

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. [For final submission](#): please carefully check your responses for accuracy; you will not be able to make changes later.

► Experimental design

1. Sample size

Describe how sample size was determined.

Response rate to docetaxel was expected to be 20-30%. The trial was powered to detect superiority of carboplatin over docetaxel with an improvement in response rates of 15%. With 90% power and 2-sided alpha of 0.05, at least 370 patients were required.

2. Data exclusions

Describe any data exclusions.

No patients were excluded from the analyses. For a priori specified biomarker subgroup analyses only those with known biomarker status were included.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

No experimental replication was attempted - N/A in a phase III randomised controlled clinical trial

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

ICR-CTSU allocated patients to carboplatin or docetaxel (1:1 ratio) utilising a computerised minimisation algorithm with a random element. Balancing factors were centre, previous adjuvant taxane chemotherapy, presence of liver or lung metastasis, performance status (0/1 vs 2) and recurrent locally advanced vs metastatic carcinoma.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Treatment allocation was not blinded; this would have been impractical due to different side effect profiles and administration of the drugs. However those analysing samples to generate biomarker data for a priori planned biomarker subgroup analyses were blinded to treatment and response.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- ☐ ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- ☐ ☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☒ ☐ A statement indicating how many times each experiment was replicated
- ☐ ☒ 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 any assumptions or corrections, such as an adjustment for multiple comparisons
- ☐ ☒ Test values indicating whether an effect is present
*Provide confidence intervals or give results of significance tests (e.g. *P* values) as exact values whenever appropriate and with effect sizes noted.*
- ☐ ☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- ☐ ☒ Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

Statistical analysis was carried out in Stata v13.
Other software used for biological analyses were:
CASA 1.8 <http://www.illumina.com/>
MapSplice v12_07 <http://www.netlab.uky.edu/p/bioinfo/MapSplice>
Picard Tools v1.64 - <http://picard.sourceforge.net/>
UBU v1.0 <https://github.com/mozack/ubu>
RSEM v1.1.13 - <https://deweylab.github.io/RSEM/>
TCGA GAF - <https://api.gdc.cancer.gov/v0/data/a0bb9765-3f03-485b-839d-7dce4a9bcfeb>

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

The trial associated tissue biological resource is accessible through application to the TNT Trial Management Group Biological sub-committee where proposed research is covered by protocol and the TNT trial ethical permission and patient consent.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

ER: anti-Estrogen Receptor (ER) (SP1) Rabbit Monoclonal Ventana 790-4324/4325 antibody
PgR: anti-Progesterone Receptor (PR) (1E2) Rabbit Monoclonal Ventana 790-2223/4296 Ab
HER2: anti-HER-2/neu (4B5) Rabbit Monoclonal Ventana 790-2991 (05278368001) antibody
Ck5: Ck5 (XM26) Mouse monoclonal Novocastra NCL-CK5 antibody
EGFR: EGFR.25) Mouse monoclonal Novocastra NCL-L-EGFR-384 antibody

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used - NA

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used - NA

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used - NA

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No eukaryotic cell lines were used - NA

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

No animals were used.

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

All patients were female, aged between 26 and 81. Patients had triple negative metastatic or recurrent locally advanced breast cancer or patients had a known germline BRCA1/2 mutation and had metastatic or recurrent locally advanced breast cancer.