

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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ 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 <i>d</i> , Pearson's <i>r</i>), 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	Clinical, molecular and laboratory data were extracted from institutional databases using standard query pipelines; no custom software was developed specifically for data collection for this study.
Data analysis	All analyses were performed in R (v4.1.1) and Python(v3.13.12), using publicly available packages (Methods). Code is available at https://github.com/reznik-lab/cachexia . Requirements.txt for both R and Python were added on the github page.

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](#)

Clinical and genomic sequencing data analyzed in this study have been deposited in the cBioPortal for Cancer Genomics and are available for online browsing and download. Additional study data is also available at <https://zenodo.org/records/22045313?preview=1> (2026). Raw sequencing data are restricted to protect patient

privacy in accordance with applicable federal and state laws and institutional policies. Additional de-identified individual-level clinical data may be made available from the corresponding authors upon reasonable request and subject to institutional review and execution of an appropriate data use agreement. Code used for analysis is available at <https://github.com/reznik-lab/cachexia>

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

Sex was recorded in the electronic health record as assigned at birth and extracted from institutional clinical databases; it was not self-reported for the purposes of this study. Sex was included as a covariate in all multivariate analyses, including logistic regression models, mixed-effects models, time-dependent and fixed Cox proportional hazards models, and the time-to-cachexia predictive model. In pan-cancer demographic analyses, male sex was associated with lower odds of meeting the weight-loss episode criterion ($\log_2\text{OR} = -0.23$; Supplementary Table 2). Serologic measurements with established sex-specific reference ranges were adjusted using sex-specific thresholds accordingly (Methods). In our cohort, 27,720 patients (53.1%) were female and 24,516 (46.9%) were male; this breakdown is provided in Supplementary Table 18. Gender identity was not separately assessed, as it was not available in the source clinical data.

Reporting on race, ethnicity, or other socially relevant groupings

NA. Race and ethnicity (including ancestry) were not analyzed in this study and were not included as covariates.

Population characteristics

The study cohort comprised 52,236 adult cancer patients. Median age at diagnosis was 60.6 years (IQR 52.3–70.4). Sex distribution was 27,720 female (53.1%) and 24,516 male (46.9%). Clinical stage was I–III in 27,885 (53.4%), stage IV in 22,021 (42.2%), and missing in 2,330 (4.5%). Baseline BMI categories were underweight (<18.5) 978 (1.9%), normal (18.5–24.9) 17,132 (32.8%), overweight (25–29.9) 18,636 (35.7%), and obese (≥ 30) 15,490 (29.7%). At last follow-up, 31,348 (60.0%) were alive and 20,888 (40.0%) were deceased.

Recruitment

NA

Ethics oversight

The study is independently approved by the IRB of MSK under 12-245 Appendix C.

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

Sample size was determined by the availability of eligible patients in the institutional dataset meeting prespecified inclusion criteria. After applying exclusions, the final cohort comprised 52,236 adult patients (age ≥ 18 years) with a single primary cancer type classification and sufficient longitudinal BMI follow-up to enable episode detection. No a priori power calculation was performed for this retrospective observational study.

Data exclusions

For stratified and within cancer analyses, we restricted to cancer types with >500 patients. For gene-level analyses, within each cancer type, we tested genes observed in $>5\%$ of sequenced tumors.

Replication

Genomic findings were corroborated in an external cohort of 5785 patients at Dana Farber Cancer Institute.

Randomization

There was no experimental allocation. Patient groups came from existing clinical and genomic records, not from investigator assignment. Covariates such as age, sex, and cancer stage were adjusted for using multivariable regression.

Blinding

Investigators were not blinded during data collection and analysis. Blinding does not apply here because group membership (for example, meeting the weight-loss criterion or tumor genotype status) comes directly from each patient's existing clinical and genomic records. Investigators did not assign patients to groups

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	NCT01775072
Study protocol	https://clinicaltrials.gov/study/NCT01775072
Data collection	Tumors were sequenced using the MSK-IMPACT assay through 9/19/2024, at study locations listed in the clinical trial protocol. This is a retrospective study using data from patients enrolled in the MSK-IMPACT prospective clinical sequencing protocol at Memorial Sloan Kettering Cancer Center, a single tertiary cancer center, between 2010 and 2024. An independent external validation cohort of 5,785 patients was obtained from Dana-Farber Cancer Institute.
Outcomes	The primary outcome was occurrence of a consensus-defined cachectic weight-loss episode. Episodes were detected from longitudinal BMI measurements using a 5% weight loss threshold over a 6-month window, or 2% for patients with baseline BMI under 20. Secondary outcomes included overall survival, changes in serologic markers during episodes, and associations between tumor genotype and episode risk. These were assessed using Cox models, mixed-effects models, and competing risk regression, as described in the Methods.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA