nature portfolio

Corresponding author(s): Aziz Zaanan

Last updated by author(s): May 6, 2025

# 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<br>*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<br>*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 | N/A |
| Data analysis | Patient characteristics were summarized using descriptive statistics, and statistical significance was evaluated using Fisher's exact test or Chi-square test for categorical variables and Wilcoxon rank sum two-sided test for continuous variables, as indicated in corresponding figure legends. Survival analyses were conducted using R software v4.4.0 using packages survminer (v0.4.9) and survival (v3.2.13). Survival curves were compared using the Kaplan-Meier method. Hazard ratios (HR), associated 95% confidence intervals (CI), and p-values were calculated using Cox regression analysis (R packages survminer v0.4.9, coxphf, and survival v3.2.13). The log-rank test was used to compare two survival distributions. Overall, P values <0.05 were considered statistically significant. To account for potential immortal time bias, analyses for the prognostic value of ctDNA at the MRD Window were landmarked at 12 weeks post-surgery.<br>A multivariable Cox proportional hazards model assessed prognostic factors associated with RFS and OS. The proportional hazard assumption was tested using a global test of the Schoenfeld residuals. The receiver operating characteristic (ROC) analysis was utilized to determine an optimal ctDNA threshold at baseline by maximizing sensitivity and specificity for RFS events. |

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

The authors declare that all relevant data used to conduct the analyses are available within the article. Complete clinical information and ctDNA results are presented in the heat map and overview plot (Supplementary Figure 1). Patient characteristics and demographic information are available in Table 1 and Supplementary Data 1. Supplementary Data 2 includes the de-identified Raw data of patient characteristics, outcomes (RFS and OS), and complete ctDNA results. To protect the privacy and confidentiality of patients in this study, clinical data are not made publicly available in a repository or the supplementary material of the article but may be requested at any time from the corresponding author. Any requests will be reviewed within a time frame of 2 to 3 weeks by the corresponding author. All data shared will be de-identified. The fully documented code for the R statistical computing environment for analyses related to this manuscript is deposited in the GitHub repository and can be accessed at https://github.com/Natera-TMED/Zaanan-et-al_AGEO-PLAGAST.git

## 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 | Information on patient sex was collected. Patient sex was self-reported by study participants. Of the total 62 patients analyzed, 39 (63%) were males and 23 (37%) were females. |
| Reporting on race, ethnicity, or other socially relevant groupings | N/A |
| Population characteristics | Patient characteristics are summarized in Table 1, Supplementary Data 1, and Supplementary Figure 1 |
| Recruitment | Recruitment of eligible patients was done at the European Georges Pompidou Hospital |
| Ethics oversight | This study was conducted per the Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice. All participants provided written informed consent. This study is registered at ClinicalTrials.gov (NCT03079427). |

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 | This is a retrospective evaluation of ctDNA for patients enrolled in the PLAGAST study. The sample size was determined by the available number of patients who had evaluable tissue available for whole exome sequencing and blood samples at prespecified time points as mentioned in study protocol. |
| Data exclusions | Complete detail on exclusions in provided in Figure 1. |
| Replication | Due to sample size and biospecimen availability, ctDNA measurements were performed once on each sample and were not repeated. Note that the ctDNA assay run on the patient samples has been analytically validated previously. |
| Randomization | This is an observational study and is not a randomized trial. |
| Blinding | The ctDNA measurements were conducted by Natera in a manner that kept them blinded to clinical data including RFS and OS. All treatment and clinical management decisions were made blinded to ctDNA results. All clinical outcome data were collected by individuals blinded to the ctDNA results. |

# 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 |

# 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 | This study is registered at ClinicalTrials.gov (NCT03079427). |
| Study protocol | The study protocol is submitted along with the manuscript. |
| Data collection | Plasma samples (n=280) were collected between 07/2015 and 12/2022 from 62 patients with LAR G/GEJ adenocarcinoma at baseline (anytime post-diagnostic biopsy/enrollment and before NAT or before surgery for NAT-naive), during-NAT, post-NAT, and post-surgery at the molecular residual disease (MRD) window (within 2-12 weeks, before adjuvant treatment). |
| Outcomes | The primary endpoint was recurrence-free survival (RFS), defined as the time from surgery until radiological or clinical recurrence. Recurrence was determined based on diagnostic imaging or any other diagnostic procedure if imaging was not confirmative (e.g., gastroscopy to diagnose local recurrence). The secondary endpoints were OS, tumor regression score (TRG), and pathological tumor stage. OS was defined as the time between the date of surgery and the date of death due to any cause or latest clinical follow-up. TRG was defined according to the Mandard classification.15 To account for potential immortal time bias, analyses for the prognostic value of ctDNA at the MRD Window were landmarked at 12 weeks post-surgery. Patient characteristics were summarized using descriptive statistics, and statistical significance was evaluated using Fisher's exact test or Chi-square test for categorical variables and Wilcoxon rank sum two-sided test for continuous variables, as indicated in corresponding figure legends. Survival analyses were conducted using R software v4.4.0 using packages survminer (v0.4.9) and survival (v3.2.13). Survival curves were compared using the Kaplan-Meier method. Hazard ratios (HR), associated 95% confidence intervals (CI), and p-values were calculated using Cox regression analysis (R packages survminer v0.4.9, coxphf, and survival v3.2.13). The log-rank test was used to compare two survival distributions. Overall, P values <0.05 were considered statistically significant. A multivariable Cox proportional hazards model assessed prognostic factors associated with RFS and OS. The proportional hazard assumption was tested using a global test of the Schoenfeld residuals. The receiver operating characteristic (ROC) analysis was utilized to determine an optimal ctDNA threshold at baseline by maximizing sensitivity and specificity for RFS events. |

# Plants

| | |
|---|---|
| Seed stocks | N/A |
| Novel plant genotypes | N/A |
| Authentication | N/A |