Analytic approaches to clinical validation of results from preclinical models of glioblastoma: a systematic review

Beth Fitt[1¶], Grace Loy[1¶], Edward Christopher[2], Paul M. Brennan[3,4,5], Michael T.C. Poon[3,4,5,6*,#a]

[¶]These authors contributed equally to this work

[1] Edinburgh Medical School, University of Edinburgh, Edinburgh, UK

[2] Royal Infirmary of Edinburgh, NHS Lothian, Edinburgh, UK

[3] Brain Tumour Centre of Excellence, Cancer Research UK Edinburgh Centre, Edinburgh, UK

[4] Translational Neurosurgery, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

[5] Department of Clinical Neuroscience, Royal Infirmary of Edinburgh, Edinburgh, UK

[6] Centre for Medical Informatics, Usher Institute, University of Edinburgh, Edinburgh, UK

[#a] Centre for Medical Informatics, Building Nine, Edinburgh BioQuarter, Edinburgh, EH16 4UX, United Kingdom

* Corresponding author

E-mail: michael.poon@ed.ac.uk

Abstract

Introduction: Analytic approaches to clinical validation of results from preclinical models are important in assessment of their relevance to human disease. This systematic review examined consistency in reporting of glioblastoma cohorts from The Cancer Genome Atlas (TCGA) or Chinese Glioma Genome Atlas (CGGA) and assessed whether studies included patient characteristics in their survival analyses.

Methods: We searched Embase and Medline on 02Feb21 for studies using preclinical models of glioblastoma published after Jan2008 that used data from TCGA or CGGA to validate the association between at least one molecular marker and overall survival in adult patients with glioblastoma. Main data items included cohort characteristics, statistical significance of the survival analysis, and model covariates.

Results: There were 58 eligible studies from 1,751 non-duplicate records investigating 126 individual molecular markers. In 14 studies published between 2017 and 2020 using TCGA RNA microarray data that should have the same cohort, the median number of patients was 464.5 (interquartile range 220.5-525). Of the 15 molecular markers that underwent more than one univariable or multivariable survival analyses, five had discrepancies between studies. Covariates used in the 17 studies that used multivariable survival analyses were age (76.5%), pre-operative functional status (35.3%), sex (29.4%) MGMT promoter methylation (29.4%), radiotherapy (23.5%), chemotherapy (17.6%), IDH mutation (17.6%) and extent of resection (5.9%).

Conclusion: Preclinical glioblastoma studies that used TCGA for validation did not provide sufficient information about their cohort selection and there were inconsistent results. Transparency in reporting and the use of analytic approaches that adjust for clinical variables can improve the reproducibility between studies.

Keywords

GBM, animal models, cell lines, laboratory, survival analysis

Introduction

Glioblastoma, the most common primary brain cancer, is a fatal disease with patients' median survival of 6-8 months [1,2]. Novel therapies from translational research are desperately needed because current therapeutic options have only a modest and temporary impact on survival [3,4]. Discovery science has advanced our understanding of cancer cell biology and is a step towards developing novel therapies [5]. These discoveries are usually based on preclinical models, from which the relevance to human disease must be established. Demonstrating relevance requires quality clinical and biological data. The Cancer Genome Atlas (TCGA) [6] and the Chinese Glioma Genome Atlas (CGGA) [7] are two open-access resources from which laboratory scientists can interrogate human data to verify their findings in preclinical glioblastoma research. These resources are valuable for the molecular characterisation of glioblastoma and can also be used to examine the associations between molecular markers of interest and survival. An association with survival might implicate a molecular marker as a potential drug target.

Survival analyses using only genomic data are unlikely to have adequate clinical relevance because clinical factors also affect survival. An imbalance of clinical characteristics between comparison groups can confound the association between the molecular marker and survival. Univariable survival analyses that take on only one molecular marker do not account for other markers or clinical characteristics [8]. The resulting associations from such analyses are subjected to confounding effects, which may render them unreliable. Confounding is a fundamental issue that affects observational health-related research, and it should be controlled for when possible [9]. Multivariable analyses are methods to control for confounders and are, therefore, preferable. Open access policies for data and code sharing should facilitate the re-use of data and reproducibility of results [10]. Transparent and detailed reporting of the analytic approach is crucial for replicability and comparison of analyses. These methodological aspects can ensure the science that progresses to clinical trials is well-founded.

Clinical validation of results from preclinical glioblastoma studies using TCGA or CGGA data represents a common experimental step to substantiate research findings. This systematic review examined these studies for their consistency in reporting of cohorts from TCGA and CGGA and whether they included patient characteristics in their survival analyses.

Methods

Eligibility criteria

This review included studies that used data from TCGA or CGGA to examine the association between at least one molecular marker and overall survival in adult patients aged ≥18 years diagnosed with non-recurrent histopathologically confirmed glioblastoma. Studies using any molecular data type from TCGA or CGGA were eligible. Studies using both TCGA and CGGA were eligible if they had separately reported results for TCGA and CGGA. We only included studies that used cell or animal models to first identify molecular markers associated with tumour biology, then examined the association between these markers and overall survival in humans using TCGA or CGGA data. We excluded case reports, reviews, editorials and conference abstracts (S1 Supporting Information).

Study selection

We searched Embase and Medline on 02 February 2021 for potentially eligible studies published after January 2008 using search terms relating to “glioma”, “survival”, “TCGA” and “CGGA” (S2 Supporting Information). The lower limit of the search period was set because data from TCGA first became available in 2008. After removing duplicate studies, two independent reviewers (B.F. and G.L.) performed screening using titles and abstracts followed by full-text eligibility assessment. Any disagreements at each stage were resolved through discussion with a third reviewer (M.T.C.P.).

Data extraction and data items

Two reviewers (B.F. and G.L.) independently collected data from each study using the online systematic review management software Covidence (Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org). Disagreements were resolved by discussion between the two reviewers or by involving a third reviewer (M.T.C.P.). Data items included study characteristics, TCGA cohort characteristics, CGGA cohort characteristics, genomic data used, molecular markers, and details of survival analysis. Molecular markers included expression, variants, or methylation of genes, RNAs and microRNAs. A set of molecular markers was defined by the analysis of >1 molecular markers together.

Each study can report results from multiple survival analyses using the overall cohort or specific subgroups (S1 Figure). We collected information on all survival analyses performed in the studies. We categorised survival analysis into univariable and multivariable analysis, and we collected the covariates entered into the multivariable analysis. To describe the association between molecular markers and survival, we considered the reported p value of <0.05 as statistical significance. If a study reported results from both TCGA and CGGA cohorts, we extracted the statistical significance of these results separately. Data on effect sizes and their corresponding 95% confidence intervals (CI) were

not collected because studies using log-rank (Mantel-Cox) tests to compare survival between study-specific groups do not provide these data and there was no plan for meta-analysis.

Quality assessment

There was no risk of bias assessment tool directly relevant to studies in this review. However, we assessed components of the study design relating to risk of bias. These measures of quality included types and size of cohorts used for survival analysis, types of genomic data used from TCGA or CGGA, and the criteria used to select patients for survival analysis. We did not quantify the quality of study based on risk of bias items because this review aimed to assess the reporting and approach to analyses rather than to summarise effect sizes.

Summary statistics

We presented study characteristics, results and quality measures using descriptive statistics with stratification by type of survival analysis, univariable and multivariable, where available. The availability of data in TCGA increased over time and there are different numbers of patients in whom various types of data are available. To assess the reproducibility of cohort selection from TCGA, we summarised the number of patients in studies published between 2017-2020 using TCGA RNA microarray. These studies should have the same number of patients because they all used the same RNA microarray dataset from TCGA when there was no further accrual of patients. There were occasions when two or more survival analyses within or between studies investigated the association between a molecular marker and survival. We presented findings on these molecular markers that underwent two or more analyses to demonstrate the consistencies of results. There was no meta-analysis of any association between molecular markers and overall survival.

Results

Study characteristics

This review included 58 eligible studies from 1,751 non-duplicate records retrieved from our systematic search (Fig 1 and S1 References). Individual study characteristics are presented in S1 Table. These studies investigated 126 individual molecular markers and 32 sets of molecular markers. Most (62.1%) studies were published in 2017-2020 and were from research teams based in the United States (34.5%), China (27.6%) and Europe (24.1%). The pre-clinical glioblastoma models used were cell lines and orthotopic mouse models in 51.7% and 48.3% studies, respectively. All studies used a form of data from TCGA with various combination with other data sources and two studies used data

from CGGA (Table 1). RNA microarray data was the most common data type, used in 45 (77.6%) studies. Three (5.2%) studies did not specify the data type used. Six studies (five using TCGA data and one using both TCGA and CGGA data) did not provide the number of patients included.

Fig 1. PRISMA flowchart of study selection.

Table 1. **Characteristics of 58 included studies that used TCGA or CGGA data to validate findings from experiments using pre-clinical models of glioblastoma**

	Overall N = 58	Survival analysis type	
		Univariable N=41	Multivariable N=17
Year of publication			
2009-2012	4 (6.9%)	2 (4.9%)	2 (11.8%)
2013-2016	18 (31.0%)	13 (31.7%)	5 (29.4%)
2017-2020	36 (62.1%)	26 (63.4%)	10 (58.8%)
Country / region			
United States	20 (34.5%)	13 (31.7%)	7 (41.2%)
Europe (inc. UK)	14 (24.1%)	9 (22.0%)	5 (29.4%)
China	16 (27.6%)	13 (31.7%)	3 (17.6%)
Other countries[a]	8 (13.8%)	6 (14.6%)	2 (11.8%)
Pre-clinical model			
Cell lines	30 (51.7%)	24 (58.5%)	6 (35.3%)
Orthotopic mouse models	28 (48.3%)	17 (41.5%)	11 (64.7%)
Data source			
TCGA only	34 (58.6%)	26 (63.4%)	8 (47.1%)
TCGA & CGGA	1 (1.7%)	0 (0.0%)	1 (5.9%)
TCGA and other public sources	9 (15.5%)	6 (14.6%)	3 (17.6%)
TCGA and own patients	13 (22.4%)	8 (19.5%)	5 (29.4%)
TCGA, CGGA and other public sources	1 (1.7%)	1 (2.4%)	0 (0.0%)
Experimental strategy			
RNA microarray only	27 (46.6%)	24 (58.5%)	3 (17.6%)
RNA sequencing only	7 (12.1%)	4 (9.8%)	3 (17.6%)
miRNA microarray only	2 (3.4%)	1 (2.4%)	1 (5.9%)
RNA microarray and RNA sequencing	4 (6.9%)	3 (7.3%)	1 (5.9%)
RNA microarray and miRNA microarray	10 (17.2%)	6 (14.6%)	4 (23.5%)
RNA sequencing and miRNA microarray	1 (1.7%)	0 (0.0%)	1 (5.9%)
RNA microarray and DNA methylation	1 (1.7%)	0 (0.0%)	1 (5.9%)
RNA sequencing, RNA microarray and miRNA microarray	2 (3.4%)	0 (0.0%)	2 (11.8%)
RNA sequencing, RNA microarray and DNA methylation	1 (1.7%)	0 (0.0%)	1 (5.9%)
Unspecified	3 (5.2%)	3 (7.3%)	0 (0.0%)
Prognostic marker of interest			
One marker only	21 (36.2%)	20 (48.8%)	1 (5.9%)
>1 individual markers	13 (22.4%)	10 (24.4%)	3 (17.6%)
Set(s) of markers only	7 (12.1%)	4 (9.8%)	3 (17.6%)
One marker and set(s) of markers	2 (3.4%)	2 (4.9%)	0 (0.0%)
>1 individual markers and set(s) of markers	10 (17.2%)	4 (9.8%)	6 (35.3%)
One marker and sets of markers with clinical variable(s)	4 (6.9%)	1 (2.4%)	3 (17.6%)
Sets of markers and markers with clinical variable(s)	1 (1.7%)	0 (0.0%)	1 (5.9%)

[a]Other countries included Brazil, Canada, India, Israel, Republic of Korea and Taiwan. UK = United Kingdom; TCGA = The Cancer Genome Atlas; CGGA = Chinese Glioma Genome Atlas; miRNA = micro-RNA.

When investigating the association between their markers of interest from pre-clinical models and survival using genomic data, more studies used univariable survival analyses only (70.7%) compared to those that used multivariable analyses (29.3%). All univariable analyses used the non-parametric log-rank (Mantel-Cox) method and all multivariable analyses used the Cox proportional hazards regression. There were 16 (27.6%) studies that described additional criteria for patient inclusion within the selected TCGA cohort.

Reproducibility and survival analysis

The date and requested data type of query in TCGA can result in a different number of patients available for survival analysis. To assess reproducibility of cohort selection from TCGA in the included studies, we summarised the numbers of patients in studies with similar data specifications. In 14 studies published between 2017 and 2020 using TCGA RNA microarray data without additional patient inclusion criteria, the median number of patients included was 464.5 (interquartile range [IQR] 220.5-525). Of these studies, 12 studies did not perform a multivariable survival analysis, therefore all should have the same number of patients included; the median number of patients included in the univariable survival analysis was 467 (IQR 196.75-528.75).

Among the 126 distinct molecular markers investigated in the included studies, 15 markers underwent more than one univariable or multivariable survival analysis (Table 2). The association of these markers with outcomes were consistent between different analyses most of the time. However, there were discrepancies between results for C-X-C Motif Chemokine Ligan 14 (CXCL14), epidermal growth factor receptor (EGFR), netrin 4 (NTN4), SRY-Box transcription factor 2 (SOX2), serglycin (SRGN) and miRNA-17-5p microRNA (Table 2). These discrepancies appear to relate to the type of survival analysis used (CXCL14, SOX2, SRGN) or the data type (EGFR, NTN4).

Table 2. Results of molecular markers that were reported in two or more separate survival analyses

Molecular marker	Consistency	Author	Data type	Analysis type	Direction of association
CXCL14	✘	Zeng 2018	RNA-Seq, RNA microarray and miRNA microarray	U	▼
				M	□
EGFR	✘	Kuang 2018	RNA microarray only	U	▼
		Li 2018	RNA-Seq only	U	□
HOTAIR	✓	Xavier-Magalhaes 2018	RNA-Seq, RNA microarray and DNA methylation	U	▼
				M	▼
IDO1	✓	Zhai 2017	RNA microarray and RNA-Seq	U	▼
				M	▼
IL-8	✓	Hasan 2019	RNA microarray only	U	▼
				M	▼
MARCKS	✓	Jarboe 2012	RNA microarray and DNA methylation	U	▲
				M	▲
miR-17-5p	✘	Zeng 2018	RNA-Seq, RNA microarray and miRNA microarray	U	▲
				M	□
miR-181d	✓	Genovese 2012	RNA microarray and miRNA microarray	U	□
		Ho 2017	RNA-Seq, RNA microarray and miRNA microarray	U	□
miR-34a	✓	Genovese 2012	RNA microarray and miRNA microarray	U	▼
				M	▼
NTN4	✘	Hu 2012	RNA microarray only	U	▲
		Li 2018	RNA-Seq only	U	□
PD-L1	✓	Nduom 2016	RNA-Seq only	U	▼
				M	▼
POSTN	✓	Mega 2020	RNA microarray only	U	▼
		Liu 2019	RNA microarray and miRNA microarray	U	▼
		Mega 2020	RNA microarray only	M	▼
SFRP1	✓	Delic 2014	RNA microarray and miRNA microarray	U	▲
				M	▲
Sox2	✘	Sathyan 2015	RNA microarray and miRNA microarray	U	▲
				M	□
SRGN	✘	Mega 2020	RNA microarray only	U	▼
				M	□

Consistency refers to the association between a molecular marker and survival being statistically significant in different analyses. Inconsistencies of associations with survival: same analysis type and different data type (EGFR, NTN4) and different analysis type on same data type (CXCL14, miR-17-5p, Sox2, SRGN). Molecular markers ordered alphabetically. Full references available in Supplementary Materials. RNA-Seq = RNA sequencing; û = not consistent between different analyses; ü = consistent between different analyses; U = univariable survival analysis; M = multivariable survival analysis; ▲ = positive association i.e. higher levels of the molecular marker associated with better survival and $p<0.05$; ▼= negative association i.e. lower levels of molecular marker associated with worse survival and $p<0.05$; □ = statistical significance not demonstrated ($p\geq0.05$)

There were 17 studies that investigated the association between their molecular markers of interest and overall survival using a multivariable survival analysis. All these studies used TCGA data, which have clinical data available. The most frequently included clinical variable in the multivariable model was age (76.5%) (Fig 2). Other variables included pre-operative functional status (35.3%), sex (29.4%), MGMT promoter methylation (29.4%), radiotherapy (23.5%), chemotherapy (17.6%), IDH mutation (17.6%) and extent of resection (5.9%).

Fig 2. Clinical variables entered analyses in 17 studies that used a multivariable survival model. Rows represent studies that used a multivariable model for survival analysis (S1 References). Columns are clinical variables relevant to survival in patients with glioblastoma.

Discussion

There were studies in glioblastoma research that used data from publicly available genomic repositories to correlate pre-clinical experimental findings with clinical survival benefit in humans. These studies often had different numbers of patients included despite using the same data source and data type. Survival analyses often did not include other critical clinical variables associated with survival such as extent of resection [11], chemotherapy and radiotherapy [3,12]. In studies that performed a multivariable survival analysis, most clinical variables such as extent of resection and oncological treatment were not included. This yielded some inconsistent results between studies. Other results were subject to confounding effects by clinical variables that were not accounted for.

Reproducibility

Research reproducibility encompasses several aspects: consistent results based on the same data and analysis, consistent results based on the same data but different analyses, consistent results from new data based on previous study design of another study, and consistent results from another study with a similar study design [13,14]. Our review addressed the first two of these aspects. Development of novel cancer therapies relies on reproducible results from preclinical research. The need for improving reproducibility is not new [15]. In cancer research, there is a heavy reliance on the preclinical literature for drug development [16]. However, issues with reporting bias, suboptimal reporting quality, varying reproducibility and preclinical model representation of disease impede the success in finding new therapies [17]. The availability of survival data in publicly available data from cancer genomics programmes presents an opportunity for researchers to assess the association between molecular markers and patient survival in a reproducible manner. These open access data sources provide data on the same cohort of patients, which encourages reproducibility between studies.

However, our findings demonstrate that patient selection was not adequately described, resulting in different numbers of patients between studies that supposedly used the same dataset. There are reproducible ways of querying TCGA data, for example, using the 'TCGABiolinks' R/Bioconductor package [18] where code-based commands can be shared as supplementary materials. Adopting relevant aspects of reporting guidelines such as Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) [19], Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) [20] and REporting recommendations for tumour MARKer prognostic studies (REMARK) [21] can further improve transparency in reporting.

Confounding effects of clinical variables

Confounding is an important consideration in analysing observation data. A confounder can diminish or exaggerate the association between the exposure and the outcome, leading to spurious results [22]. Confounding effects may be controlled by design or by analysis - the latter is most relevant in this review. Control by analysis refers to adopting an analysis method that adjusts for confounders. There are many ways to achieve this, such as stratified and various regression models [9]. The most commonly used multivariable survival analysis is the Cox regression [8]. Most studies in this review did not consider clinical variables as potential confounders to the association between the molecular marker of interest and survival. There are nevertheless examples of associations that no longer exhibit a statistical significance after adjustment to clinical variables in a multivariable analysis (Table 2). Therefore, it is important to explore and consider confounders when assessing the effect of molecular markers on survival [23]. This is not a simple task because of data missingness, relatively small numbers of patients available, as well as correlations between clinical variables. Both data driven and clinically informed choice of covariates would be a reasonable approach [24].

Strengths and limitations

This systematic review assessed all pre-clinical studies that used data from TCGA or CGGA to validate findings from their laboratory experiments. Our data collection allowed comparison of findings between and within studies, which allowed our evaluation of replicability.

Clinical studies that examined associations of previously investigated molecular markers with survival were not included in this review. These studies may provide more detailed descriptions of cohort selection and may be more likely to consider confounding effects from clinical variables. This would mean an overestimation of inconsistencies and suboptimal analytic approaches in our review. However, any omission of consideration about patients being more than their tumours should be highlighted to re-orientate research focus to patient benefits. Collecting data on p values only to denote statistical significance was a pragmatic approach to describing associations reported in the included

studies, since most studies did not report any effect sizes. This does not represent our views on the appropriate statistical approach and reporting of findings. We advocate reporting of effect sizes with their corresponding precision, adjusting for confounders. P values should not be used as a cut-off for the significance of an association [25]. There are other aspects of survival analyses that we did not assess, such as whether included studies tested for the proportional hazard assumption when using a Cox regression [26]. While these analytic procedures are important, reporting of these would not affect our findings. We were unable to perform meta-analyses of the associations between molecular markers and survival because studies were not comparable and there were few effect sizes reported. This limitation prevented us from quantifying the consistency based on heterogeneity and variance measures.

Conclusions

Translational studies in glioblastoma research should increase their transparency to facilitate replicability. The validation of laboratory experimental findings using human data is important to demonstrate translational value; but this should be done with consideration of patient characteristics. Integration of expertise in pre-clinical, genomic and clinical studies may help to address the challenge of producing replicable and meaningful research through collaboration between scientists in different fields.

References

1. Ostrom QT, Patil N, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013-2017. Neuro Oncol. 2020;22: iv1–iv96. doi:10.1093/neuonc/noaa200

2. Poon MTC, Brennan PM, Jin K, Sudlow CLM, Figueroa JD. Might changes in diagnostic practice explain increasing incidence of brain and central nervous system tumors? A population-based study in Wales (United Kingdom) and the United States. Neuro Oncol. 2021;23: 979–989. doi:10.1093/neuonc/noaa282

3. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med. 2005;352: 987–996. doi:10.1056/NEJMoa043330

4. Poon MTC, Sudlow CLM, Figueroa JD, Brennan PM. Longer-term survival in patients with glioblastoma in population-based studies pre- and post-2005: a systematic review and meta-analysis. Sci Rep. 2020;10: 11622. doi:10.1038/s41598-020-68011-4

5. O'Duibhir E, Carragher NO, Pollard SM. Accelerating glioblastoma drug discovery: Convergence of patient-derived models, genome editing and phenotypic screening. Molecular and Cellular Neuroscience. 2017;80: 198–207. doi:10.1016/j.mcn.2016.11.001

6. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155: 462–77. doi:10.1016/j.cell.2013.09.034

7. Zhao Z, Zhang K-N, Wang Q, Li G, Zeng F, Zhang Y, et al. Chinese Glioma Genome Atlas (CGGA): A Comprehensive Resource with Functional Genomic Data from Chinese Glioma Patients. Genomics, Proteomics & Bioinformatics. 2021; S1672022921000450. doi:10.1016/j.gpb.2020.10.005

8. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. Br J Cancer. 2003;89: 431–436. doi:10.1038/sj.bjc.6601119

9. Greenland S, Morgenstern H. Confounding in Health Research. Annu Rev Public Health. 2001;22: 189–212. doi:10.1146/annurev.publhealth.22.1.189

10. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3: 160018. doi:10.1038/sdata.2016.18

11. Molinaro AM, Hervey-Jumper S, Morshed RA, Young J, Han SJ, Chunduru P, et al. Association of Maximal Extent of Resection of Contrast-Enhanced and Non–Contrast-Enhanced Tumor With Survival Within Molecular Subgroups of Patients With Newly Diagnosed Glioblastoma. JAMA Oncol. 2020;6: 495. doi:10.1001/jamaoncol.2019.6143

12. Perry JR, Laperriere N, O'Callaghan CJ, Brandes AA, Menten J, Phillips C, et al. Short-Course Radiation plus Temozolomide in Elderly Patients with Glioblastoma. New England Journal of Medicine. 2017. doi:10.1056/NEJMoa1611977

13. Samsa G, Samsa L. A Guide to Reproducibility in Preclinical Research: Academic Medicine. 2019;94: 47–52. doi:10.1097/ACM.0000000000002351

14. Amaral OB, Neves K. Reproducibility: expect less of the scientific paper. Nature. 2021;597: 329–331. doi:10.1038/d41586-021-02486-7

15. Baker M. 1,500 scientists lift the lid on reproducibility. Nature. 2016;533: 452–454. doi:10.1038/533452a

16. Rubin EH, Gilliland DG. Drug development and clinical trials—the path to an approved cancer drug. Nat Rev Clin Oncol. 2012;9: 215–222. doi:10.1038/nrclinonc.2012.22

17. Begley CG, Ellis LM. Raise standards for preclinical cancer research. Nature. 2012;483: 531–533. doi:10.1038/483531a

18. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Research. 2016;44: e71–e71. doi:10.1093/nar/gkv1507

19. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. The Lancet. 2007;370: 1453–1457. doi:10.1016/S0140-6736(07)61602-X

20. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350: g7594–g7594. doi:10.1136/bmj.g7594

21. Sauerbrei W, Taube SE, McShane LM, Cavenagh MM, Altman DG. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration. JNCI: Journal of the National Cancer Institute. 2018;110: 803–811. doi:10.1093/jnci/djy088

22. Davey Smith G. Data dredging, bias, or confounding. BMJ. 2002;325: 1437–1438. doi:10.1136/bmj.325.7378.1437

23. Wang X, Lin Y, Song C, Sibille E, Tseng GC. Detecting disease-associated genes with confounding variable adjustment and the impact on genomic meta-analysis: With application to major depressive disorder. BMC Bioinformatics. 2012;13: 52. doi:10.1186/1471-2105-13-52

24. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. Br J Cancer. 2003;89: 605–611. doi:10.1038/sj.bjc.6601120

25. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31: 337–350. doi:10.1007/s10654-016-0149-3

26. Austin PC. Statistical power to detect violation of the proportional hazards assumption when using the Cox regression model. Journal of Statistical Computation and Simulation. 2018;88: 533–552. doi:10.1080/00949655.2017.1397151

Supporting information

S1. Supporting information. List of eligibility criteria.

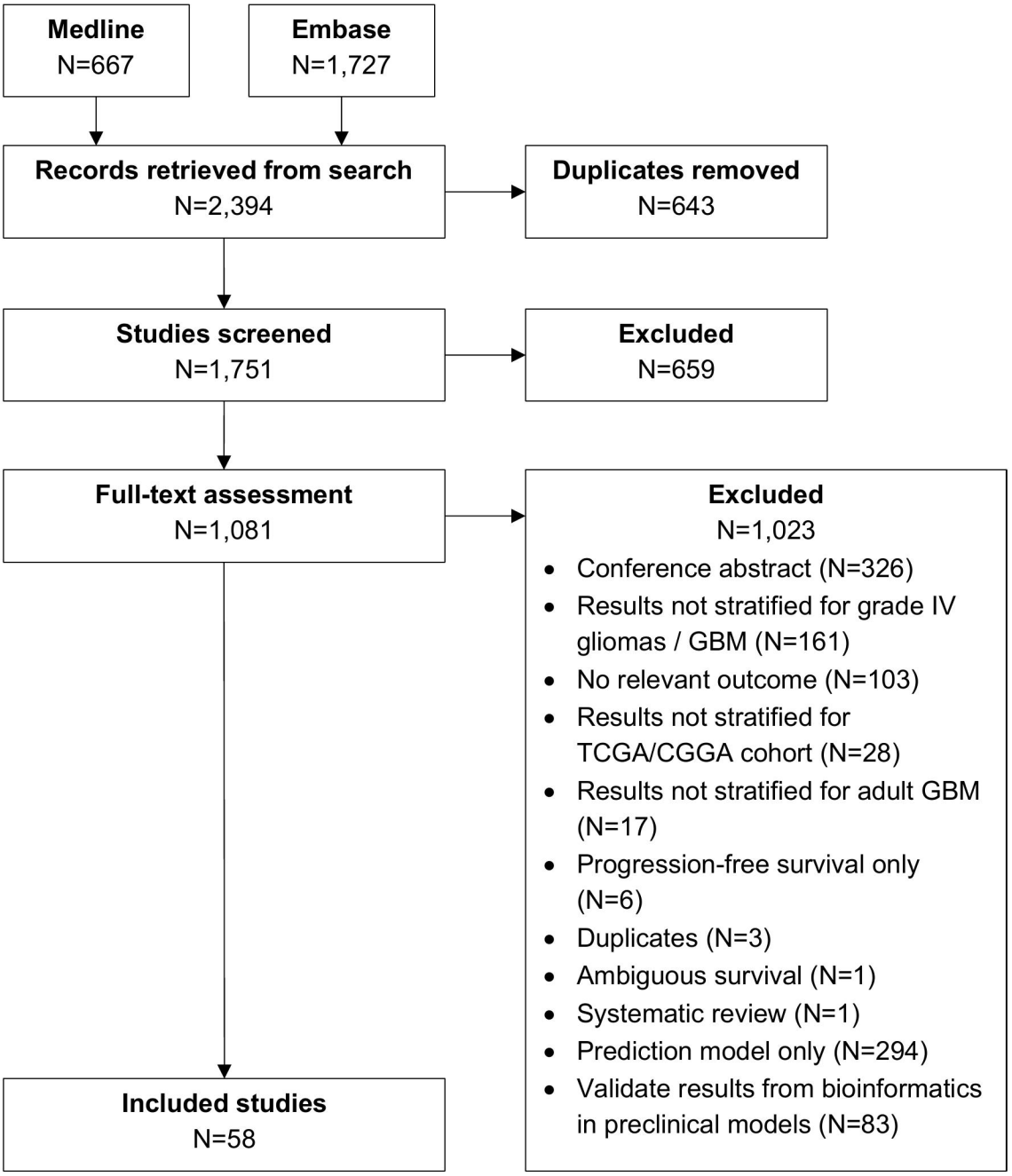
S2 Supporting information. Search strategy in Medline and Embase

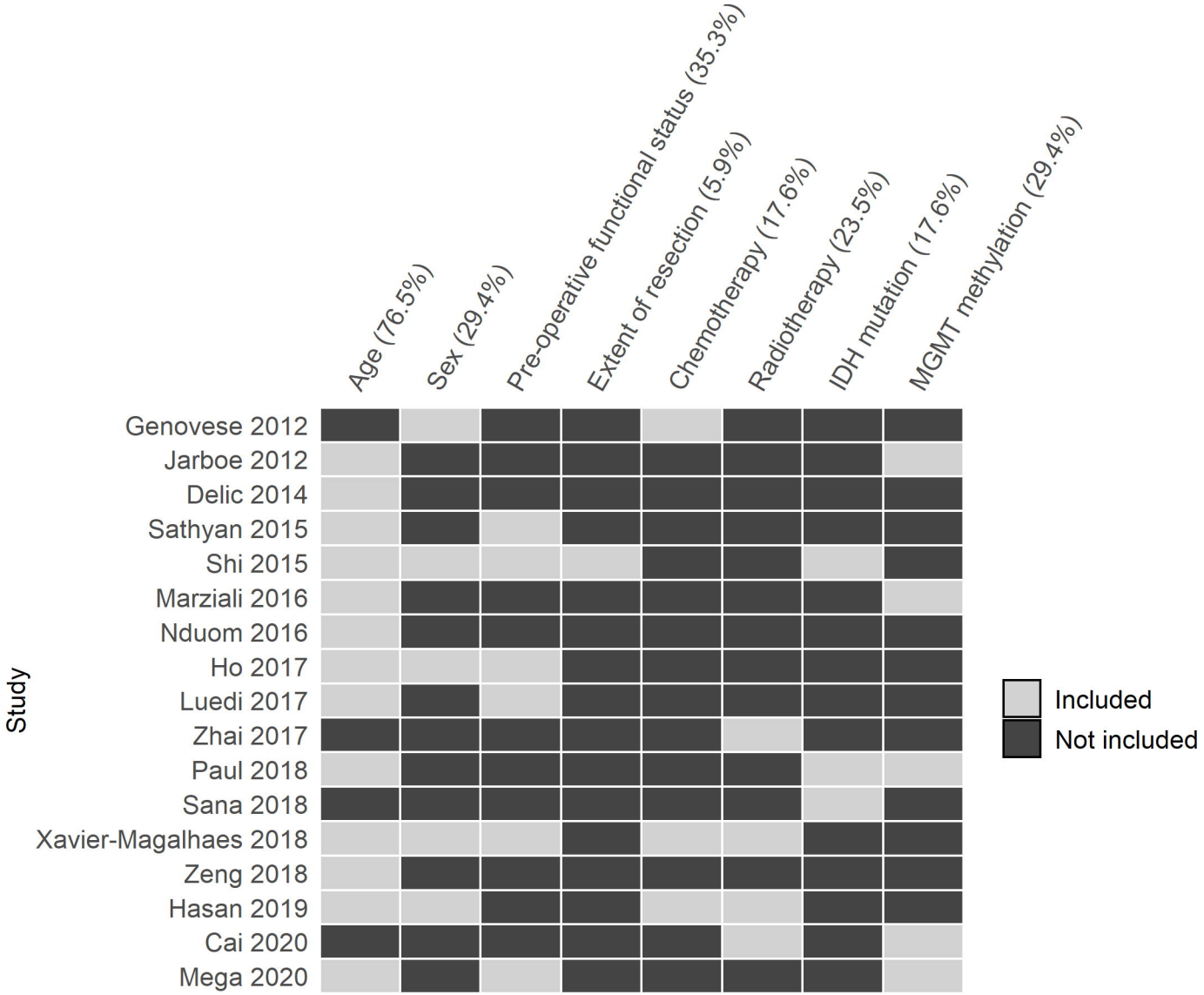
S1 Fig. Common analytic strategy used by included studies.

S1 Table. Characteristics of 58 included studies. Data type: A = RNA microarray only, B = RNA microarray and miRNA microarray, C = RNA sequencing only, D = RNA microarray and RNA sequencing, E = miRNA microarray only, F = RNA sequencing, RNA microarray and miRNA microarray, G = RNA sequencing and miRNA microarray, H = RNA microarray and DNA methylation, I = RNA sequencing, RNA microarray and DNA methylation, J = Unspecified. If a study used a data source but not specified the number of patients, the column for data source would be "Yes [NS]" indicating number of patients not specified

S2 Table. References to specific analyses extracted for comparison of results on molecular markers. Molecular markers ordered alphabetically accompanied with the location of analysis in the original manuscript. U = univariable survival analysis; M = multivariable survival analysis; ▲ = positive association i.e. higher levels of the molecular marker associated with better survival and $p<0.05$; ▼= negative association i.e. lower levels of molecular marker associated with worse survival and $p<0.05$; □ = statistical significance not demonstrated ($p\geq0.05$)

S1 References. Full references of included studies.


Medline
N=667
Embase
N=1,727
Records retrieved from search
N=2,394
Duplicates removed
N=643
Studies screened
N=1,751
Excluded
N=659
Full-text assessment
N=1,081
Excluded
N=1,023
Conference abstract (N=326)
Results not stratified for grade IV gliomas / GBM (N=161)
No relevant outcome (N=103)
Results not stratified for TCGA/CGGA cohort (N=28)
Results not stratified for adult GBM (N=17)
Progression-free survival only (N=6)
Duplicates (N=3)
Ambiguous survival (N=1)
Systematic review (N=1)
Prediction model only (N=294)
Validate results from bioinformatics in preclinical models (N=83)
Included studies
N=58


Age (76.5%)
Sex (29.4%)
Pre-operative functional status (35.3%)
Extent of resection (5.9%)
Chemotherapy (17.6%)
Radiotherapy (23.5%)
IDH mutation (17.6%)
MGMT methylation (29.4%)
Study
Genovese 2012
Jarboe 2012
Delic 2014
Sathyan 2015
Shi 2015
Marziali 2016
Nduom 2016
Ho 2017
Luedi 2017
Zhai 2017
Paul 2018
Sana 2018
Xavier-Magalhaes 2018
Zeng 2018
Hasan 2019
Cai 2020
Mega 2020
Included
Not included