxoSmithKline, AstraZeneca, Pfizer, Lilly, GenMab and Pierre Fabre and holds ownership interest in Uniti Card, Neon Therapeutics and Forty Seven, all outside this submitted work. K.E.d.V. reports research funding from Roche and is a consultant for Macomics, outside the scope of this work. R.S. reports non-financial support from Merck and Bristol Myers Squibb (BMS), research support from Merck, Puma Biotechnology and Roche and personal fees from Roche, BMS and Exact Sciences for advisory boards, all outside the scope of this paper. L.F.A.W. reports funding to the institute from Genmab BV. V.C.G.T.-H. reports research funding to the institute from Roche, Eisai, Pfizer, Novartis, Lilly, Daiichi Sankyo/AstraZeneca and Gilead Sciences, a consulting role from Pfizer, Lilly, Accord Healthcare and Novartis and honoraria from Novartis, Roche, Lilly and AstraZeneca, all outside this submitted work. T.N.S. is advisor for Allogene Therapeutics, Asher Bio, Merus, Neogene Therapeutics, and Scenic Biotech; is a stockholder in Allogene Therapeutics, Asher Bio, Cell Control, Celsius, Merus, and Scenic Biotech; and is venture partner at Third Rock Ventures, all outside of the current work. M.K. reports research funding to the institute from BMS, Roche and AstraZeneca/MedImmune and an advisory role/speakers' fee (all compensated to the institute) for Alderaan, BMS, Domain Therapeutics, Medscape, Roche, MSD and Daiichi Sankyo, outside the submitted work. Natera provided non-financial support to this study.

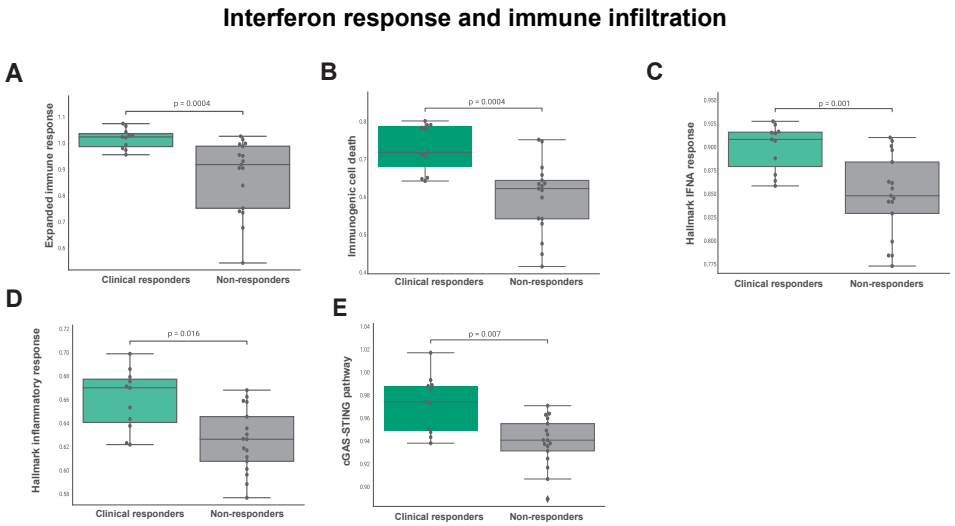
Supplementary Material



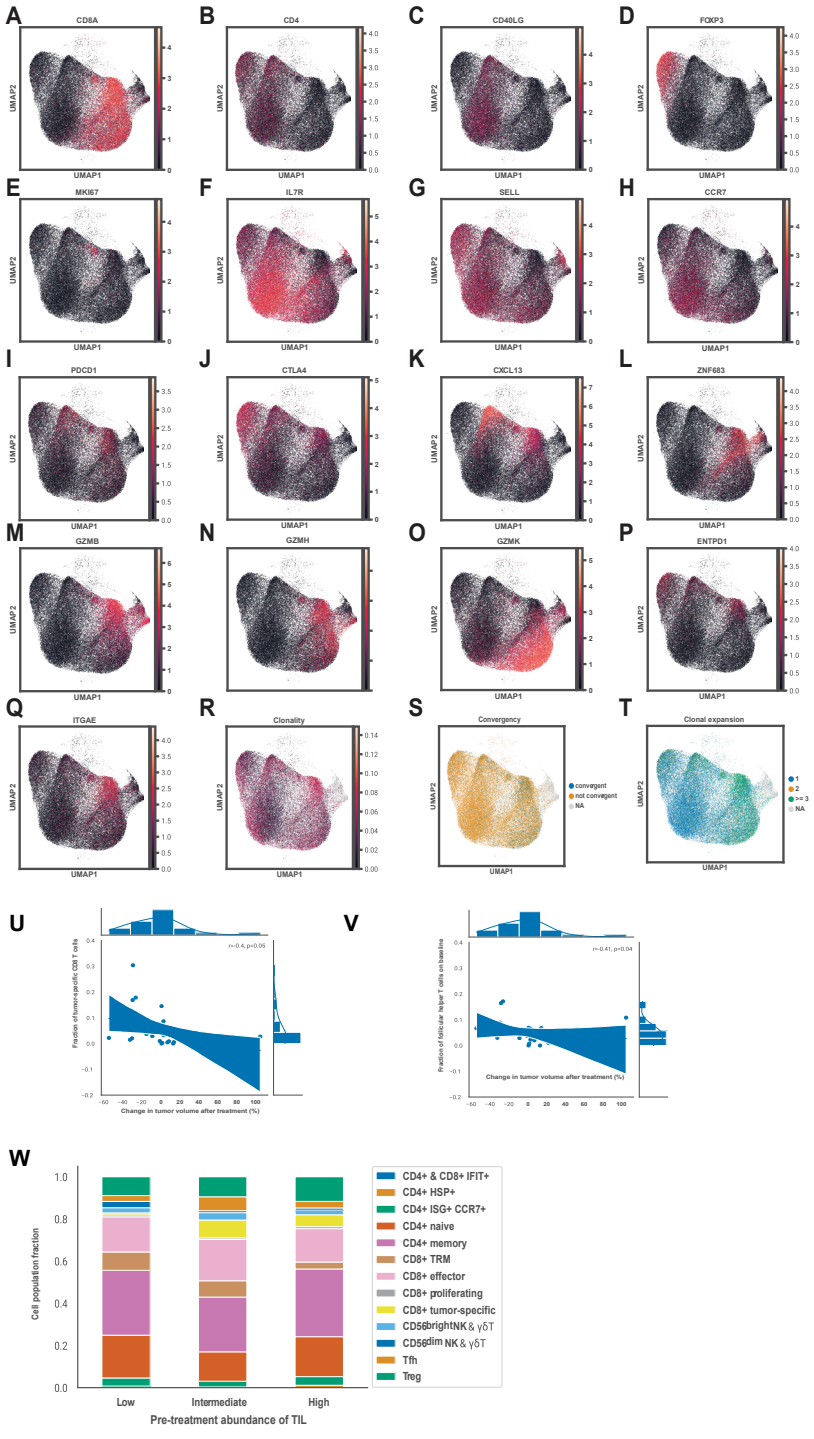
◀◀◀ ***Extended Data Fig. 1 IHC CD8 + T cell analysis. A.*** *CONSORT Flow Diagram. Consort diagram of patients eligible, recruited, numbers followed up and included in analysis. *max 15 patients per cohort analyzed for primary end point.* ***B.*** *H&E-stained image, corresponding to CD8/PD-1 stained tissue under* ***C. C.*** *Representative example of a CD8/PD-1 double-stained tissue (haematoxylin = blue, PD-1 = yellow, CD8 = purple).* ***D.*** *Representative example of the performance of the AI-based tumor cell classifier Tumor classification (red) and nontumor cells (green).* ***E.*** *Example of cell segmentation and tumor phenotype assignment. Cell with purple border = CD8+ cell, yellow border = PD-1+ cell, orange border = PD-1 + CD8+ cell.* ***F.*** *Corresponding distance analysis in the same tissue area as under* ***D*** *and* ***E.*** *The grey lines represent the shortest distance from a tumor cell to its nearest CD8 + T cell.* ***G.*** *Proportions of patients reaching immune activation stratified according to TIL levels at inclusion in cohorts A and B. 10 patients had 5–10% TILs, 10 patients 11–49% TILs and 10 patients had 50% or more TILs.* ***H.*** *Pretreatment and post-treatment MRI images of patient #3 with a pathological complete response (pCR) at surgery after ICI only (cT2N0, ypT0N0). Figure* ***A*** *was created with BioRender.com. In* ***A-B****, one biopsy was analyzed per patient.*



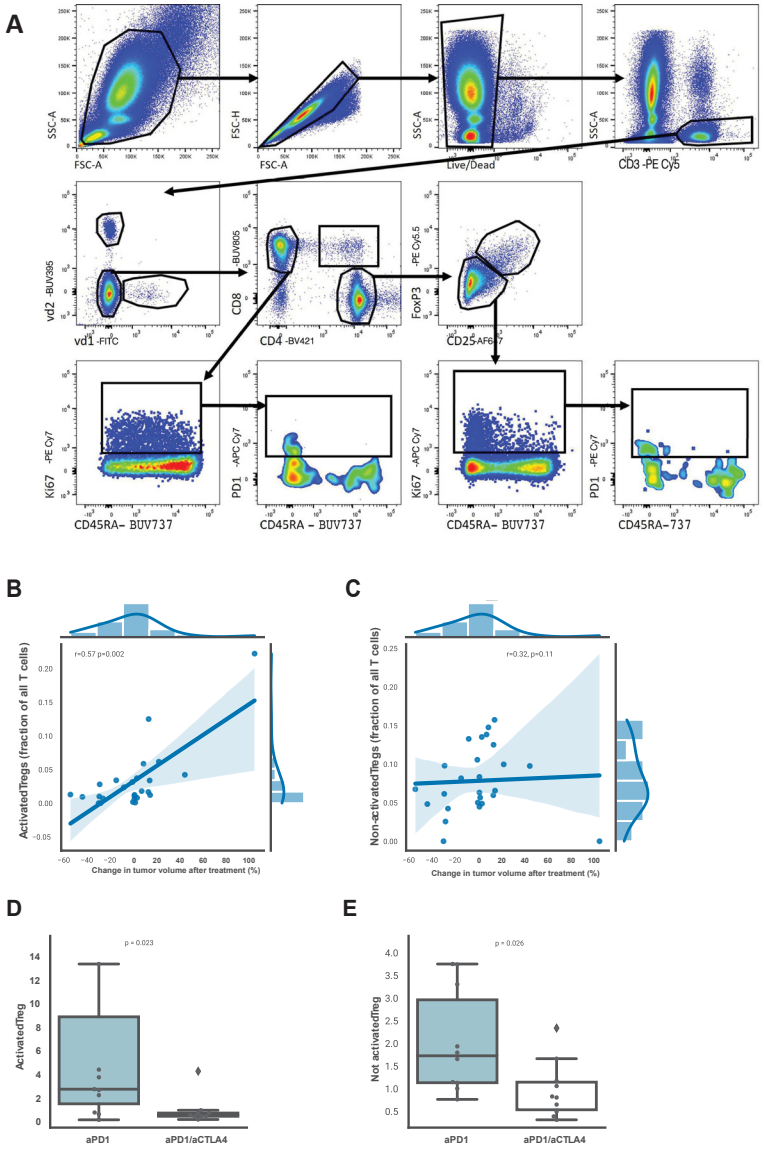
Extended Data Fig. 2 Baseline tumor microenvironment features and genomic profile of cohorts A and B. A. *PD-1 + CD8 + T cell density in pretreatment biopsies of patients with and without who did and did not experience clinical response in cohorts A and B. n = 31 patients.* ***B.*** *PD-1+ cell density in pretreatment biopsies of patients with and without who did and did not experience clinical response in cohorts A and B. n = 31 patients.* ***C.*** *Tumor mutational burden (TMB) in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n = 30 patients. Boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. Data were analyzed by a two-sided Mann–Whitney test.* ***D.*** *Oncoplot of TMB (mutations per megabase (Mb)) and top mutated genes in cohorts A and B.* ***E.*** *Proportions of Lehmann et al. subtypes33 in patients with and without clinical response in cohorts A and B. MSL, mesenchymal stem-like; LAR, luminal androgen receptor.*



◀◀◀ ***Extended Data Fig. 3 Gene signatures in pretreatment biopsies associated with clinical response in cohorts A and B. A–L****. Gene set expression scores in pretreatment biopsies of patients with and without clinical response in cohorts A and B. n = 28 patients.* ***A.*** *Expanded immune signature from Ayers et al.56* ***B.*** *Immunogenic cell death signature57.* ***C.*** *Hallmark IFNA response gene set.* ***D.*** *Hallmark inflammatory response gene set.* ***E.*** *cGAS-STING pathway gene set58.* ***F.*** *Effector CD8 + T cell gene set59.* ***G.*** *Exhausted T cell gene set59.* ***H.*** *Checkpoint molecules gene set59.* ***I.*** *Naive T cell gene set60.* ***J.*** *Tertiary lymphoid structures gene set61.* ***K.*** *Hallmark TGF-beta signaling gene set.* ***L.*** *Hallmark Notch signaling. In* ***A–L****, boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. P values were derived using a two-sided Mann–Whitney test. Reported p values were significant after Benjamini–Hochberg (FDR) correction at 10% significance level.*



◀◀◀ ***Extended Data Fig. 4 Single-cell RNA-Seq pretreatment tumor microenvironment profile of the cohorts A and B. A-Q****. UMAP representations of the marker gene expression in the dataset.* ***A.*** *CD8A.* ***B.*** *CD4.* ***C.*** *CD40LG* ***D.*** *FOXP3* ***E.*** *MKI67* ***F.*** *IL7R.* ***G.*** *SELL.* ***H.*** *CCR7.* ***I.*** *PDCD1.* ***J.*** *CTLA4.* ***K.*** *CXCL13.* ***L.*** *ZNF683.* ***M.*** *GZMB.* ***N.*** *GZMH.* ***O.*** *GZMK.* ***P.*** *ENTPD1.* ***Q.*** *ITGAE.* ***R.*** *UMAP representation of the T cell clonality in the dataset.* ***S.*** *UMAP representation of the T cell clone convergence in the dataset.* ***T.*** *UMAP representation of the T cell clonal expansion in the dataset.* ***U.*** *Fractions of tumor-reactive CD8 + T cells relative to all T cells in pretreatment biopsies of patients based on single-cell RNA-Seq data in relation to the change in tumor volume after treatment based on RECIST 1.1 in cohorts A and B.* ***V.*** *Fractions of Tfh cells relative to all T cells in pretreatment biopsies of patients based on single-cell RNA-Seq data in relation to the change in tumor volume after treatment based on RECIST 1.1 in cohorts A and B.* ***W.*** *Fractions of different T cell clusters relative to all T cells based on single-cell RNA-Seq data in pretreatment biopsies of patients who had low (5–10%), intermediate (11–49%) and high (>=50%) presence of tumor-infiltrating lymphocytes before treatment in cohorts A and B. In* ***U-V****, correlation was estimated with Spearman's rank correlation coefficient, two-sided, with 95% confidence interval for the regression estimate.*



Extended Data Fig. 5 Gating strategy used for the flow cytometry data analysis and activated and non-activated Tregs in cohorts A and B. A. *Gating strategy used for the flow cytometry data analysis.* ***B.*** *Spearman correlation between fraction of activated Tregs and the change in tumor size on MRI (%).* ***C.*** *Spearman correlation between fraction of non-activated Tregs and the change in tumor size on MRI (%). Activated Tregs were defined as activated by the expression of CD137.* ***D–E.*** *Fold change in activated (**D**) and non-activated (**E**) Tregs after anti-PD-1 or anti-PD-1/anti-CTLA4 therapy. n = 22 patients. In **B-C**, correlation was estimated with Spearman's rank correlation coefficient, two-sided, with 95% confidence interval for the regression estimate. In **D–E**, boxplots display a minimum (Q0), a maximum (Q4), a median (Q2) and the interquartile range. P values were derived using a two-sided Mann–Whitney test.*

***Extended Data Table 1. Full lis**t of adverse events.*

	A: Nivolumab (N=16)			B: Nivo+Ipi 4 wks (N=15)			Nivo+Ipi 6 wks (N=15)		
Adverse event	Grade 1-2	Grade 3	Grade 4	Grade 1-2	Grade 3	Grade 4	Grade 1-2	Grade 3	Grade 4
Worst grade adverse event per patient	12 (75%)	1 (6%)	0 (0%)	13 (86.6%)	1 (7%)	1 (7%)	9 (60%)	5 (33%)	1 (7%)
Endocrine disorders									
Thyroid dysfunction	7 (44%)	1 (6%)[a]	0 (0%)	9 (60%)	0 (0%)	0 (0%)	9 (60%)	0 (0%)	0 (0%)
Hypothyroidism	6 (38%)	0 (0%)	0 (0%)	7 (47%)	0 (0%)	0 (0%)	6 (40%)	0 (0%)	0 (0%)
Hyperthyroidism	4 (25%)	1 (6%)	0 (0%)	8 (53%)	0 (0%)	0 (0%)	9 (60%)	0 (0%)	0 (0%)
Adrenal insufficiency*	1 (6%)	0 (0%)	0 (0%)	1 (7%)	1 (7%)[b]	0 (0%)	2 (13%)	1 (7%)[f]	0 (0%)
Diabetes Mellitus (Immune mediated)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)[b]	0 (0%)	0 (0%)	0 (0%)
Gastro-intestinal									
Abdominal pain	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Diarrhea	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Nausea	1 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (13%)	0 (0%)	0 (0%)
Colitis	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)[g]	0 (0%)
Rectal Ulcer	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Laboratory test									
Elevated liver function tests	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	7 (47%)	0 (0%)	0 (0%)
Hyperphosphatemia	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Hypophosphatemia	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Lymphocyte count decreased	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Troponine T increased	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Immune related hepatitis	0 (0%)	0 (0%)	0 (0%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	2(13%)[c,d]	1 (7%)[h]
Musculo-skeletal									
Arthralgia	3 (19%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Myalgia	2 (12%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Back pain	1 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Immune mediated polymyalgia rheumatica	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Cardiopulmonal									
Chest-pain	0 (0%)	0 (0%)	1 (6%)[a]	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Ejection fraction decreased #	1 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Pneumonitis	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	1 (7%)[e]	0 (0%)
Upper respiratory infection	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Cough	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Other									
Allergic reaction	1 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Anemia	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (20%)	0 (0%)	0 (0%)
Dry eye	1 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Dry mouth	2 (12%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (27%)	0 (0%)	0 (0%)
Dry skin	1 (6%)	0 (0%)	0 (0%)	2 (13%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Fatigue	1 (6%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Flu like symptoms	2 (12%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Headache	2 (12%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Infusion related reaction	2 (12%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Skin rash	2 (12%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	4 (27%)	0 (0%)	0 (0%)
Photosensitivity	1 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Itching	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
Cervical lymphadenopathy	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)

Extended Data Table 2. Antibody overview

Human flow cytometry antibodies					
Antigen	**Fluorochrome**	**Clone**	**Dilution**	**Company**	**Catalogue number**
CD3	PE Cy5	UCHT1	1:200	BD Bioscience	555334
CD4	BV421	RPA-T4	1:100	BD Bioscience	562424
CD8	BUV805	SK1	1:200	BD Bioscience	612754
Pan γδ TCR	PE	11F2	1:100	BD Bioscience	555717
vδ1	FITC	TS8.2	1:100	Thermofisher	TCR2730
vδ2	BUV395	B6	1:100	BD Bioscience	748582
FoxP3	PE Cy5.5	FJK-16s	1:50	Thermofisher	35-5773-82
CCR7	APC R700	150503	1:50	BD Bioscience	565868
CD45RA	BUV737	HI100	1:400	BD Bioscience	612846
CD25	AF647	BC96	1:100	BioLegend	302618
PD-1	APC Cy7	EH12.2H7	1:100	BioLegend	329922
CTLA-4	PE CF594	BNI3	1:200	BD Bioscience	562742
IL-17	PerCP Cy5.5	N49-653	1:50	BD Bioscience	560799
IFNγ	BV785	4S.B3	1:200	BioLegend	502542
TNFα	PE Cy7	Mab11	1:400	BioLegend	502930
CD27	BV786	L128	1:100	BD Bioscience	563327
TIGIT	PerCP Cy5.5	A151536	1:100	BioLegend	372718
Ki-67	PE Cy7	B56	1:50	BD Bioscience	561283
CTLA-4	PE CF594	PE/Dazzle594	1:200	BioLegend	369616

References

1. Schmid, P. *et al.* Event-free Survival with Pembrolizumab in Early Triple-Negative Breast Cancer. *N. Engl. J. Med.* **386**, 556–567 (2022).
2. Loibl, S. *et al.* A randomised phase II study investigating durvalumab in addition to an anthracycline taxane-based neoadjuvant therapy in early triple-negative breast cancer: clinical results and biomarker analysis of GeparNuevo study. *Ann. Oncol.* **30**, 1279–1288 (2019).
3. Mittendorf, E. A. *et al.* Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 trial. *Lancet* **396**, 1090–1100 (2020).
4. Schmid, P. *et al.* Pembrolizumab for Early Triple-Negative Breast Cancer. *N. Engl. J. Med.* **382**, 810–821 (2020).
5. Gustafson, C. E. *et al.* Immune cell repertoires in breast cancer patients after adjuvant chemotherapy. *JCI Insight* **5**, (2020).
6. Mariniello, A. *et al.* Platinum-based chemotherapy attenuates the effector response of CD8 T cells to concomitant PD-1 blockade. *Clin. Cancer Res.* (2023) doi:10.1158/1078-0432.CCR-23-1316.
7. Blank, C. U. *et al.* Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. *Nat. Med.* **24**, 1655–1661 (2018).
8. Cascone, T. *et al.* Neoadjuvant nivolumab or nivolumab plus ipilimumab in operable non-small cell lung cancer: the phase 2 randomized NEOSTAR trial. *Nat. Med.* **27**, 504–514 (2021).
9. Vos, J. L. *et al.* Neoadjuvant immunotherapy with nivolumab and ipilimumab induces major pathological responses in patients with head and neck squamous cell carcinoma. *Nat. Commun.* **12**, 7348 (2021).
10. van Dijk, N. *et al.* Preoperative ipilimumab plus nivolumab in locoregionally advanced urothelial cancer: the NABUCCO trial. *Nat. Med.* **26**, 1839–1844 (2020).
11. Chalabi, M. *et al.* Neoadjuvant immunotherapy in locally advanced mismatch repair-deficient colon cancer. *N. Engl. J. Med.* **390**, 1949–1958 (2024).
12. Bianchini, G., De Angelis, C., Licata, L. & Gianni, L. Treatment landscape of triple-negative breast cancer - expanded options, evolving needs. *Nat. Rev. Clin. Oncol.* **19**, 91–113 (2022).
13. Gianni, L., Huang, C., Egle, D. & Bermejo, B. … or without atezolizumab followed by an adjuvant anthracycline regimen in high-risk triple negative breast cancer (TNBC): NeoTRIP Michelangelo randomized study. *Annals of* (2023).
14. 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).
15. Robert, C. *et al.* Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N. Engl. J. Med.* **364**, 2517–2526 (2011).
16. Hodi, F. S. *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010).
17. Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat. Med.* **26**, 566–576 (2020).
18. Adams, S. *et al.* A Multicenter Phase II Trial of Ipilimumab and Nivolumab in Unresectable or Metastatic Metaplastic Breast Cancer: Cohort 36 of Dual Anti-CTLA-4 and Anti-PD-1 Blockade in Rare Tumors (DART, SWOG S1609). *Clin. Cancer Res.* **28**, 271–278 (2022).
19. de Jong, V. M. T. *et al.* Prognostic Value of Stromal Tumor-Infiltrating Lymphocytes in Young, Node-Negative, Triple-Negative Breast Cancer Patients Who Did Not Receive (neo)Adjuvant Systemic Therapy. *J. Clin. Oncol.* **40**, 2361–2374 (2022).
20. Loi, S. *et al.* Tumor-Infiltrating Lymphocytes and Prognosis: A Pooled Individual Patient Analysis of Early-Stage Triple-Negative Breast Cancers. *J. Clin. Oncol.* **37**, 559–569 (2019).
21. Denkert, C. *et al.* Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol.* **19**, 40–50 (2018).
22. Salgado, R. *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann. Oncol.* **26**, 259–271 (2015).

23. Park, J. H. *et al.* Prognostic value of tumor-infiltrating lymphocytes in patients with early-stage triple-negative breast cancers (TNBC) who did not receive adjuvant chemotherapy. *Ann. Oncol.* **30**, 1941–1949 (2019).
24. Loi, S. *et al.* Association Between Biomarkers and Clinical Outcomes of Pembrolizumab Monotherapy in Patients With Metastatic Triple-Negative Breast Cancer: KEYNOTE-086 Exploratory Analysis. *JCO Precis Oncol* **7**, e2200317 (2023).
25. Voorwerk, L. *et al.* Publisher Correction: Immune induction strategies in metastatic triple-negative breast cancer to enhance the sensitivity to PD-1 blockade: the TONIC trial. *Nat. Med.* **25**, 1175 (2019).
26. Blank, C. U. *et al.* Neoadjuvant Nivolumab and Ipilimumab in Resectable Stage III Melanoma. *N. Engl. J. Med.* (2024) doi:10.1056/NEJMoa2402604.
27. Simon, R. Optimal two-stage designs for phase II clinical trials. *Control. Clin. Trials* **10**, 1–10 (1989).
28. Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).
29. Higgs, B. W. *et al.* Interferon Gamma Messenger RNA Signature in Tumor Biopsies Predicts Outcomes in Patients with Non-Small Cell Lung Carcinoma or Urothelial Cancer Treated with Durvalumab. *Clin. Cancer Res.* **24**, 3857–3866 (2018).
30. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
31. Menzies, A. M. *et al.* Pathological response and survival with neoadjuvant therapy in melanoma: a pooled analysis from the International Neoadjuvant Melanoma Consortium (INMC). *Nat. Med.* **27**, 301–309 (2021).
32. Verschoor, Y. L. *et al.* Neoadjuvant atezolizumab plus chemotherapy in gastric and gastroesophageal junction adenocarcinoma: the phase 2 PANDA trial. *Nat. Med.* (2024) doi:10.1038/s41591-023-02758-x.
33. Lehmann, B. D. *et al.* Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One* **11**, e0157368 (2016).
34. Oliveira, G. *et al.* Phenotype, specificity and avidity of antitumour CD8 T cells in melanoma. *Nature* **596**, 119–125 (2021).
35. Lowery, F. J. *et al.* Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. *Science* **375**, 877–884 (2022).
36. van Gulijk, M. *et al.* PD-L1 checkpoint blockade promotes regulatory T cell activity that underlies therapy resistance. *Sci Immunol* **8**, eabn6173 (2023).
37. Geurts, V. C. M. *et al.* Tumor-Infiltrating Lymphocytes in Patients With Stage I Triple-Negative Breast Cancer Untreated With Chemotherapy. *JAMA Oncol* (2024) doi:10.1001/jamaoncol.2024.1917.
38. Unger, J. M. *et al.* Sex Differences in Risk of Severe Adverse Events in Patients Receiving Immunotherapy, Targeted Therapy, or Chemotherapy in Cancer Clinical Trials. *J. Clin. Oncol.* **40**, 1474–1486 (2022).
39. Nguyen, V. P. *et al.* A pilot study of neoadjuvant nivolumab, ipilimumab and intralesional oncolytic virotherapy for HER2-negative breast cancer. *Cancer Res. Commun.* (2023) doi:10.1158/2767-9764.crc-23-0145.
40. Maher, V. E. *et al.* Analysis of the Association Between Adverse Events and Outcome in Patients Receiving a Programmed Death Protein 1 or Programmed Death Ligand 1 Antibody. *J. Clin. Oncol.* **37**, 2730–2737 (2019).
41. 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* **6**, 519–527 (2020).
42. Beaufils, M. *et al.* Dysthyroidism during immune checkpoint inhibitors is associated with improved overall survival in adult cancers: data mining of 1385 electronic patient records. *J Immunother Cancer* **11**, (2023).
43. Street, S. *et al.* The positive effect of immune checkpoint inhibitor-induced thyroiditis on overall survival accounting for immortal time bias: a retrospective cohort study of 6596 patients. *Ann. Oncol.* **32**, 1050–1051 (2021).
44. Groha, S. *et al.* Germline variants associated with toxicity to immune checkpoint blockade. *Nat. Med.* **28**, 2584–2591 (2022).
45. Reijers, I. L. M. *et al.* IFN-γ signature enables selection of neoadjuvant treatment in patients with stage III melanoma. *J. Exp. Med.* **220**, (2023).

46. Bullock, A. J. *et al.* Botensilimab plus balstilimab in relapsed/refractory microsatellite stable metastatic colorectal cancer: a phase 1 trial. *Nat. Med.* (2024) doi:10.1038/s41591-024-03083-7.
47. Blomberg, O. S. *et al.* Neoadjuvant immune checkpoint blockade triggers persistent and systemic T activation which blunts therapeutic efficacy against metastatic spread of breast tumors. *Oncoimmunology* **12**, 2201147 (2023).
48. Simpson, T. R. *et al.* Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. *J. Exp. Med.* **210**, 1695–1710 (2013).
49. Sharma, A. *et al.* Anti-CTLA-4 Immunotherapy Does Not Deplete FOXP3+ Regulatory T Cells (Tregs) in Human Cancers. *Clin. Cancer Res.* **25**, 1233–1238 (2019).
50. van der Leun, A. M. *et al.* Dual immune checkpoint blockade induces analogous alterations in the dysfunctional CD8+ T cell and activated Treg compartment. *Cancer Discov.* (2023) doi:10.1158/2159-8290.CD-22-0851.
51. Loibl, S. *et al.* Neoadjuvant durvalumab improves survival in early triple-negative breast cancer independent of pathological complete response. *Ann. Oncol.* **33**, 1149–1158 (2022).
52. Pathologic complete response (pCR) to neoadjuvant treatment with or without atezolizumab in triple-negative, early high-risk and locally advanced breast cancer: NeoTRIP Michelangelo randomized study. *Ann. Oncol.* **33**, 534–543 (2022).
53. Chalabi, M., Verschoor, Y. L. & Van den Berg, J. LBA7 Neoadjuvant immune checkpoint inhibition in locally advanced MMR-deficient colon cancer: The NICHE-2 study. *Annals of* (2022).
54. Rozeman, E. A. *et al.* Survival and biomarker analyses from the OpACIN-neo and OpACIN neoadjuvant immunotherapy trials in stage III melanoma. *Nat. Med.* **27**, 256–263 (2021).
55. Pusztai, L. *et al.* Event-free survival by residual cancer burden with pembrolizumab in early-stage TNBC: exploratory analysis from KEYNOTE-522. *Ann. Oncol.* **35**, 429–436 (2024).
56. Magbanua, M. J. M. *et al.* Circulating tumor DNA in neoadjuvant-treated breast cancer reflects response and survival. *Ann. Oncol.* **32**, 229–239 (2021).
57. Ayers, M. *et al.* IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017).
58. Garg, A. D., De Ruysscher, D. & Agostinis, P. Immunological metagene signatures derived from immunogenic cancer cell death associate with improved survival of patients with lung, breast or ovarian malignancies: A large-scale meta-analysis. *Oncoimmunology* **5**, e1069938 (2016).
59. Hu, X. E. *et al.* Clinical and biological heterogeneities in triple-negative breast cancer reveals a non-negligible role of HER2-low. *Breast Cancer Res.* **25**, 34 (2023).
60. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845–865.e7 (2021).
61. Gangaev, A. *et al.* Identification and characterization of a SARS-CoV-2 specific CD8+ T cell response with immunodominant features. *Nat. Commun.* **12**, 1–14 (2021).
62. Cabrita, R. *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* **577**, 561–565 (2020).
63. Oken, M. M. *et al.* Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am. J. Clin. Oncol.* **5**, 649–655 (1982).
64. Litton, J. K. *et al.* Standardized Definitions for Efficacy End Points in Neoadjuvant Breast Cancer Clinical Trials: NeoSTEEP. *J. Clin. Oncol.* **41**, 4433–4442 (2023).
65. Blakely, C. M. & McCoach, C. E. Role of MPR as an Early Signal for Efficacy in Neoadjuvant Studies. *Clinical cancer research: an official journal of the American Association for Cancer Research* vol. 26 3499–3500 (2020).
66. Cascone, T. *et al.* A Phase I/II Study of Neoadjuvant Cisplatin, Docetaxel, and Nintedanib for Resectable Non-Small Cell Lung Cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research* vol. 26 3525–3536 (2020).
67. [No title]. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae_v5_quick_reference_5x7.pdf.
68. Website. R Core Team. 2022. 'R: A Language and Environment for Statistical Computing.' Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
69. slidescore.com
70. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
71. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).

72. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
73. Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res.* **7**, 1338 (2018).
74. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
75. Picard. https://broadinstitute.github.io/picard.
76. Chen, X. *et al.* TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer. *Cancer Inform.* **11**, 147–156 (2012).
77. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
78. Fang, Z., Liu, X. & Peltz, G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**, btac757 (2022).
79. van Rossum, G. *Python Reference Manual*. (1995).
80. pandas-dev/pandas: Pandas. (2023) doi:10.5281/zenodo.8239932.
81. McKinney, W. Data Structures for Statistical Computing in Python. in *Proceedings of the 9th Python in Science Conference* (SciPy, 2010). doi:10.25080/majora-92bf1922-00a.
82. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
83. Waskom, M. *et al. Mwaskom/seaborn: v0.8.1 (September 2017)*. (Zenodo, 2017). doi:10.5281/ZENODO.883859.
84. Charlier, F. *et al.* trevismd/statannotations: v0.5. (2022) doi:10.5281/zenodo.7213391.
85. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
86. van der Auwera, G. & O'Connor, B. D. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. (O'Reilly Media, Incorporated, 2020).
87. Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C. & Koeffler, H. P. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **28**, 1747–1756 (2018).
88. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at https://www.R-project.org/ (2022).
89. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 1–5 (2018).
90. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* 1–12 (2023).
91. Fleming, S. J. *et al.* Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* 1–13 (2023).
92. Bassez, A. *et al.* A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat. Med.* **27**, 820–832 (2021).
93. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184–197 (2015).
94. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
95. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–7 (2016).
96. Sturm, G. *et al.* Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics* **36**, 4817–4818 (2020).
97. GitHub - schillerlab/sc-toolbox: A collection of project templates and useful functions for single-cell data analysis with Scanpy. *GitHub* https://github.com/schillerlab/sc-toolbox.
98. Blomberg, O. S. *et al.* IL-5-producing CD4 T cells and eosinophils cooperate to enhance response to immune checkpoint blockade in breast cancer. *Cancer Cell* **41**, 106–123.e10 (2023)