EMBO Molecular Medicine

Prediction of continuous amyloid PET with fluid measures of phosphorylated tau and beta-amyloid

Niklas Mattsson-Carlgren, Linda Karlsson, Weizhong Tang, Kaj Blennow, Henrik Zetterberg, Randall Bateman, Suzanne Schindler, Nicolas Barthelemy, Sebastian Palmqvist, Erik Stomrud, Shorena Janelidze, and Oskar Hansson

Corresponding authors: Niklas Mattsson-Carlgren (niklas.mattsson-carlgren@med.lu.se) , Oskar Hansson (oskar.hansson@med.lu.se)

Review Timeline:

Submission Date:	18th Jun 25
Editorial Decision:	12th Aug 25
Revision Received:	28th Oct 25
Editorial Decision:	5th Nov 25
Revision Received:	7th Nov 25
Accepted:	13th Nov 25

Editor: Jingyi Hou

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

12th Aug 2025

Dear Dr. Mattsson-Carlgren,

Thank you again for submitting your work to EMBO Molecular Medicine. We have now received the reports from the two reviewers and as you will see below, the reviewers think that the study is interesting. The reviewers have raised several mostly minor concerns, which we ask you to address in a revision.

I think the referees' recommendations are clear and need not be repeated here. All issues raised by the referees need to be satisfactorily addressed. As you may already know, our editorial policy allows in principle a single round of major revision so it is essential to provide responses to the referees' comments that are as complete as possible. Please feel free to contact me in case you would like to discuss in further detail any of the issues raised by the referees.

EMBO Molecular Medicine has a "scooping protection" policy, whereby similar findings that are published by others during review or revision are not a criterion for rejection. Should you decide to submit a revised version, I do ask that you get in touch after three months if you have not completed it, to update us on the status.

Please also contact us as soon as possible if similar work is published elsewhere. If other work is published, we may not be able to extend the revision period beyond three months.

Please read below for important editorial formatting and consult our author's guidelines for proper formatting of your revised article for EMBO Molecular Medicine.

I look forward to receiving your revised manuscript soon.

Use this link to login to the manuscript system and submit your revision: https://embomolmed.msubmit.net/cgi-bin/main.plex

Sincerely,
Jingyi

Jingyi Hou
Senior Editor
EMBO Molecular Medicine

*****

When submitting your revised manuscript, please carefully review the instructions that follow below. We perform an initial quality control of all revised manuscripts before re-review; failure to include requested items will delay the evaluation of your revision.

We require:

1) A .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) Individual production quality figure files as .eps, .tif, .jpg (one file per figure). For guidance, download the 'Figure Guide PDF': (https://www.embopress.org/page/journal/17574684/authorguide#figureformat).

3) A .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point responses to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

4) A complete author checklist, which you can download from our author guidelines (https://www.embopress.org/page/journal/17574684/authorguide#submissionofrevisions). Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

5) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript.

6) It is mandatory to include a 'Data Availability' section after the Materials and Methods. Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public database, and the accession numbers and database listed under 'Data Availability'. Please remember to provide a reviewer password if the datasets are not yet public (see https://www.embopress.org/page/journal/17574684/authorguide#dataavailability).

In case you have no data that requires deposition in a public database, please state so in this section. Note that the Data Availability Section is restricted to new primary data that are part of this study.

7) For data quantification: please specify the name of the statistical test used to generate error bars and P values, the number (n) of independent experiments (specify technical or biological replicates) underlying each data point and the test used to calculate p-values in each figure legend. The figure legends should contain a basic description of n, P and the test applied. Graphs must include a description of the bars and the error bars (s.d., s.e.m.). See also 'Figure Legend' guidelines: https://www.embopress.org/page/journal/17574684/authorguide#figureformat

8) At EMBO Press we ask authors to provide source data for the main manuscript figures. You will receive a separate email with instructions for providing source data with your revised manuscript, including how to upload and organize the files.

9) Our journal encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at .

10) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online. A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc... in the text and their respective legends should be included in the main text after the legends of regular figures.

- For the figures that you do NOT wish to display as Expanded View figures, they should be bundled together with their legends in a single PDF file called *Appendix*, which should start with a short Table of Content. Appendix figures should be referred to in the main text as: "Appendix Figure S1, Appendix Figure S2" etc.

- Additional Tables/Datasets should be labeled and referred to as Table EV1, Dataset EV1, etc. Legends have to be provided in a separate tab in case of .xls files. Alternatively, the legend can be supplied as a separate text file (README) and zipped together with the Table/Dataset file.

See detailed instructions here:

.

11) The paper explained: EMBO Molecular Medicine articles are accompanied by a summary of the articles to emphasize the major findings in the paper and their medical implications for the non-specialist reader. Please provide a draft summary of your article highlighting

- the medical issue you are addressing,

- the results obtained and

- their clinical impact.

This may be edited to ensure that readers understand the significance and context of the research. Please refer to any of our published articles for an example.

12) Author contributions: You will be asked to provide CRediT (Contributor Role Taxonomy) terms in the submission system. These replace a narrative author contribution section in the manuscript.

13) A Conflict of Interest statement should be provided in the main text.

14) Please provide a 'Synopsis' to further enhance discoverability. Synopses are displayed on the journal webpage and are freely accessible to all readers. They include a short stand first (maximum of 300 characters, including space) as well as 2-5 one-sentences bullet points that summarizes the paper. Please write the bullet points to summarize the key NEW findings. They should be designed to be complementary to the abstract - i.e. not repeat the same text. We encourage inclusion of key

acronyms and quantitative information (maximum of 30 words / bullet point). Please use the passive voice. Please attach these in a separate file or send them by email, we will incorporate them accordingly.

Please also suggest a visual abstract to illustrate your article as a PNG file 550 px wide x 300-600 px high.

15) All Materials and Methods need to be described in the main text using our 'Structured Methods' format. According to this format, the Methods section includes a Reagents and Tools Table (listing key reagents, experimental models, software and relevant equipment and including their sources and relevant identifiers) followed by a Methods and Protocols section describing the methods, ideally using a step-by-step protocol format. The aim is to facilitate adoption of the methodologies across labs.

Please download and fill our Reagents and Tools Table template (.docx), which you can find in our author guidelines:
https://www.embopress.org/page/journal/17574684/authorguide#structuredmethods

When submitting your revised manuscript, please do not include the Reagents and Tools Table in the Methods section of the manuscript but upload it as a separate file choosing the file type "Reagent Table".

***** Reviewer's comments *****

Referee #1 (Remarks for Author):

It is a great pleasure to participate in the review of this article. Overall, the author's approach is quite interesting. The integration of machine learning to explore optimal blood biomarkers for Alzheimer's disease makes this study distinctive. However, the following points require clarification:

The characterization of the cohort in the Results section is confusing. The remarkably low proportion of cognitively impaired participants yet a 37.1% PET-positive rate raises questions. Furthermore, the author did not clearly explain why discrepancies arose when processing data from different subpopulations.

While the presented data suggest the model is robust, I am curious about the origin of these 10 different regressor methods. Were algorithms with potentially higher predictive performance considered or evaluated?

Referee #2 (Comments on Novelty/Model System for Author):

In this manuscript, the authors use the BioFINDER-1 and -2 studies to develop machine learning algorithms to predict AB-PET using plasma and/or CSF biomarker measurements. This is an important question as understanding AB status can aid in diagnosis, prognosis, and in potential clinical trials. Furthermore, the availability of biomarkers are at lower cost and invasiveness, and could be broadly applicable where PET is not available. The study is unique in that it attempts to model not just AB positive vs negative, but also the continuous measures of AB-PET uptake. This uses a well established cohort and cutting edge technologies and analyses to address a potential important question.

Referee #2 (Remarks for Author):

Overall, the manuscript addresses an interesting and important question in the field of biomarker analysis for dementia diagnosis and monitoring. The manuscript is generally clearly written and the experimental design, results, and conclusions are well supported. I think this warrants publication, though they are a few minor revision suggestions I would recommend.

1 - the Introduction is quite densely written. A division of thoughts instead of two very long paragraphs could make for ease of reading.
2 - in the final feature selection it is stated that cognitive status/results added only minor increases. However, the number of individuals with dementia is very low in the dataset. Moreover, the number of APOE4/4 carriers is also quite low. I think that this should be addressed in the limitation section.
3 - again as a limitation to address, the introduction mentions using biomarker in a global context where AB-PET is unavailable. However, the genetic background and environmental context of the samples used in this training is quite homogenous. A statement regarding how this might impact global populations should be added.

Reviewer's comments

Referee #1 (Remarks for Author):

It is a great pleasure to participate in the review of this article. Overall, the author's approach is quite interesting. The integration of machine learning to explore optimal blood biomarkers for Alzheimer's disease makes this study distinctive. However, the following points require clarification:

The characterization of the cohort in the Results section is confusing. The remarkably low proportion of cognitively impaired participants yet a 37.1% PET-positive rate raises questions. Furthermore, the author did not clearly explain why discrepancies arose when processing data from different subpopulations.

Authors' response: We thank the reviewer for this important comment and are happy to clarify. In BioFINDER-2, Aβ-PET was generally not performed in participants with dementia, by study design. This explains the relatively low proportion of cognitively impaired participants in our sample. Moreover, the cognitively unimpaired study arms of BioFINDER-2 were deliberately enriched for APOE ε4 carriers, which further increases the expected prevalence of Aβ-positivity. Taken together with global prevalence estimates of Aβ-positivity (according to Jansen et al. (2022), the prevalence of Aβ-positivity is approximately 20% in cognitively unimpaired individuals around age 65 (average age of our sample), and ~45% in individuals with MCI), these factors explain the observed Aβ-positivity rate.

For clarity, we have now stated as a limitation that Aβ-PET was generally not performed in participants with dementia (page 9, lines 341ff):

"A fourth limitation was that, according to the BF2 study protocol, Aβ-PET was rarely performed in individuals with dementia, which limited our ability to evaluate model performance at this disease stage."

Additionally, we also now comment on the Aβ-positivity prevalence in our sample (page 4, lines 73-75):

"The overall rate of Aβ-PET positivity was 37.1% (consistent with that the cognitively unimpaired study arms of BioFINDER-2 were enriched for APOE ε4 carriers and with established global prevalence estimates)."

With regards to different subpopulations, we found relatively robust results across populations (but somewhat lower for CSF based tests for p-tau217 compared to plasma p-tau217), as described in the sensitivity analysis section (page 7, lines 223-225): "Specifically, the mean R^2 scores averaged around 0.74 and 0.72 for BF2-C-IA and BF2-C-MS respectively, in contrast to the mean R^2 scores of 0.83 and 0.78 for BF2-P-MS and BF2-P-IA."

While the presented data suggest the model is robust, I am curious about the origin of these 10 different regressor methods. Were algorithms with potentially higher predictive performance considered or evaluated?

Authors response: We appreciate the opportunity to expand on the description and discussion of the ML models. We recognize that there are many possible ML algorithms with

varying levels of complexity and predictive performance. Our goal in this study was to sample broadly across different classes of models, linear (e.g., linear regression, ridge), tree-based (e.g., random forest, gradient boosting, extra trees), and kernel-based (e.g., support vector regression), to capture a wide range of possible relationships within the data. We focused on algorithms that are well established, widely used, and known to perform robustly across diverse regression tasks.

In addition, we systematically explored different input feature combinations, performed hyperparameter tuning, and cross-validation procedures to ensure a rigorous evaluation. While it is possible that other methods (e.g., deep learning architectures) might achieve higher predictive performance, we aimed to balance model diversity with computational feasibility. Within this framework, the Extra Trees Regressor achieved the best performance. We agree that future work could extend this search to include more complex or emerging algorithms, which we now have noted in the Discussion/Limitations (page 9, lines 337ff):

"Third, while we focused on a range of well-established models and rigorously optimized their features and hyperparameters, future work could explore more advanced algorithms, such as deep learning architectures (which may need larger sample sizes for efficient training), to potentially further improve predictive performance."

Referee #2 (Comments on Novelty/Model System for Author):

In this manuscript, the authors use the BioFINDER-1 and -2 studies to develop machine learning algorithms to predict AB-PET using plasma and/or CSF biomarker measurements. This is an important question as understanding AB status can aid in diagnosis, prognosis, and in potential clinical trials. Furthermore, the availability of biomarkers are at lower cost and invasiveness, and could be broadly applicable where PET is not available. The study is unique in that it attempts to model not just AB positive vs negative, but also the continuous measures of AB-PET uptake. This uses a well established cohort and cutting edge technologies and analyses to address a potential important question.

Referee #2 (Remarks for Author):

Overall, the manuscript addresses an interesting and important question in the field of biomarker analysis for dementia diagnosis and monitoring. The manuscript is generally clearly written and the experimental design, results, and conclusions are well supported. I think this warrants publication, though they are a few minor revision suggestions I would recommend.

1 - the Introduction is quite densely written. A division of thoughts instead of two very long paragraphs could make for ease of reading.

Authors' response: We thank the reviewer for this suggestion and agree that the introduction would benefit from such a division. We have therefore now separated the introduction into five paragraphs instead of two. These paragraphs focus on 1) novel AD treatments increasing the need for methods to quantify Aβ burden, 2) existing PET, CSF and plasma biomarkers related to Aβ, 3) previous studies using fluid biomarkers to predict Aβ PET status, 4) previous efforts to predict continuous Aβ PET burden, and 5) our specific aims and hypotheses.

2 - in the final feature selection it is stated that cognitive status/results added only minor increases. However, the number of individuals with dementia is very low in the dataset. Moreover, the number of APOE4/4 carriers is also quite low. I think that this should be addressed in the limitation section.

Authors' response: We thank the reviewer for this valuable observation and fully agree that it is important to address. The number of individuals with dementia in our dataset is low, which is due to that in the BioFINDER-2 study protocol, amyloid PET was generally not performed in participants with dementia. Consequently, cognitive status may have contributed more strongly if a larger number of dementia cases had been included. The lower number of dementia patients may also be expected to somewhat reduce the number of APOE ε4 homozygotes. In addition, inclusion in BioFINDER-2 was stratified by APOE ε4 status among cognitively unimpaired individuals. Together, this may influence the apparent contribution of APOE ε4 compared to what would be expected in a population with frequencies closer to global prevalence. We agree that both of these points should be explicitly acknowledged in the Discussion/Limitations section, and we have revised the manuscript accordingly (page 9, lines 341ff):

"A fourth limitation was that, according to the BF2 study protocol, Aβ-PET was rarely performed in individuals with dementia, which limited our ability to evaluate model performance at this disease stage. Certain features, such as cognitive status, may have contributed more strongly to prediction if a larger number of dementia cases had been included. In addition, inclusion in BF2 was stratified by APOE ε4 status among cognitively unimpaired individuals, which may have influenced the contribution of this feature compared to a population with frequencies closer to global prevalence."

3 - again as a limitation to address, the introduction mentions using biomarker in a global context where AB-PET is unavailable. However, the genetic background and environmental context of the samples used in this training is quite homogenous. A statement regarding how this might impact global populations should be added.

Authors' response: We thank the reviewer for highlighting this important point. We agree that the relatively homogenous genetic and environmental background of the BioFINDER-2 cohort could limit the direct generalizability of our findings to more diverse global populations. We have now added a statement in the Discussion/Limitations section to explicitly note this and emphasize the need for validation in more diverse populations (page 9, lines 347ff):

"Another limitation was the relatively homogenous genetic and environmental background of BF2, factors that could potentially influence biomarker performance. Even though the models generalized well to the independent cohort BF1, findings should be validated across more diverse populations in future studies."

5th Nov 2025

Dear Dr. Mattsson-Carlgren,

Thank you for the submission of your revised manuscript to EMBO Molecular Medicine. We have now received the enclosed report from the referee who was asked to re-assess it. As you will see, the referee is now supportive, and I am pleased to inform you that we will be able to accept your manuscript pending the following amendments:

1. Please remove the figures from the manuscript file and upload them as separate, high-resolution figure files. The legends should stay in the manuscript text. Please add the heading "Expanded View Figure Legends" after the main tables and before the EV figure legends.

2. Please add missing callouts for panels A,B of Figre 5 and panels A,B,C of Figure EV1.

3. "Funding" section needs to be renamed to "Acknowledgments". Please verify that the funding details (including project numbers) in the manuscript match those entered in the submission system. The line-by-line entries provided in our system will be automatically linked upon publication, so accuracy and completeness are essential.

4. Appendix: please add page numbers in the Table of Contents.

5. Please move the information in "Ethics" to the "Methods" section where appropriate.

6. Source data: please upload then as one (Zip) file per figure.

7. Data availability : Please revise the following sentence "By request, anonymized data can be shared to qualified academic investigators for the purpose of replicating procedures and results presented in the article."

to:

"By request, anonymized data will be shared to qualified academic investigators."

8. Please download and fill our Reagents and Tools Table template (.docx), which you can find in our author guidelines: https://www.embopress.org/page/journal/17574684/authorguide#structuredmethods

When submitting your revised manuscript, please do not include the Reagents and Tools Table in the Methods section of the manuscript but upload it as a separate file choosing the file type "Reagent Table".

9. "Disclosure" should be renamed to "Disclosure and Competing Interests Statement".

10. Please move the "summary" to the manuscript file and change the heading to "The Paper Explained".

11. Please upload the EV tables as two separate files and remove them from the manuscript text.

12. The references need to be formatted according to the EMBO Molecular Medicine reference style. Citations should be listed in alphabetical order. Please list up to 10 co-authors of a paper before adding et al. in the reference list. Remove the DOIs for all published articles.

13. Please correct the order of the manuscript sections to: Abstract / Keywords / The Paper Explained / Introduction /Results / Discussion / Methods / Data Availability /Acknowledgements / Disclosure and Competing Interests Statement / References / Figure Legends / Tables / Expanded View Figure Legends

14. Please note that the box plots need to be defined in terms of minima, maxima, centre, bounds of box and whiskers, and percentile in the legends of figures 2C, 4, 5, EV1 A-C.

Please submit your revised manuscript within two weeks.

I look forward to reading a new revised version of your manuscript as soon as possible.

Kind regards,
Jingyi

Jingyi Hou
Senior Editor
EMBO Molecular Medicine

*** Instructions to submit your revised manuscript ***

*** PLEASE NOTE *** As part of the EMBO Publications transparent editorial process initiative (see our Editorial at https://www.embopress.org/doi/pdf/10.1002/emmm.201000094), EMBO Molecular Medicine will publish online a Review Process File to accompany accepted manuscripts.

In the event of acceptance, this file will be published in conjunction with your paper and will include the anonymous referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript. If you do NOT want this file to be published, please inform the editorial office at contact@embomolmed.org.

***** Reviewer's comments *****

Referee #1 (Comments on Novelty/Model System for Author):

This study presents an innovative machine learning approach using BioFINDER-1 and -2 data to predict Aβ-PET from plasma and CSF biomarkers. The work is methodologically sound, clinically meaningful, and clearly presented.

Referee #1 (Remarks for Author):

The authors have responded carefully and effectively to all reviewer comments. The revised manuscript is much improved and suitable for consideration.

The authors addressed the remaining editorial issues.

13th Nov 2025

Dear Dr. Mattsson-Carlgren,

We are pleased to inform you that your manuscript is accepted for publication and is now being sent to our publisher to be included in the next available issue of EMBO Molecular Medicine.

Your manuscript will be processed for publication by EMBO Press. It will be copy edited and you will receive page proofs prior to publication. Please note that you will be contacted by Springer Nature Author Services to complete licensing and payment information.

You may qualify for financial assistance for your publication charges - either via a Springer Nature fully open access agreement or an EMBO initiative. Check your eligibility: https://www.embopress.org/page/journal/17574684/authorguide#chargesguide

Should you be planning a Press Release on your article, please get in contact with embo_production@springernature.com as early as possible in order to coordinate publication and release dates.

If you have any questions, please do not hesitate to contact the Editorial Office. Thank you for your contribution to EMBO Molecular Medicine.

Yours sincerely,
Jingyi

Jingyi Hou
Senior Editor
EMBO Molecular Medicine

>>> Please note that it is EMBO Molecular Medicine policy for the transcript of the editorial process (containing referee reports and your response letter) to be published as an online supplement to each paper. If you do NOT want this, you will need to inform the Editorial Office via email immediately. More information is available here: https://www.embopress.org/transparent-process#Review_Process

EMBO Press Author Checklist

Corresponding Author Name: Niklas Mattsson-Carlgren	
Journal Submitted to: EMBO Molecular Medicine	
Manuscript Number: EMM-2025-22132	

USEFUL LINKS FOR COMPLETING THIS FORM
The EMBO Journal - Author Guidelines
EMBO Reports - Author Guidelines
Molecular Systems Biology - Author Guidelines
EMBO Molecular Medicine - Author Guidelines

Reporting Checklist for Life Science Articles (updated January
This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your
Please note that a copy of this checklist will be published alongside your article.

Abridged guidelines for figures
1. Data
The data shown in figures should satisfy the following conditions:
- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical
- if n<5, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data

2. Captions
Each figure caption should contain the following information, for each panel where they are relevant:
- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements.
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Please complete ALL of the questions below.
Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Not Applicable	

Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For **antibodies** provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number - Non-commercial: RRID or citation	Not Applicable	

DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Not Applicable	

Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID.	Not Applicable	
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Not Applicable	
Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination.	Not Applicable	

Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID.	Not Applicable	
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail **housing and husbandry conditions**.	Not Applicable	

Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Not Applicable	
Microbes: provide species and strain, unique accession number if available, and source.	Not Applicable	

Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Yes	Table 1 and Table 2.

Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgments section?	Not Applicable	

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI.	Not Applicable	
Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable.	Yes	Methods - Participants.

Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if **external detailed step-by-step protocols** are available.	Not Applicable	

Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about **sample size** estimate even if no statistical methods were used.	Yes	Table 1 and Table 2.
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described?	Not Applicable	
Include a statement about **blinding** even if no blinding was done.	Yes	Methods - Cerebrospinal fluid (CSF) and plasma biomarkers
Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established? If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.	Yes	Results - Model development and feature engineering
For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Figure 2, Figure 4, Figure 5.

Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was **replicated** in laboratory.	Not Applicable	
In the figure legends: define whether data describe **technical or biological replicates**.	Not Applicable	

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval.	Yes	Methods - ethics
Studies involving **human participants**: Include a statement confirming that **informed consent** was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Yes	Methods - ethics
Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained.	Not Applicable	
Studies involving experimental **animals**: State details of **authority granting ethics approva**l (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Not Applicable	
Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	

Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm	Not Applicable	
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	
If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript?	Not Applicable	

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided.	Not Applicable	
For **tumor marker prognostic studies,** we recommend that you follow the **REMARK** reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Not Applicable	
For **phase II and III randomized controlled trials**, please refer to the **CONSORT** flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	Not Applicable	

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section?	Yes	Data availability
Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Yes	Data availability
Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Yes	Methods - Statistical analyses
If publicly available data were reused, provide the respective **data citations in the reference list**.	Not Applicable	