nature research

Corresponding author(s): Ingo K. Mellinghoff
Michael F. Berger

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Sequence data for all CSF samples were generated in the MSK research sequencing core (IGO) in the Center for Molecular Oncology. Tumor sequence data for comparison were generated in the MSK clinical sequencing laboratory in the Department of Pathology and were obtained from the cBioPortal for Cancer Genomics (http://cbioportal.org/msk-impact).
Data analysis	The MSK-IMPACT data analysis pipeline can be found here: https://github.com/rhshah/IMPACT-Pipeline and is composed entirely of open source software. Multivariable survival analysis was performed using the PHREG procedure in SAS (Cary, NC) to implement Cox regression modeling, a standard statistical modeling technique not unique to SAS.

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/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Data Availability: All genomic results and associated clinical data for all patients in this study are publically available in the cBioPortal for Cancer Genomics at the following URL: http://www.cbioportal.org/study?id=glioma_msk_2018

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	No sample size calculation was performed for this retrospective analysis, however all consecutive samples within 1/12/15-4/5/2017 were obtained. The sample size was sufficient since we were able to detect associations between tumor-derived DNA in CSF with both disease burden and adverse outcome.
Data exclusions	The complete patient cohort consisted of 85 patients on whom all analyses were based with the exception of the multivariable Cox regression model. This model was based on a subset of 63 patients and is the only analysis where exclusion criteria were applied. Due to study design, 22 patients were excluded because the original tumor resection had been performed at an outside hospital and MRIs were not available. To ensure there were no differences in the complete patient cohort (n=85) and the subset on whom the multivariable Cox regression model was based (n=63), we compared the two groups on overall survival experience as well as all covariables in the model: CSF positivity, IDH status, percent extent of resection at diagnosis, and tumor burden at LP/CSF. There were no statistically significant differences between the two groups indicating that the subset of patients on whom the Cox regression model was based was comparable to the complete patient cohort.
Replication	For multiple patients in our study, a contemporaneous tumor biopsy or second CSF collection was available for comparisons with the CSF ctDNA sample. This data is included in Figure 2 and shows the reproducibility of our assay.
Randomization	This was not a setting in which randomization was required since there was no intervention being assessed. Samples from all comers during a specific date range were assessed for signal.
Blinding	Conventional blinding was not used in our design since no intervention was being assessed. However, we note in the Methods section that: “Brain MRIs prior to and directly following the initial resection were reviewed by an experienced neuroradiologist without knowledge of the CSF ctDNA results."

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

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

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	The study includes data from tumor samples and cerebro-spinal fluid (CSF) from glioma patients. These unique materials are not available from the authors or from standard commercial sources.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Patients were between 22 and 90 years old. The patient population included 54 men and 31 women. 46 patients had glioblastoma (GBM). 39 patients had lower grade glioma (LGG). 84/85 patients had received prior radiation therapy. 78/85 patients had received therapy with alkylating agents as part of their systemic therapy.
Recruitment	All patients in our study were treated and followed at Memorial Sloan Kettering and underwent lumbar puncture for clinical indications. There are no expected biases based on recruitment since all consecutive samples were obtained during a specified time period. There was no self-selection into the study.

Magnetic resonance imaging

Experimental design

Design type	N/A
Design specifications	N/A
Behavioral performance measures	N/A

Acquisition

Imaging type(s)	Structural
Field strength	1.5T and 3T
Sequence & imaging parameters	MRI included the following parameters: axial contrast-enhanced T1-weighted images with gadolinium contrast (gadobutrol, 0.1 mmol/kg max 14 ml) and field-of-view, 24; slice thickness, 3-5mm; gap, 0; time-to-echo (TE), minimum; time-to-repetition (TR), 400-2000; flip angle, auto; matrix, 256x224; and axial fluid attenuated inversion recovery (FLAIR) images with field-of-view, 24; slice thickness, 3-5mm; gap, 0; TE, 120; TR, 8000-9000; flip angle, 90 degrees; time-to-inversion (TI), 2000-2500; matrix, 256x256.
Area of acquisition	whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	N/A
Normalization	N/A
Normalization template	N/A
Noise and artifact removal	N/A
Volume censoring	N/A

Statistical modeling & inference

Model type and settings	N/A
Effect(s) tested	N/A

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	N/A
Correction	N/A

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis