

Parsimony over panel size: a stability-selected four-gene mitoxyperilysis-related score is an independent, externally reproduced prognostic factor in lower-grade glioma

Huiling Zhai

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0001-4945-8392>

Jingyi Chen

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0002-4482-4661>

Bin Lian

drmi1o@gkd.edu.cn

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0000-0002-1477-9137>

Yanlin Qiu

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0007-8349-0322>

Yingxin Wu

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0008-3486-8795>

Yiyi Zhong

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0007-0523-3451>

Yini Sun

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0003-9751-7232>

Jiayi Kuang

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0005-0826-2341>

Jiawei Zheng

Guangzhou Vocational and Technical University of Science and Technology <https://orcid.org/0009-0007-5237-0504>

Jinyuan Liu

Research Article

Keywords: lower-grade glioma, mitoxyperilysis-related genes, prognostic signature, stability selection, external validation

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11293196/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Purpose. Mitoxyperilysis is a lytic cell death modality: innate immune stimulation plus metabolic disruption drive BAX/BAK1/BID-dependent mitochondrial oxidative stress and membrane rupture. A related colorectal cancer signature exists, so priority is unavailable; whether they prognosticate in glioma is open.

Methods. A 101-gene mouse mitoxyperilysis-related gene (MRG) panel (GSE235046) gave 89 evaluable genes in TCGA-LGG (n = 510, 125 deaths). A 25-gene ridge Cox index was re-derived by LASSO-Cox (10-fold cross-validation) and 200-bootstrap stability selection ($\pi \geq 0.75$), plus one post hoc elimination ($P < 0.05$), reproduced in CGGA-325 (WHO II–III, n = 172, 90 deaths).

Results. Stability selection retained BMP1, KIF15, TRAF3, PDGFA, OSGIN2 ($\pi = 0.775–0.875$); only OSGIN2 was non-significant ($P = 0.38$), leaving MRG-4 (apparent C-index 0.794, cross-validated 0.796 ± 0.019). In CGGA-325 the external C-index rose from 0.664 (0.610–0.715) to 0.769 (0.722–0.812), $\Delta + 0.105$ (95% CI 0.054–0.162), exceeding 99.3% of 3,000 random four-gene draws; decrement-weighted AUC 0.845 vs 0.731 ($\Delta + 0.114$, 95% CI 0.054–0.177; zero excluded at 18 of 19 horizons).

Conclusions. The score reproduced rank ordering, not a transportable absolute-risk instrument: its advantage over a four-variable clinical model was not established (+ 0.045, – 0.018 to 0.111). At $\pi \geq 0.80$ they fall out only in the original run — an independent seed selects PBX2, drops KIF15 — so no error control is claimed for it. It marks a proliferation-high, matrix-remodelling LGG state, not mitoxyperilysis.

1. Introduction

Diffuse gliomas are graded on a molecular axis: the 2021 WHO classification separates IDH-mutant from IDH-wildtype tumours, with CDKN2A/B deletion and 1p/19q codeletion defining risk strata within IDH-mutant disease ^{1,2}. Standard management is maximal safe resection plus radiotherapy and temozolomide ³, yet age, grade, resection extent and molecular class leave wide residual spread ^{1,2}.

Ferroptosis, pyroptosis and necroptosis have each become expression indices separating glioma patients by survival ^{4,5}. Most share three weaknesses: the gene set intersects a literature pathway with differential expression; the index often proves a generic proliferation or inflammatory score; and the model is sized to the panel, not to events, so reported discrimination does not reproduce.

Mitoxyperilysis is a perturbed pathway of the first kind: innate immune stimulation plus metabolic disruption — lipopolysaccharide with carbon starvation — kills macrophages independently of caspases, pyroptosis, PANoptosis, necroptosis, ferroptosis and oxeiptosis ⁶, and execution runs from BAX/BAK1/BID-dependent mitochondrial oxidative stress ⁷ to prolonged mitochondria–plasma membrane contact, the NINJ1 pore being the regulated executor of rupture ⁸. NINJ1 is not among the panel genes used here, so no score component is execution machinery, and mTORC2 is only permissive ^{6,9,10}. A mitoxyperilysis-related colorectal cancer signature stratifies prognosis ¹¹, so this study cannot

claim priority. In LGG the rationale is analogous rather than established: IDH-mutant cells accumulate 2-hydroxyglutarate, rewiring redox metabolism ¹², and gliomas are infiltrated by macrophages and microglia, but co-occurrence of the two arms in one cell is not shown here.

We derive an index from the full 89-gene panel in TCGA-LGG under three tests fixed before external validation: adjustment for IDH status and WHO grade, head-to-head comparison with a proliferation score, and external validation in CGGA-325 with frozen coefficients. It passes all three, yet external discrimination falls to 0.664 with coefficients unstable in sign, so we treat it as a hypothesis to compress: re-derive it by LASSO-Cox with stability selection, freeze a four-gene score, and re-run validation.

2. Materials and methods

2.1 Data sources and cohort assembly

TCGA-LGG expression (RNA-Seq V2) and clinical data (*lgg_tcga_pan_can_atlas_2018*) were obtained via the cBioPortal API ¹³, and CGGA-325 RSEM expression and annotation ¹⁴, requiring primary tumours with complete OS and panel expression: 510 patients (125 deaths, 24.5%), 502 fully annotated (123 deaths) (Table 1). Analysis-set derivation (510 → 502 → 410; CGGA-325 325 → 313 → 182 → 172) is deposited in Supplementary Tables S21–S22; validation is **172 patients (90 deaths)**, with median follow-up 27.8 versus 130.2 months (event rates 24.5% versus 52.3%). Both cohorts predate WHO 2021, so many of the 136 IDH-wildtype cases would be graded glioblastoma under the current classification (Supplementary Note S2).

2.2 Construction of the mitoxyperilysis-related gene (MRG) panel

The panel came from GSE235046 (mouse bone-marrow-derived macrophages; media, LPS, carbon starvation [CS], LPS+CS, LPS+CS+Torin; three replicates per arm) by dual-signal synergy and mTOR-sensitivity criteria. Torin inhibits mTORC1 and mTORC2 alike, so this cannot establish mTORC2 dependence, and no criterion measured the death phenotype: the genes are immune-plus-metabolic-stress responders in a macrophage system, not validated effectors of mitoxyperilysis. This yielded 101 core genes (Supplementary Table S1); 92 had a human orthologue and three were unevaluable, leaving **89 genes**.

2.3 Model A – the full-panel index (L2-penalised Cox)

Univariate Cox regression on the 89 genes ($P < 0.05$) fed a multivariable L2 ridge-penalised model ($\lambda = 1.0$; coefficients, Supplementary Table S12; median dichotomisation). Discrimination used Harrell's C-index ¹⁵, internal optimism 10-fold cross-validation, and the events-per-variable (EPV) ratio based on deaths ^{16,17}.

2.4 Model B – stability-selected four-gene score (MRG-4)

Model B re-derives the index from the same 89 genes and 510 patients, gene-wise Z-scores being re-standardised by discovery means and SDs. CoxnetSurvivalAnalysis (scikit-survival 0.28.0) was fitted over a 100-value α path¹⁸ with α set by 10-fold cross-validation and the 1-SE rule. Following Meinshausen and Bühlmann¹⁹, 200 bootstrap resamples were drawn, the LASSO refitted at the 1-SE α , and each gene's selection frequency π recorded; genes with $\pi \geq 0.75$ entered a multivariable Cox model, dropping members with $P \geq 0.05$. The apparent C-index 0.794 and cross-validated 0.796 ± 0.019 retain optimism; the fully nested 0.731 ± 0.023 is conservative. The Meinshausen–Bühlmann bound $E[V] \leq q^2/((2\pi - 1)p)$ is non-binding ($E[V] \leq 32.4$ at $\pi = 0.75$), so the threshold carries no useful false-discovery control, and OSGIN2 removal ($P = 0.382$) is a post hoc filter on top of it (Section 3.3).

2.5 Adjustment for IDH status and WHO grade (both models)

Both models were adjusted for IDH status, WHO grade, age and sex (502 complete-annotation patients); CGGA-325 added 1p/19q codeletion, MGMT promoter methylation and treatment in a sensitivity analysis. CDKN2A/B deletion and extent of resection were **not** adjusted for, because neither cohort carries a reliable field.

2.6 Comparison against a proliferation score (both models)

A 20-gene proliferation score from canonical cell-cycle markers²⁰ was assessed against each MRG score by Spearman correlation and multivariable Cox regression with and without clinical covariates.

2.7 External validation in CGGA-325 (both models)

Coefficients were frozen and the risk-group threshold is the within-cohort median, so the log-rank test addresses relative ordering, not an absolute cut-off. Under the primary cohort-z convention MRG-4 reaches C-index 0.769 (95% CI 0.722–0.812) versus 0.664 (0.610–0.715) for the 25-gene model, and 0.731 versus 0.620 under discovery-SD scaling (bootstrap 95% CIs, 2,000 resamples). The raw and z spectra of each gene differ by a monotone mapping that preserves rank strictly, so patient ordering is unchanged (Supplementary Table S17).

2.8 Head-to-head comparison of Model A and Model B

Comparisons were specified before the external validation was run. Harrell's C-index and time-dependent AUC at 12–120 months in 6-month steps used inverse-probability-of-censoring weighting, censoring estimated in the cohort evaluated. NRI and IDI at 3 and 5 years used patients with a baseline hazard only (160 at 3 years, 156 at 5 years; Table 3, Supplementary Table S30), and decision curve analysis at 3 years

used thresholds 0.20–0.80 (50 three-year deaths; prevalence 0.291) against treat-all/treat-none and a four-variable clinical model. C-index differences used a paired bootstrap (2,000 resamples), resolving differences ≥ 0.05 , and each external C-index was referenced against equal-sized random panels from the same 89 genes, fitted with the **same penalisation as the model compared**. The Model A comparator is the archived linear predictor, reproduced by the deposited coefficients at $r = 0.9148$ over the 510 discovery patients (Supplementary Tables S20, S25, S30).

2.9 Molecular subtyping and immune infiltration

Clustering (k-means, $k = 2$) on the **25 Model A genes** defines Model A gene-based subtypes, not "MRG subtypes". Eleven immune features were scored as the mean Z-score of 4–6 marker genes (ssGSEA-lite)²¹ and compared between the **25-gene Model A** risk groups by Mann–Whitney U, uncorrected for multiplicity (Supplementary Tables S26, S26b, S26c).

2.10 Somatic mutation analysis

Mutation calls for ten recurrent glioma genes were compared between **25-gene Model A** risk groups by Fisher's exact test, with per-patient and MRG-4-split results in Supplementary Tables S8 and S9b.

2.11 Single-cell localisation

Single-cell expression of the MRG-4 genes was resolved in the CZ CELLxGENE Discover census (release 2025-11-17) via its public API²², TISCH2²³ having failed. The largest 10x glioblastoma dataset (SL040, 135,482 cells, single donor, grade IV glioma; no grade II–III diffuse glioma entity is deposited) was used. Values are raw h5ad counts without normalisation, so every per-cell quantity is a mean UMI count, computed over the 15 annotated cell types and, within the malignant compartment, over the four Neftel states.

2.12 Software, reporting and reproducibility

Analyses ran in Python 3.13 with scikit-survival 0.28.0 and lifelines 0.30.3; seeds and environment are in Supplementary Table S28, the 25 Model A coefficients in Supplementary Table S12. The original stability run used an 88-gene candidate set, the replication the full 89; only SAMD8 is replication-only. The score is defined on the distributed cBioPortal z-score profile, with Supplementary Table S17 listing the four coefficients, profile identifiers and each gene's raw-versus-z correlation, and Supplementary Table S19 the Breslow baseline cumulative hazard. Reporting follows TRIPOD 2015²⁴, and risk of bias was self-assessed with PROBAST²⁵ and its current extension²⁶; both checklists are deposited as Supplementary Tables S13 and S27. The analysis code and deposited derived tables are openly available at

<https://github.com/speetle/lgg-mrg4-mitoxyperilysis> (release v1.0.0), with a SHA-256 manifest covering every deposited table.

3. Results

3.1 Univariate screening identifies 56 survival-associated MRGs

Of 89 evaluable MRGs (human orthologues of a 101-gene mouse panel: 35 up-regulated, 66 down-regulated), 56 associated with OS at raw $P < 0.05$ (50 at Benjamini–Hochberg $q < 0.05$; 57 at $q < 0.10$, 64 at $q < 0.20$). Strongest risk genes: LOXL3 (HR 1.795), BMP1 (HR 1.636), PDGFA (HR 1.609); strongest protective: CALCRL (HR 0.608), OSGIN2 (HR 0.675) (Supplementary Table S1, Supplementary Figure S1A).

3.2 Model A: a 25-gene full-panel index separates LGG into two survival groups

L2-penalised multivariable Cox on the 56 hits retained 25 genes (Supplementary Table S12). Median split (255 vs 255): worse OS high-risk (HR 3.27, 95% CI 2.21–4.83); continuous HR 2.93 per SD (2.47–3.47), C-index 0.8211, one-, three- and five-year AUCs 0.907, 0.861, 0.801 (Supplementary Figure S1D, S2A). Adjusted for age, sex, IDH-wildtype status and grade 3: HR 2.35 per SD (1.86–2.97), neither IDH-wildtype nor grade 3 significant (Table 2). An EPV of 5.0 (125 deaths, 25 terms) ^{16,17} and 10 of 25 coefficients flipping sign between univariate and multivariate analyses ¹⁵ are the two features that matter.

3.3 Model B: stability selection compresses the index to four genes

Cross-validation. The 10-fold curve was flat-topped; the 1-SE rule selected 26 genes (CV C-index 0.777) (Figure 1A; Supplementary Table S7). A pure LASSO-plus-1-SE rule was not sparse by itself, so sparsity came from stability: across 200 bootstraps the median model held 27 genes, but BMP1, KIF15, TRAF3 and PDGFA were most reproducible (Supplementary Table S31, S6; Figure 1B).

Stability, and its limit. $\pi \geq 0.75$ yielded five genes (the four plus OSGIN2), re-selected by an independent seed adding PBX2 and PPM1M; OSGIN2 was the only non-significant member (HR 0.921, $P = 0.382$). Three of the four survive Benjamini–Hochberg $q < 0.05$ (PDGFA 7.67×10^{-13} , BMP1 1.40×10^{-11} , KIF15 7.54×10^{-10}); TRAF3 does not ($q = 0.056$). $\pi \geq 0.80$ returns exactly the four genes in the **original** run only: in the replication PBX2 rises to 0.800 and is selected while KIF15 falls to 0.780 and is dropped, so that threshold is seed-dependent and carries no false-discovery control (Section 2.4) (Figure 1C).

The score (per SD, discovery-cohort standardisation): MRG-4 = **+0.258**·BMP1 **+0.464**·KIF15 **–0.238**·TRAF3 **+0.480**·PDGFA; EPV 31.2 (125 deaths, 4 terms) against 5.0 for Model A, though the search over 89 candidates is not priced; C-index 0.794, 10-fold CV 0.796 ± 0.019 , conservative fully nested 0.731 ± 0.023 (Section 3.6). Restricted mean survival times: 35.1 vs 29.4 months at 36

(difference 5.6, 95% CI 4.1–7.2) and 55.3 vs 42.1 at 60 months (13.2, 9.5–16.8) in TCGA-LGG; 10.6 (7.7–13.4) and 22.6 (17.2–27.7) in CGGA-325 (Supplementary Table S24).

3.4 MRG-4 is independent of IDH status and WHO grade in the discovery cohort

In 502 annotated patients (123 deaths), MRG-4 remained significant after adjustment for age, sex, IDH-wildtype status and grade 3 (HR 1.57 per SD, 95% CI 1.23–1.99; $P = 2.46 \times 10^{-4}$); unlike Model A, IDH-wildtype status (HR 2.35, $P = 0.0016$) and grade 3 (HR 1.63, $P = 0.031$) also remained significant (Table 2). Its IDH association is weaker — rank-biserial $r = -0.69$ against -0.84 for the 25-gene index (Supplementary Figure S2C) — and validation gives $P = 3.2 \times 10^{-5}$ ($n = 171$ with an IDH call). Adding the remaining 2021 WHO molecular variables kept the per-SD hazard ratio between 1.33 and 1.70 in both cohorts (Supplementary Table S15).

3.5 MRG-4 is proliferation-adjacent, but it dominates the proliferation score in joint models

Model A correlated only weakly with a 20-gene proliferation score ($\rho = 0.26$; Supplementary Figure S2B), whereas **MRG-4 correlated substantially ($\rho = 0.67$; $n = 502$)**. MRG-4 was not that score: it stayed significant in a joint model (HR 1.488 per SD after IDH status, grade, age and sex, $P = 0.0054$, against non-significant proliferation HR 1.095, $P = 0.507$; Table 2), as did the 25-gene index (HR 2.694 vs 1.167, $P = 0.099$; Table 2). **MRG-4 is therefore not a generic proliferation score statistically.**

3.6 External validation: MRG-4 outperforms the 25-gene model in CGGA-325

In WHO grade II–III tumours ($n = 172$, 90 deaths), MRG-4 was prognostic: C-index **0.769** (bootstrap 95% CI 0.722–0.812, cohort-internal), **0.731** (0.682–0.776) under discovery-SD scaling, against 0.664 for the 25-gene index; **1.73 per SD** unadjusted (1.49–2.01) and HR 1.65 adjusted (1.39–1.97), with median OS unreached versus 31.3 months (log-rank $P = 6.61 \times 10^{-11}$; Figure 2A). The smaller model wins head-to-head (EPV 5.0 vs **31.2**; sign flips 10/25 vs **0/4**), the external-C-index difference being **+0.105 (95% CI 0.054–0.162)** (Table 3; Figure 2B). Time-dependent AUC over 19 horizons was **0.845** against **0.731** (unweighted **0.834/0.717**), higher at 18 of the 19 horizons, with paired bootstraps giving **+0.114 (95% CI 0.054–0.177)**, excluding zero at **18 of the 19** points, and **+0.045 (95% CI -0.018 to 0.111)** against the four-variable clinical model (Figure 3A).

Two corrections. Both models scored 0.796 under 10-fold CV because the gene set was fixed before the folds; moving selection inside reverses the ranking (0.832 ± 0.021 vs 0.731 ± 0.023). Against a matched null, **0.68% of 2,287 valid unpenalised four-gene draws matched or beat 0.769**, while at $\lambda = 1.0$ only **348 of 3,000 25-gene panels fitted and 97.8% matched or beat 0.664**; unpenalised, the 25-gene null is

beaten by only **49.3%** (Supplementary Table S30). The calibration slope was **0.548 (95% CI 0.400–0.697)**, by construction the cohort's own per-SD log hazard ratio, so **no absolute risks should be read from it** (Figure 3C).

Where MRG-4 does not win. NRI at 3 years -0.291 (IDI -0.047) and 5 years -0.266 (IDI -0.016), from **160 patients with 50 three-year events**, favour Model A, neither significantly, and validation refitting reverses the sign (NRI **+0.440**, IDI **+0.041**). MRG-4 beats the 25-gene model on net benefit at **20 of the 31 thresholds**, while a four-variable clinical model **refitted in the same 172 patients** is highest at **21 of the 31** (Figure 3B). We claim improved *discrimination and transportability*, not reclassification, clinical utility, or superiority over standard clinical variables (Table 3).

3.7 Subgroup effects

MRG-4 stratified every TCGA-LGG subgroup except WHO grade II (borderline, $P = 0.055$), retaining significance in IDH-wildtype tumours ($P = 0.0068$), where the 25-gene model did not ($P = 0.463$). In CGGA-325 it stratified all four subgroups — IDH-mutant ($P = 2.2 \times 10^{-6}$), IDH-wildtype ($P = 0.0022$), WHO grade II ($P = 0.0018$), grade III ($P = 3.1 \times 10^{-5}$) — whereas the 25-gene model stratified grade III only. The 20 stratum tests are uncorrected for multiplicity (Supplementary Table S29; Supplementary Figure S3A–C).

3.8 Single-cell localisation: the score is not a macrophage signal, and not tumour-exclusive either

We resolved the four MRG-4 genes (MKI67 as proliferation reference) in glioma single-cell data; the census²² has no grade II–III diffuse glioma dataset, so grade IV glioma is used with that caveat (Supplementary Figure S4A,B; Supplementary Table S10). In dataset SL040²⁷ — 135,482 cells from a **single donor (n = 1)**, with 48,281 malignant cells (35.6%) — transcript share makes all three positively weighted genes tumour-weighted (PDGFA **73.9%**, BMP1 **77.0%**, KIF15 **79.7%**), while TRAF3 sits at **23.9%**. By per-cell mean expression, however, PDGFA is marginally higher in vascular smooth-muscle cells than in malignant cells and BMP1 is higher in pericytes and fibroblasts, so **only KIF15 is tumour-restricted on both metrics**. The composite score ranks malignant cells first ($+0.430$) and myeloid/lymphoid compartments lowest (MoTAM -0.349): stromal-low, not stromal-free (Supplementary Tables S11, S16).

3.9 What the four genes are, and what they are not

Two analyses bear on the score's relation to the mitoxyperilysis pathway; **neither tests pathway activity**. Across all 89 evaluable genes the Spearman correlation between the mouse induction $\log_2$ fold-change and the human univariate log hazard ratio was **$\rho = -0.320$ ($P = 0.0023$, $n = 89$)** — significantly *opposite* to the expectation that genes induced in the model are risk genes in patients. In CGGA-325 the index remained significant when added to a 20-gene proliferation score ($\chi^2 = 14.9$, $P = 1.1 \times 10^{-4}$). A planned

comparison against a 15-gene extracellular-matrix score cannot be reported: its member genes are in no deposited object.

4. Discussion

The result is not that a mitoxyperilysis-related score predicts glioma survival, but that the same panel yields a 25-gene index that does not transport and a four-gene one that does.

The full-panel index came first (56 univariate hits, a ridge penalty) and passed our three tests, yet two things were wrong: an EPV of 5.0, half the conventional threshold^{16,17}, with 10 of 25 coefficients opposite in sign, and external discrimination of 0.664 against an apparent 0.821. Our first internal validation missed this, returning 0.796 for both models, because the gene set was fixed before the folds; repeating selection inside the folds favours the larger model, and 510 patients and 125 events cannot settle this. Randomised-panel benchmarking did: the four-gene score beat 99.3% of penalised random four-gene panels, the 25-gene index falling below the mean of its own matched null. Sparsity came from repeated selection, not the LASSO 1-SE rule (its flat curve selected 26 genes): BMP1, KIF15, TRAF3 and PDGFA were chosen in 80.5–87.5% of 200 bootstraps, and all four cleared $\pi \geq 0.75$ under an independent seed, though the stricter $\pi \geq 0.80$ subset is not reproduced.

The four-gene score wins on discrimination and fit (external C-index 0.769 versus 0.664; higher time-dependent AUC at 18 of the 19 horizons) but correlates with a 20-gene proliferation score at $\rho = 0.67$ against 0.26; reclassification is not improved. Biological specificity is not fixed: PDGFA and BMP1 are growth-factor and matrix genes (the amplified 4q12 locus carries **PDGFRA**, not PDGFA at 7p22, a correction¹); KIF15 is a mitotic kinesin; TRAF3 fails Benjamini–Hochberg correction ($q = 0.056$). The immune result is least interpretable — all 11 features were elevated in the high-risk group, and our marker-gene Z-score is not a deconvolution estimate²¹ — and the mTORC2 arm is weakest-supported, untested^{6,28–30}. The secondary characterisation of the Model A genes — k-means subtypes, immune feature scores and mutation profiling — is reported in Supplementary Note S2.

Limitations

Six limitations are load-bearing; the full list is in Supplementary Note S1. The four genes were selected on all 510 discovery patients, so the apparent 0.794 and cross-validated 0.796 retain optimism, the nested 0.731 being conservative and 0.731–0.769 the external range. Compression cost specificity: MRG-4 is more proliferation-correlated ($\rho = 0.67$) than the full-panel index ($\rho = 0.26$), and cross-species direction is reversed ($\rho = -0.32$). Reclassification is not improved, and the index is never compared with a published LGG signature. The within-cohort slope is 0.548 (95% CI 0.400–0.697): probabilities are not absolute risks. Discovery and validation systems are mismatched — mouse-macrophage panel, human bulk glioma target, fitted genes that are not the pathway effectors, and no tumour-versus-normal comparison. Follow-up is immature (125 deaths; median 27.8 months), and the external cohort is small

with an event rate (52.3%) over twice discovery (24.5%). What reproduces is rank ordering, not absolute transportability.

Conclusion

A mitoxyperilysis-related panel yields a lower-grade glioma index: 25 genes with an EPV of 5.0, an apparent C-index of 0.821 and an external C-index of 0.664, or, with stability selection, four genes (PDGFA, KIF15, BMP1, TRAF3) with an EPV of 31.2, an external C-index of 0.769 and independence from IDH status and grade. The score marks a proliferation-high, matrix-remodelling LGG state, not the mitoxyperilysis pathway; its gain is discrimination and ranking consistency, not clinical utility. Priority does not belong to this study ¹¹; the contribution is parsimony, external reproducibility, and showing these differ from a cross-validated C-index.

Declarations

Ethics approval and consent to participate. Not applicable: only publicly available, de-identified human transcriptomic data (TCGA-LGG, CGGA-325) were analysed, generating no new material and requiring no individual consent.

Consent for publication. Not applicable.

Availability of data and materials. All primary data are public: TCGA-LGG from cBioPortal (study *lgg_tcga_pan_can_atlas_2018*), CGGA-325 from the Chinese Glioma Genome Atlas (<http://www.cgga.org.cn>), and single-cell data from 21722308-5091-4b63-9c07-4f116ab9a7b3 (SL040) in CZ CELLxGENE Discover. Derived data accompany the manuscript as Supplementary Tables S1–S30, Supplementary Figures S1–S5 and Supplementary Notes S1–S2 (per-patient scores S4, S5; extended CGGA-325 adjusted models S15; MRG-4 mutation analysis S9b; immune feature scores S26), with a SHA-256 manifest covering every deposited table.

Availability of code. The analysis code and the deposited derived tables are openly available at <https://github.com/speetle/lgg-mrg4-mitoxyperilysis> (release v1.0.0): code under the MIT licence, derived tables under CC BY 4.0, with a SHA-256 checksum manifest covering every deposited table.

Competing interests. The authors declare no competing interests.

Funding. No specific grant.

Authors' contributions (CRediT). Huiling Zhai: investigation, formal analysis, writing – review and editing. Jingyi Chen: investigation, formal analysis, writing – review and editing. Bin Lian: conceptualisation, methodology, software, formal analysis, investigation, data curation, writing – original draft, writing – review and editing; guarantor. Yanlin Qiu, Yingxin Wu, Yiyi Zhong, Yini Sun, Jiayi Kuang,

Jiawei Zheng and Jinyuan Liu: investigation, writing — review and editing. All authors approved the final manuscript.

Acknowledgements. Not applicable.

Declaration of generative AI and AI-assisted technologies. An AI-assisted tool was used for language editing; the authors have reviewed the content and take full responsibility.

Declaration on figure provenance. All figures are vector graphics generated from the analysed data with Matplotlib; no generative AI image models were used.

References

1. Cancer Genome Atlas Research Network (2015) Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med* 372(26):2481–2498 PMID: 26061751. 10.1056/NEJMoa1402121
2. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D et al (2021) The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol* 23(8):1231–1251 PMID: 34185076. 10.1093/neuonc/noab106
3. Weller M, van den Bent M, Preusser M, Le Rhun E, Tonn JC, Minniti G et al (2021) EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat Rev Clin Oncol* 18(3):170–186 PMID: 33293629. 10.1038/s41571-020-00447-z
4. Zheng Y, Ji Q, Xie L, Wang C, Yu CN, Wang YL et al (2021) Ferroptosis-related gene signature as a prognostic marker for lower-grade gliomas. *J Cell Mol Med* 25(6):3080–3090 PMID: 33594759. 10.1111/jcmm.16368
5. Sun P, Wang X, Zhong J, Yu D, Xuan H, Xu T et al (2023) Development and validation of a pyroptosis-related genes signature for risk stratification in gliomas. *Front Genet.* ;14:1087563. PMID: 36861130. 10.3389/fgene.2023.1087563
6. Wang Y, Lu J, Carisey AF, Chadchan SB, Lee HW, Malireddi RKS et al (2025) Innate immune and metabolic signals induce mitochondria-dependent membrane lysis via mitoxyperiosis. *Cell.* ;188(25):7155–7174.e25. PMID: 41317732. 10.1016/j.cell.2025.11.002
7. Ke FS, Holloway S, Uren RT, Wong AW, Little MH, Kluck RM et al (2022) The BCL-2 family member BID plays a role during embryonic development in addition to its BH3-only protein function by acting in parallel to BAX, BAK and BOK. *EMBO J* 41(15):e110300 PMID: 35758142. 10.15252/embj.2021110300
8. Ramos S, Hartenian E, Broz P (2024) Programmed cell death: NINJ1 and mechanisms of plasma membrane rupture. *Trends Biochem Sci* 49(8):717–728 PMID: 38906725. 10.1016/j.tibs.2024.05.007
9. Chen Z, Xu D (2026) Mitochondria at the membrane provide a route to inflammatory cell death. *Cell Metab* 38(2):260–262 PMID: 41638191. 10.1016/j.cmet.2025.12.019

10. Al-Zidan R, Gautam M, Man SM (2026) Mitoxyperilysis: fasting-induced cell death in immunometabolism and disease. *Trends Biochem Sci* 51(4):313–315 PMID: 41856854. 10.1016/j.tibs.2026.01.005
11. Zhang Y, Sun B, Lin J, Ding C, Zhou X, Xu X et al (2026) A mitoxyperilysis-related signature stratifies prognosis and identifies an aggressive colorectal cancer ecosystem with immune remodeling. *Front Cell Dev Biol.* ;14:1851988. PMID: 42491158. 10.3389/fcell.2026.1851988
12. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W et al (2009) IDH1 and IDH2 mutations in gliomas. *N Engl J Med* 360(8):765–773 PMID: 19228619. 10.1056/NEJMoa0808710
13. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA et al (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2(5):401–404 PMID: 22588877. 10.1158/2159-8290.CD-12-0095
14. Zhao Z, Zhang KN, Wang Q, Li G, Zeng F, Zhang Y et al (2021) Chinese Glioma Genome Atlas (CGGA): A Comprehensive Resource with Functional Genomic Data from Chinese Glioma Patients. *Genomics Proteom Bioinf* 19(1):1–12 PMID: 33662628. 10.1016/j.gpb.2020.10.005
15. Harrell FE Jr, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 15(4):361–387 PMID: 8668867. 10.1002/(sici)1097-0258(19960229)15:4%3C361::aid-sim168%3E3.0.co;2-4
16. Vittinghoff E, McCulloch CE (2007) Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol* 165(6):710–718 PMID: 17182981. 10.1093/aje/kwk052
17. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 49(12):1373–1379 PMID: 8970487. 10.1016/s0895-4356(96)00236-3
18. Tibshirani R (1997) The lasso method for variable selection in the Cox model. *Stat Med* 16(4):385–395 PMID: 9044528. 10.1002/(sici)1097-0258(19970228)16:4%3C385::aid-sim380%3E3.0.co;2-3
19. Meinshausen N, Bühlmann P, *Stat Soc Series JR, B Stat Methodol* (2010). ;72(4):417–473. 10.1111/j.1467-9868.2010.00740.x. (Not indexed in PubMed; verified against Crossref.)
20. Gobin M, Nazarov PV, Warta R, Timmer M, Reifemberger G, Felsberg J et al (2019) A DNA Repair and Cell-Cycle Gene Expression Signature in Primary and Recurrent Glioblastoma: Prognostic Value and Clinical Implications. *Cancer Res* 79(6):1226–1238 PMID: 30674534. 10.1158/0008-5472.CAN-18-2076
21. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF et al (2009) Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462(7269):108–112 PMID: 19847166. 10.1038/nature08460
22. CZI Cell Science Program, Abdulla S, Aevermann B, Assis P, Badajoz S, Bell SM et al (2025) CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. *Nucleic Acids Res.* ;53(D1):D886-D900. PMID: 39607691. 10.1093/nar/gkae1142
23. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J et al (2023) TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. *Nucleic Acids Res.*

- ;51(D1):D1425-D1431. PMID: 36321662. 10.1093/nar/gkac959
24. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS et al (2019) PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 170(1):51–58 PMID: 30596875. 10.7326/M18-1376
 25. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P et al (2025) PROBAST + AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ* 388:e082505 PMID: 40127903. 10.1136/bmj-2024-082505
 26. Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. ;350:g7594. PMID: 25569120. 10.1136/bmj.g7594
 27. Mossi Albiach A, Janusauskas J, Kjaer Jacobsen J, Kapustová I, Karamzadeh R, Kvedaraite E et al (2026) Tissue-reactive zonation of mesenchymal-like phenotypes in human glioblastoma. *Med*. ;7(8):101228. PMID: 42492522. 10.1016/j.medj.2026.101228
 28. Iskandar K, Foo J, Liew AQX, Zhu H, Raman D, Hirpara JL, Autophagy et al (2024) ;20(6):1418–1441. PMID: 38261660. 10.1080/15548627.2024.2307224
 29. Martinez-Lopez N, Mattar P, Toledo M, Bains H, Kalyani M, Aoun ML et al (2023) mTORC2-NDRG1-CDC42 axis couples fasting to mitochondrial fission. *Nat Cell Biol* 25(7):989–1003 PMID: 37386153. 10.1038/s41556-023-01163-3
 30. Wang Y, Tian Q, Hao Y, Yao W, Lu J, Chen C et al (2022) The kinase complex mTORC2 promotes the longevity of virus-specific memory CD4(+) T cells by preventing ferroptosis. *Nat Immunol* 23(2):303–317 PMID: 34949833. 10.1038/s41590-021-01090-1

Tables

Table 1. Baseline characteristics of the discovery and validation cohorts.

Characteristic	TCGA-LGG (all, n = 510)	TCGA-LGG (complete annotation, n = 502)	CGGA-325, WHO II–III (n = 172)
Age, median (range), years	41 (14–87)	41 (14–87)	39 (10–74)
Patients aged < 18 years, n	2	2	2
Male sex, n (%)	283 (55.5)	279 (55.6)	106 (61.6)
WHO grade 2 / 3, n	246 / 263	241 / 261	98 / 74
IDH-mutant / wildtype / unannotated, n	—	410 / 92 / 0	127 / 44 / 1
IDHmut-codel / IDHmut-non- codel / IDHwt, n	—	165 / 245 / 92	—
Deaths, n (%)	125 (24.5)	123 (24.5)	90 (52.3)
Median follow-up (reverse KM), months	27.8	27.8	130.2
Median OS, low- vs high-risk MRG-4 median split, months	117.4 vs 51.9	130.8 vs 50.9	not reached vs 31.3

Grade was not available for 1 patient and IDH status/subtype for 8 patients in TCGA-LGG. In CGGA-325 one of the 172 analysed patients (CGGA_1046) has no IDH annotation. That patient is retained in all whole-cohort analyses and in the median-split survival analysis, and is excluded only from the two IDH-stratified rows of Supplementary Table S29, so those rows are based on 127 and 44 patients respectively rather than on 128 and 44. Patients are the unit of analysis; one primary tumour sample per patient. The median-OS row uses each cohort's own MRG-4 median split, and is given separately for the all-patient and complete-annotation columns since the two splits differ. Two patients in each cohort are under 18 years (youngest 14 and 10 years); removing them changes the external C-index from 0.769 to 0.770 and leaves every conclusion unchanged.

Table 2. Cox regression for overall survival. All models are fitted in the discovery cohort (TCGA-LGG) unless stated otherwise. Blocks 1–4 are **MRG-4 (Model B)** and blocks 5–8 are **Model A (25 genes)**. Every risk-score hazard ratio is per 1 SD of that model's own score as standardised in the analysis set of the model concerned; the SD is 0.903 for MRG-4 (n = 510) and 1.062 for Model A (n = 502) in TCGA-LGG, so a row labelled "per SD" does not refer to the same physical increment across blocks. Model 1: score with age and sex (n = 510). Model 2: the same model plus IDH status and WHO grade in patients with complete annotation (n = 502). Model 3 (Model A only): binary risk group with age, sex, IDH status and grade (n = 502). Model 1 is tabulated in full for Model A and by its score term only for MRG-4, where that term is the unadjusted per-SD estimate; the age and sex terms of the MRG-4 Model 1 are not tabulated. Models 5 and 6: the index against a 20-gene proliferation score, without and with clinical covariates, on the same 502 patients so that the two indices are compared on identical subsets. External-validation models for both scores are plotted in Figure 2C and deposited as Supplementary Tables S14 and S15.

Block	Model	Variable	HR	95% CI	P
<i>MRG-4 (four genes, Model B)</i>					
1	1	MRG-4 score (per SD)	2.467	2.097–2.903	1.4×10^{-27}
1	2	MRG-4 score (per SD)	1.563	1.231–1.985	2.5×10^{-4}
1	2	Age (per year)	1.044	1.028–1.061	6.2×10^{-8}
1	2	Male sex	1.196	0.825–1.733	0.346
1	2	IDH-wildtype	2.349	1.381–3.996	0.0016
1	2	WHO grade 3	1.627	1.046–2.531	0.031
2	5	MRG-4 score (per SD)	2.304	1.864–2.847	1.1×10^{-14}
2	5	Proliferation score (per SD)	1.117	0.897–1.391	0.321
2	6	MRG-4 score (per SD)	1.488	1.125–1.970	0.0054
2	6	Proliferation score (per SD)	1.083	0.857–1.368	0.507
2	6	Age (per year)	1.044	1.028–1.060	6.9×10^{-8}
2	6	Male sex	1.205	0.830–1.749	0.328
2	6	IDH-wildtype	2.404	1.410–4.100	0.0013
2	6	WHO grade 3	1.563	0.987–2.473	0.057
<i>Model A (25 genes)</i>					
3	1	Model A score (per SD)	2.698	2.258–3.224	8.1×10^{-28}
3	1	Age (per year)	1.047	1.031–1.064	4.5×10^{-9}
3	1	Male sex	1.184	0.823–1.703	0.362

3	2	Model A score (per SD)	2.348	1.857– 2.968	1.0×10^{-12}
3	2	Age (per year)	1.045	1.029– 1.062	3.5×10^{-8}
3	2	Male sex	1.237	0.855– 1.788	0.259
3	2	IDH-wildtype	1.357	0.808– 2.281	0.249
3	2	WHO grade 3	1.421	0.911– 2.216	0.122
3	3	Risk group (high vs low)	1.938	1.218– 3.084	0.0053
3	3	Age (per year)	1.048	1.032– 1.064	3.8×10^{-9}
3	3	Male sex	1.231	0.848– 1.788	0.275
3	3	IDH-wildtype	2.991	1.892– 4.728	2.8×10^{-6}
3	3	WHO grade 3	1.941	1.286– 2.928	0.0016
3	5	Model A score (per SD)	2.694	2.212– 3.282	7.0×10^{-23}
3	5	Proliferation score (per SD)	1.166	0.972– 1.400	0.099
<i>Unadjusted single covariates (n = 502)</i>					
4	—	IDH-wildtype, unadjusted	6.665	4.587– 9.684	2.5×10^{-23}
4	—	WHO grade 3, unadjusted	3.324	2.238– 4.938	2.7×10^{-9}
4	—	Proliferation score (per SD), unadjusted	1.840	1.555– 2.176	1.1×10^{-12}
4	—	MRG-4 score (per SD), unadjusted	2.467	2.097– 2.903	1.4×10^{-27}
4	—	Model A score (per SD), unadjusted	2.927	2.467– 3.472	8.0×10^{-35}

Table 2 reports both models together. Values were recomputed from the deposited fits: the adjusted MRG-4 P value in Model 2 is 2.46×10^{-4} and the WHO grade III P value in the CGGA-325 adjusted model is 8.54×10^{-5} , the same value shown in Figure 2C. The unadjusted proliferation hazard ratio is 1.840

(1.555–2.176). The Model A and MRG-4 covariate rows of every block, and the whole of Model 3, were recomputed from the deposited per-patient scores and reproduce the tabulated hazard ratios, confidence intervals and *P* values to the digits shown (Supplementary Table S25).

Table 3. Head-to-head performance of Model A (25 genes), the 26-gene LASSO model, and MRG-4. C-index values are Harrell's C. External values are in CGGA-325 WHO II–III; the primary column is cohort-internal standardisation, and the alternative conventions are given in Supplementary Table S30. The two cross-validation rows are different quantities and must not be compared across rows: the first fixes the gene set before the folds are formed, the second repeats the entire selection inside every fold (Section 2.4). The time-dependent AUC rows use the censoring distribution estimated in the cohort in which the AUC is evaluated, not in the discovery cohort; both censoring conventions are tabulated at all 19 horizons in Supplementary Table S23. The two TCGA apparent C-indices for the models this paper reports are rebuildable from the deposited per-patient scores (0.8211 for the 25-gene index from Supplementary Table S2; 0.7938 for MRG-4 from Supplementary Table S4). † The 26-gene LASSO apparent C-index is the one value in this table that is not rebuildable from the deposited objects, because the 26-gene coefficient vector at the 1-SE α was not retained; it is shown for context, its cross-validated (0.777) and external (0.726) counterparts are rebuildable and are the values used in the text, and the same value is printed in Figure 2B.

Metric	Model A (25 genes)	LASSO (26 genes)	MRG-4 (4 genes)
Events per fitted term	5.0 (25 terms)	4.8 (26 terms)	31.2 (4 terms)
Coefficient sign flips	10 / 25	—	0 / 4
TCGA apparent C-index	0.821	0.870 †	0.794
TCGA 10-fold CV, gene set fixed before folding	0.796	0.777	0.796
TCGA 10-fold CV, selection repeated inside folds	0.832 ± 0.021	—	0.731 ± 0.023
CGGA-325 external C-index (cohort z)	0.664 (0.610–0.715)	0.726	0.769 (0.722–0.812)
CGGA-325 external C-index (discovery-SD scaling)	0.620	—	0.731 (0.682–0.776)
CGGA-325 external C-index (rank-inverse normal)	0.676	—	0.744
Difference in external C-index against Model A (paired bootstrap)	reference	—	+0.105 (0.054–0.162), <i>P</i> < 0.0001
Within-cohort regression slope in CGGA-325	—	—	0.548 (0.400–0.697); equals the cohort's own per-SD log HR, so it does not test transfer
Integrated Brier score, 6–60 months (MRG-4 vs the 25-gene index; lower is better)	—	—	0.287 vs 0.316
Randomised-panel benchmark, penalisation matched	97.8% of valid random 25-gene panels match or beat it (null mean 0.731)	—	0.68% of valid random 4-gene panels match or beat it (null mean 0.638)
Leave-one-gene-out, external C-index range	—	—	0.722–0.751 (all four = 0.769)
CGGA-325 partial log-likelihood	–423.59	—	–404.09
CGGA-325 AIC (1 covariate)	849.18	—	810.18
Time-dependent AUC, mean over the 19 horizons of 12–120 mo weighted by the Kaplan–Meier survival decrement, cohort z (unweighted mean 0.717 vs 0.834)	0.731	—	0.845

Difference in time-dependent AUC against Model A, paired bootstrap (2,000 resamples)	reference	—	+0.114 (0.054–0.177) ; the per-time-point 95% interval excludes zero at 18 of the 19 time points, one point (18 months) excepted
Difference in time-dependent AUC against the four-variable clinical model, paired bootstrap	—	—	+0.045 (–0.018–0.111), not established at this sample size
Continuous NRI, 3-year mortality (n = 160, 50 events)	reference	—	–0.291 (–0.625 to 0.098); +0.440 when the coefficients are refitted in the validation cohort
IDI, 3-year mortality	reference	—	–0.047
Category-based NRI (tertiles), 3 years	reference	—	+0.429
Net benefit, 3 years, n = 172 (thresholds 0.20–0.80)	higher than MRG-4 at 0/31 thresholds; strictly highest at 3 (0.76–0.80)	—	higher than the 25-gene index at 20/31 (0.20–0.50, 0.54–0.60); strictly highest at 7; the four-variable clinical model is strictly highest at 21/31

Figures

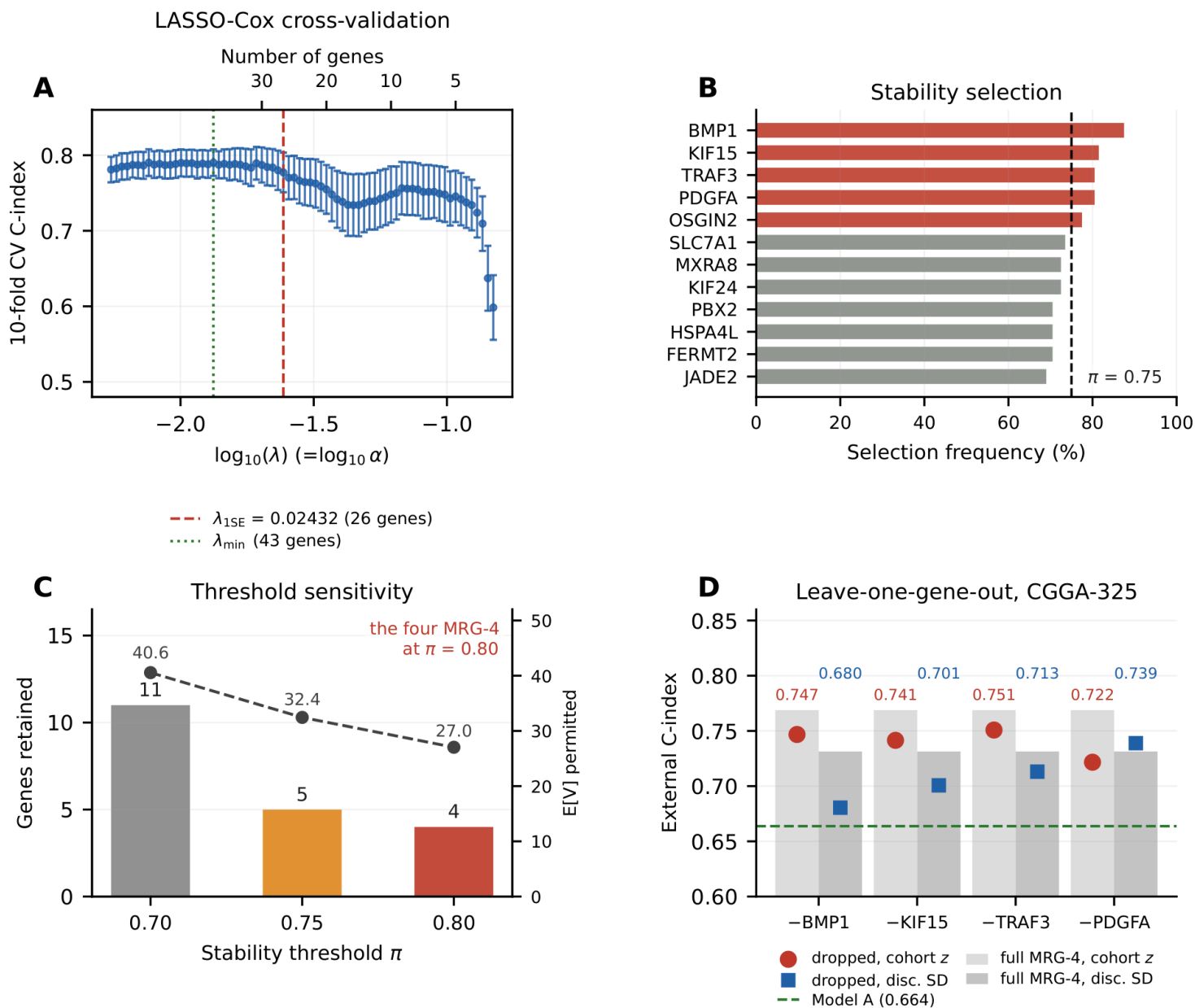


Figure 1

Derivation of the four-gene score, MRG-4. (A) LASSO-Cox 10-fold cross-validation curve: mean C-index ($\pm$ SE) against $\log_{10}(\alpha)$, with the number of genes in the model on the upper axis; dotted line, λ_{min} (43 genes); dashed line, λ_{1SE} (26 genes). The plateau shows that model size is poorly determined by the penalty path alone. (B) Stability selection: per-gene selection frequency across 200 bootstrap resamples at λ_{1SE} , original run. Red bars exceed the specified $\pi = 0.75$ threshold (dashed line) and define the five-gene set, from which OSGIN2 was subsequently removed for non-significance. The independent replication run is tabulated in Supplementary Table S6 and is not plotted here. (C) Sensitivity to the stability threshold. Bars, number of genes retained at $\pi = 0.70, 0.75$ and 0.80 ; the four-gene set is returned by $\pi = 0.80$ in the original run only, and no elimination step is applied at that threshold. Line (right axis), the Meinshausen–Bühlmann bound $E[V] \leq q^2 / ((2\pi - 1)p)$ with $q = 38$ (the largest selected-gene count over the 200 resamples) and $p = 89$ candidates, which is the number of false selections that

threshold would permit; the bound is non-binding at every threshold shown, so the threshold itself carries no useful false-discovery control. (D) Leave-one-gene-out. Bars, the full-model external C-index under each of the two standardisation conventions (cohort z 0.769; discovery-SD scaling 0.731); points, the external C-index with each gene removed in turn. The ranking of the four genes is convention-dependent – PDGFA is the most important gene under cohort z (0.722) and the least important under discovery-SD scaling (0.739) – so no claim is made about the relative contribution of individual genes. Dashed line, the 25-gene comparator at 0.664. Removing TRAF3 gives the smallest change (0.751).

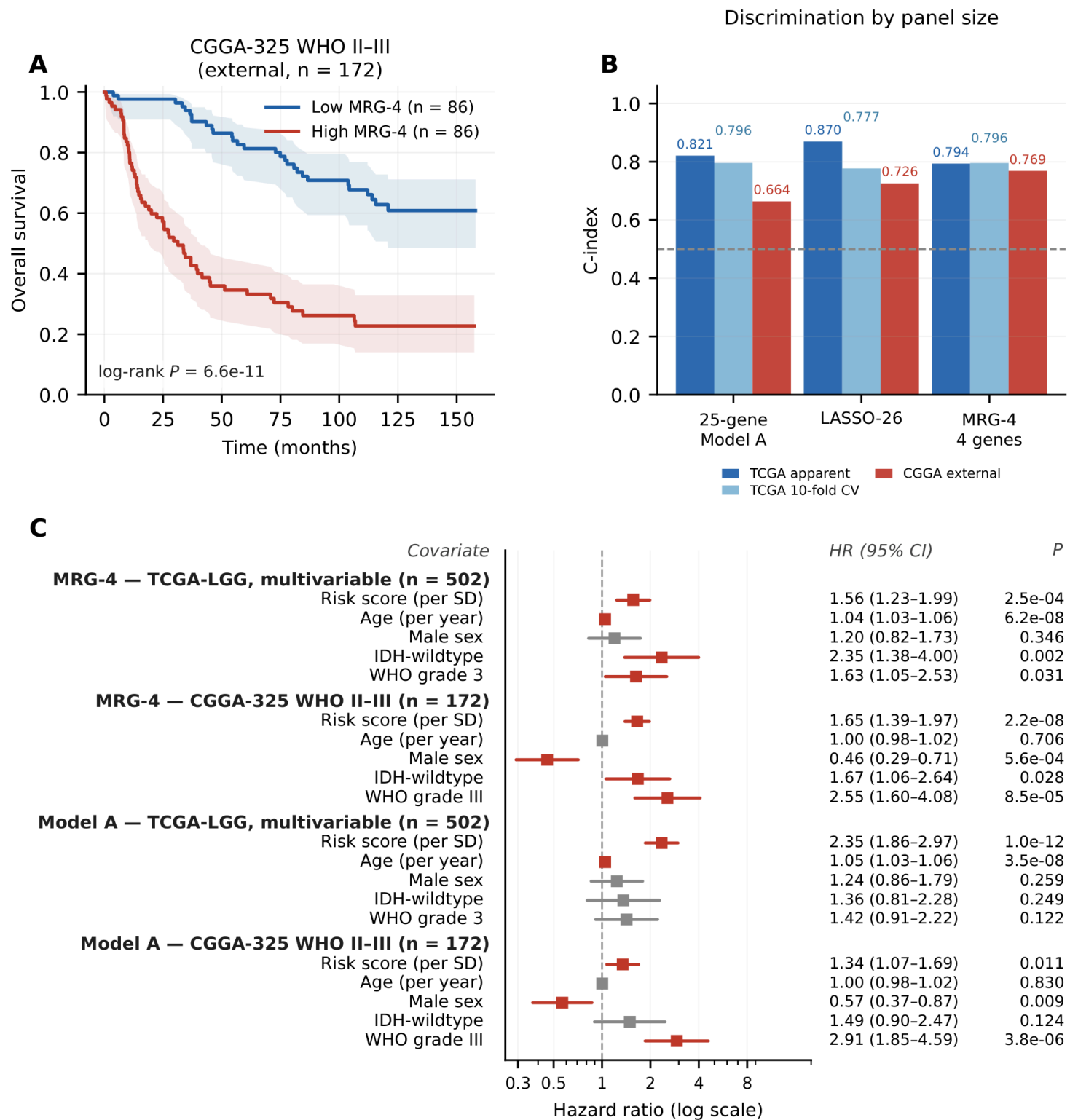


Figure 2

External validation of MRG-4 and the cost of panel size. (A) Kaplan–Meier overall survival by MRG-4 median split in CGGA-325 WHO grade II–III (n = 172, 90 deaths); median overall survival not reached vs 31.3 months; log-rank P = 6.61×10^{-11} . (B) C-index for the 25-gene model, the 26-gene LASSO model and MRG-4 across apparent (TCGA, n = 510), 10-fold cross-validated (TCGA, gene set fixed before folding) and external (CGGA-325) evaluation. Internal cross-validation does not separate the models (0.796 vs 0.777 vs 0.796); external validation does (0.664 vs 0.726 vs 0.769). (C) Adjusted models for both scores. Forest plot of hazard ratios (95% CI, log scale) for MRG-4 and for Model A, each in TCGA-LGG (n = 502, 123 deaths) and in CGGA-325 WHO II–III (n = 172, 90 deaths), with the covariates age, sex, IDH status and grade shown for each block. MRG-4, IDH-wildtype status and grade are all independently significant in both cohorts; for Model A neither IDH status nor grade is significant in the discovery cohort. The CGGA-325 blocks are also deposited as Supplementary Table S14, with each score's SD stated so that the scaling of every hazard ratio is unambiguous.

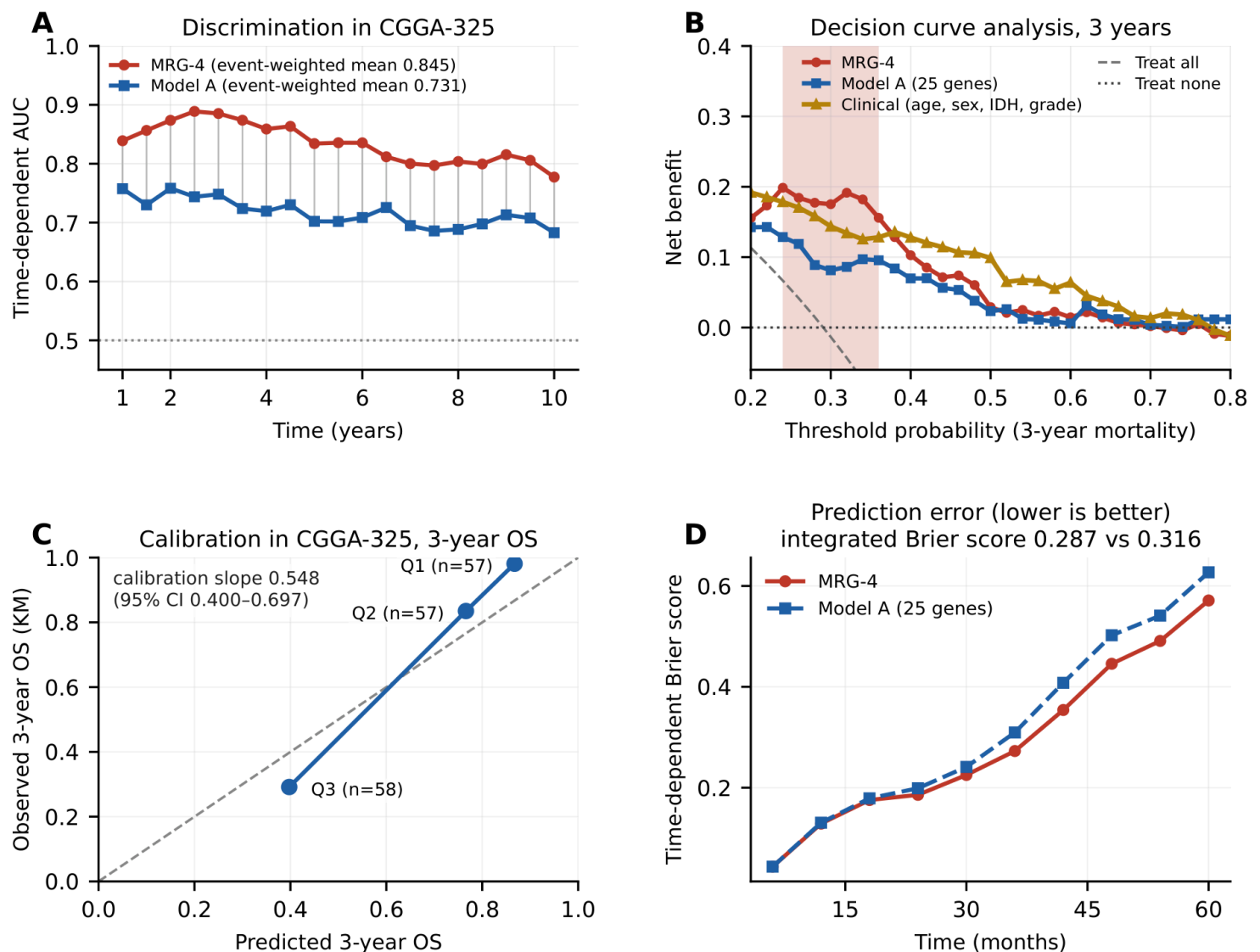


Figure 3

Head-to-head comparison, calibration and prediction error. (A) Time-dependent AUC from 12 to 120 months in CGGA-325; mean weighted by the Kaplan–Meier survival decrement at each horizon, 0.845 for MRG-4 versus 0.731 for the 25-gene model (unweighted means 0.834 and 0.717); the weights are the survival decrements $\Delta S(t)$ at the 19 horizons, not the numbers of events, with the censoring distribution estimated in CGGA-325 itself. A paired bootstrap over patients puts the mean difference at +0.114 (95% CI 0.054–0.177), and the per-time-point interval excludes zero at 18 of the 19 horizons. Both censoring conventions are tabulated at all 19 horizons in Supplementary Table S23; every value quoted here and in Figure 3A uses the cohort-evaluated convention. (B) Decision curve analysis for 3-year mortality in the 172 validation patients, with all four comparators computed in the same 172 patients. MRG-4 has the higher net benefit than the 25-gene model at 20 of the 31 thresholds (0.20–0.50 and 0.54–0.60); the 25-gene model is strictly highest at only three, 0.76, 0.78 and 0.80, where every strategy is below 0.012. The four-variable clinical model (age, sex, IDH status, grade) has the highest net benefit at 21 of the 31 thresholds, from 0.38 to 0.74 and at 0.20–0.22; MRG-4 is ahead only across 0.24–0.36, shaded. (C) Calibration at three years by tertile of the cohort-standardised index: predicted 3-year survival (from a baseline hazard re-estimated in the validation cohort) against Kaplan–Meier observed survival, with the diagonal marking perfect calibration. Predicted against observed values are Q1 0.867/0.981, Q2 0.766/0.835 and Q3 0.397/0.292; the within-cohort regression slope is 0.548 (95% CI 0.400–0.697); being computed on the cohort-standardised index, it equals the cohort's own unadjusted per-SD log hazard ratio and does not test transfer of the frozen model (Section 2.7). (D) Time-dependent Brier score for MRG-4 and for the 25-gene model, both with the baseline hazard re-estimated in the validation cohort so that the comparison is on the same footing; the integrated Brier score from 6 to 60 months is 0.287 against 0.316.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.zip](#)
- [Supplementary.docx](#)