nature portfolio

Corresponding author(s): Eiji Oki

Last updated by author(s): Aug 15, 2024

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	All data collection was done by input into an EDC systems; the EDC was TrialMaster version 5.0 (update 6) from Anju LifeSciences software.
Data analysis	Survival analyses were carried out using R software v4.4.0 using packages survival (v.3.7.0), survminer (v.0.4.9) and coxphf (v.1.13.4). The exact P values from Cox regression hazard models were calculated using package Rmpfr (v.0.9.5). The Kaplan–Meier method was used to estimate the survival distribution. Differences between the groups were tested using the log-rank test. A multivariable Cox proportional hazards model was used to assess prognostic factors associated with DFS.

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, non-proprietary data used to conduct the analyses are available within the article. Supplementary Table 2 includes the

deidentified Raw data of patient characteristics, outcomes (DFS and OS), complete ctDNA results (at the MRD window, 3- and 6-months timepoints and surveillance windows, ctDNA clearance post MRD window, and molecular recurrence). To protect the privacy and confidentiality of patients in this study under the Japanese Act on the Protection of Personal Information, identifiable clinical data are not made publicly available in a repository or the supplementary material of the article but can be requested at any time from the corresponding author [Takayuki Yoshino (tyoshino@east.ncc.go.jp) or Eiji Oki (oki.eiji.857@m.kyushu-u.ac.jp)]. Any requests will be reviewed within a time frame of 2 to 3 weeks by the CIRCULATE-Japan study steering committee to verify whether the request is subject to any intellectual property or confidentiality obligations. All data shared will be de-identified and will be provided to researchers with access limited for scientific verification purposes and strict prohibitions on secondary use.

The fully documented code and raw outputs for the R statistical computing environment for analyses related to this manuscript are deposited at the github repository and can be accessed at https://github.com/Natera-TMED/Nakamura-et-al_CIRCULATE-JP-Galaxy-2024.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 2,240 patients analyzed, 1,149 (51%) were males and 1,091 (49%) were females.
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	Patient characteristics are summarized in Fig. 1b. Extended Data Tables 1a and 1b describe characteristics for patients with stage I-III and stage IV disease, respectively, and Extended Data Table 1c lists organs involved in metastasis for stage IV patients.
Recruitment	All registration was done on a voluntary basis by researchers at each participating site. The possibility of bias in the selection by researchers cannot be ruled out. However, since enrollment was prospective and each site enrolled most of the cases that fit the criteria, we believe that bias was suppressed as much as possible.
Ethics oversight	All patients provided written informed consent before participation in the study. The clinical protocol was approved by the Institutional Review Board of the National Cancer Center Japan and authorized by the head of each participating institution. The GALAXY study has been registered in the Japan Registry of Clinical Trials (UMIN000039205). The study was conducted in accordance with the Declaration of Helsinki.

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 an updated analysis of the planned GALAXY study to evaluate overall survival. The data cutoff for this analysis was performed to allow for sufficient follow-up to evaluate the impact of circulating tumor DNA-based molecular residual disease (MRD) status on the overall survival. No other sample size calculation was performed for ctDNA analysis. 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 1B. Specifically, Patients with stage I-III rectal cancer were excluded from the analysis, as the clinical management of these patients differs significantly from those with stage I-III colon cancer. Additionally, patients enrolled in associated interventional phase III trials (ALTAIR and VEGA), and patients whose sample(s) failed quality control metrics were excluded.
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 DFS, OS and objective response. 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	The study has been registered in the University hospital Medical Information Network (UMIN000039205).
Study protocol	The study protocol is submitted along with the manuscript.
Data collection	Between May 2020 and February 2024, a total of 6,061 patients with surgically resectable, stage I-IV CRC were prospectively enrolled in the GALAXY study at 92 institutions. Of these 6,061 patients, 2,240 patients with stage IV rectal or any stage colon cancer were included in this analysis. Blood samples were collected at 4, 12, 24, 36, 48, 72, and 96 weeks after surgery until recurrence. Computed tomography (CT) imaging was performed every 6months after surgery. All data collection was done by input into an EDC systems; the EDC was TrialMaster version 5.0 (update 6) from Anju LifeSciences software.
Outcomes	A statistical analysis plan was developed prior to data analysis. The primary endpoints were DFS and OS. DFS was defined as the time between the date of landmark and the date of diagnosis with recurrence or death due to any cause or the latest radiological assessment. For DFS, an event was defined as recurrence or death due to any cause. Recurrence was determined based on diagnostic imaging or any other diagnostic procedure if imaging was not confirmative (e.g. colonoscopy to diagnose local recurrence). OS was defined as the time between the date of landmark and the date of death due to any cause or latest clinical follow-up. The chi-squared test was used to compare categorical variables. Survival analyses were carried out using R software v4.4.0 using packages survival (v.3.7.0), survminer (v.0.4.9) and coxphf (v.1.13.4). The exact P values from Cox regression hazard models were calculated using package Rmpfr (v.0.9.5). The Kaplan–Meier method was used to estimate the survival distribution. Differences between the groups were tested using the log-rank test. A multivariable Cox proportional hazards model was used to assess prognostic factors associated with DFS. Clinically relevant cutoffs were applied for demographic variables, wherever appropriate. MRD analyses were landmarked at the date of MRD time point to account for immortal time bias. Surveillance analyses were landmarked at 10 weeks post-surgery, since patients with CRC typically receive an adjuvant therapy around 8-10 weeks post-surgery per current clinical guidelines. Thus 10-week landmark ensures that both the ACT-treated and non ACT-treated (Observation) populations are alive and free of clinical events independent of whether they had ctDNA time point and adjuvant treatment, The secondary endpoint was ctDNA clearance after ACT, and the exploratory endpoint was molecular recurrence analysis (conversion of ctDNA status to positive post MRD window in ctDNA MRD-negative patients) and Cox regression was used to compare the survival benefit between the ACT and observation groups, was landmarked at 2 months post-surgery, and was adjusted for age, sex, stage, performance status, and MSI status. ctDNA clearance at 3 months and at 6 months were defined as ctDNA clearance from MRD time point to the 3 months (70–112 days post-surgery) and 6 months (160–200 days post-surgery) time point, respectively, in patients treated with ACT. The definition of clearance at 6 months was independent of the clearance status at 3 months. ctDNA clearance analyses were landmarked at 3 months and 6 months, respectively. In patients who were ctDNA positive during the MRD window and received ACT, sustained clearance was defined as ctDNA clearance that was achieved post-MRD and persisted for two or more subsequent ctDNA time points, while transient clearance was defined as ctDNA clearance post-MRD which reverted to ctDNA positive at any subsequent time point. The association of ctDNA-based MRD with recurrence risk according to each actionable biomarker was evaluated by comparing DFS between patients with a specific actionable biomarker who had either positive or negative ctDNA at the MRD timepoint. P values <0.05 were considered statistically significant.

Plants

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