nature portfolio

Corresponding author(s): Marleen Kok

Last updated by author(s): *YYYY-MM-DD*

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	For DNA sequencing, sequencing reads were aligned to the human reference GRCh38 (Ensemble, v. 105) using BWA90 0.7.17. Duplicated reads were marked using Picard MarkDuplicates 2.25.0, after which quality scores were recalibrated using GATK4 BaseRecalibrator 4.2.2.0. Single-nucleotide variants (SNVs) and short insertions and deletions (indels), were called using GATK4 Mutect2 4.2.2.0 on the tumor samples matched with germ line samples. Subsequently, variants were filtered by the PASS filter, and annotated using Ensembl Variant Effect Predictor. maftools 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 3.10.5 and R 4.1.3. Pandas 2.0.0 was used for data handling. maftools 2.10.5, Matplotlib 3.5.2, seaborn 0.12.2, and statannotations 0.5.0 were used for plotting. For ctDNA analysis, a proprietary bioinformatics tissue variant calling pipeline (Natera Inc) 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. 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. For single-cell RNA-Seq, 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 Cell Ranger (I0x Genomics, v2.0). The data were analyzed with scanpy 1.9.3 and Seurat v3. Cellbender 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. 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 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 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. PDl +Ki67+CD4+ cells were defined based on the level of MKl67 gene expression >0 in the Tfh cell population. Scirpyl 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 3.10.5. Pandas 2.0.0 and numpy 1.22.4 were used for data handling. Matplotlib 3.5.2, seaborn 0.12.2, sc-toolboxlOl 0.12.3 and statannotations 0.5.0 were used for plotting.
For bulk RNA-Seq, data were aligned to GRCh38 with STAR 2.7.1a, with the twopassMode='Basic'. FPKM were obtained with RSeQC 4.0.0 FPKM_count.py and subsequently normalized to transcripts per million. Data quality was assessed with FastQC 0.11.5, FastQ Screen 0.14.0, the Picard CollectRnaSeqMetrics and RSeQC 4.0.0 read_distribution.py and read_duplication.py and were found to be suitable for the downstream analysis. TNBCtype was used for the Lehmann subtype classification. The Gseapy 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 3.10.5. Pandas 2.0.0 and numpy 1.22.4 were used for data handling. Matplotlib 3.5.2, seaborn 0.12.2 and statannotations 0.5.0 were used for plotting.

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

Data

Policy information about availability of data

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

DNA and RNA sequencing data will be stored in the European Genome–Phenome Archive upon publication. 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 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.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	All participants are of female sex, by DNA and by identification
Reporting on race, ethnicity, or other socially relevant groupings	NA, we do not report on race, ethnicity or socially relevant groupings.
Population characteristics	Patients in cohorts A and B were eligible for enrollment if they were at least 18 years of age and had stage 1-111 (clinical tumor stage Tlc-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<l0% 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 score68 of 0 or 1 and adequate organ functions. The Tl Ls percentage needed to be 5% or more. 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 NO, tumor stage Tlc-T2, and Tlls had to be 50% or more. 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.
Recruitment	The BELLINI trial (full title: Pre-operative Trial for Breast Cancer With Nivolumab in Combination With Novel 10; 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. Patients were asked to be screened for BELLINI at the routine diagnostic visits at the treatment center. Patients willing to participate may have had motivation to de-escalate chemotherapy, even though this was not the primary endpoint of the trial. We do not know if this could affect the outcomes.
Ethics oversight	All patients provided written informed consent before enrollment. This investigator-initiated trial was designed by the

Ethics oversight	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).

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

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	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 R73 v.4.2.1.
Data exclusions	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 score68 of 0 or 1 and adequate organ functions. The TILs percentage needed to be 5% or more. 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. 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. No patients were excluded from analysis after registration for the trial.
Replication	The current study was not attempted to be replicated as this is a human phase 2 clinical trial, stage I. However, the most important findings will be tested in the second stage of the trial.
Randomization	Patients were not randomized in this phase II non-randomized trial.
Blinding	Since the BELLINI is the first trial testing apD1/aCTLA-4 in early stage TNBC, physicians needed to know what patients received in order to ensure safety. Pathologists were blinded for clinical outcome and treatment group upon scoring of the biopsies.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used

Human flow cytometry antibodies
Antigen Fluorochrome Clone Dilution Company Catalog 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

Immunohistochemistry
PDL1 clone 22C3 (1/40 dilution, 1 hour at RT, Agilent/DAKO, Lot11654144)
PD1 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) followed by Anti-NP AP (Ventana Medical systems, Ready to Use dispenser, LotJ23971)

CD8 clone C8/144B (1/200 dilution, 32 minutes at 37C, Agilent, Lot41527763).
PD-L1 (22C3, DAKO) and PD-1 (NAT105, Roche Diagnostics). CD8 was visualized using Anti-Mouse HQ (Ventana Medical systems, Ready to Use dispenser, LotK20711), followed by Anti-HQ HRP (Ventana Medical systems, Ready to Use dispenser, LotK22062)

Validation

All flow cytometry antibodies are commercially available and described in extended data table 2, including company, dilution and catalogue number for further information and background on each antibody.
For immunohistochemistry: CD8 and PD-L1 are diagnostic markers and were validated on human diagnostic tissue by the local pathology department. PD-1 has been validated for research purposes and validated on human tonsil and appendix tissue by the local pathology department. All antibodies used are commercially available and validated by manufacturer. Antibodies for flow cytometry were further validated for target species (human) using FMO or isotype controls where necessary. All antibodies were titrated to identify optimal staining concentration

Clinical data

Policy information about clinical studies

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

Clinical trial registration

NCT03815890

Study protocol

The full trial protocol can be requested for academic purposes at IRB of the Netherlands Cancer Institute and the corresponding author (Marleen Kok)

Data collection

All patients were enrolled in the Netherlands Cancer Institute, Amsterdam. The first patient was enrolled on the 19th of September 2019 and the last on the 24th of January 2023. The study is still recruiting new patients for new cohorts.

Outcomes

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)
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)69. Major pathologic response (MPR, secondary endpoint) is a frequently used surrogate endpoint for efficacy in neoadjuvant trials evaluating immune checkpoint blockade across cancer types9,12,27. MPR was defined as ≤10% of residual viable tumor in the surgical specimen 18,70,71 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.572. 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.

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

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).

Instrument

Data acquisition was performed on an LSRII SORP flow cytometer (BD Biosciences)

Software

Diva software and data analysis was performed using FlowJo 10.6.2. Gating strategy is displayed in Extended Data Fig. 5A.

Cell population abundance

We report the percentage of a proliferating T cells within a pre-specified T-cell population in figure 3G and H; it;s therefore a relative measure and not an absolute count.

Gating strategy

Gating strategy is displayed in extended Data fig. 5A
) T cell panel gating strategy to identify proliferating T cells
T cells (CD3+,) vd1 negative and vd2 negative T cells
(CD3+, vd1-vd2+)
CD8 T cells (CD3+, vd1-vdd2-, CD8+, CD4-) --> next, Ki67High for proliferating fraction of CD8 T cells (CD3+, vd1-vdd2-, CD8+, CD4-, Ki67High) --> Last, PD1 positive fraction (CD3+, vd1-vdd2-, CD8+, CD4-, Ki67High, PD1High)
conventional CD4 T cells (CD3+, vd1-vd2-, CD8-, CD4+, FoxP3-) --> Next, Ki67High for proliferating fraction of CD4 T cells (CD3 +, vd1-vd2-, CD8-, CD4+, FoxP3-, Ki67high) --> last, PD1 positive fraction (CD3+, vd1-vd2-, CD8-, CD4+, FoxP3-, Ki67high,

PD1high)

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.