Neuro-Oncology Advances

6(1), vdae193, 2024 | https://doi.org/10.1093/noajnl/vdae193 | Advance Access date 28 December 2024



# A machine learning–assisted systematic review of preclinical glioma modeling: Is practice changing with the times?

**Theodore C. Hirst, Emma Wilson, Declan Browne, and Emily S. Sena**

All author affiliations are listed at the end of the article

Corresponding Author: Theodore C. Hirst, MBChB, BMedSci (hons), Patrick G Johnson Centre for Cancer Research, Queens University Belfast, 97 Lisburn Road, Belfast BT9 7AE, UK (t.hirst@qub.ac.uk).

**Abstract**

**Background**. Despite improvements in our understanding of glioblastoma pathophysiology, there have been no major improvements in treatment in recent years. Animal models are a vital tool for investigating cancer biology and its treatment, but have known limitations. There have been advances in glioblastoma modeling techniques in this century although it is unclear to what extent they have been adopted.

**Methods**. We searched Pubmed and EMBASE using terms designed to identify all publications reporting an animal glioma experiment, using a machine learning algorithm to assist with screening. We reviewed the full text of a sample of 1000 articles and then used the findings to inform a screen of all included abstracts to appraise the modeling applications across the entire dataset.

**Results**. The search identified 26 201 publications of which 13 783 were included at screening. The automated screening had high sensitivity but limited specificity. We observed a dominance of traditional cell line paradigms and the emergence of advanced tumor model systems eclipsed by a large increase in the volume of cell line experiments. Few studies used more than 1 model in vivo and most publications did not verify critical genetic features.

**Conclusions**. Advanced models have clear advantages in terms of tumor and disease recapitulation and have largely not replaced traditional cell lines which have a number of critical deficiencies that limit their viability in modern animal research. The judicious use of advanced models or more relevant cell lines might improve the translational relevance of future animal glioblastoma experimentation.

## Key Points

- Animal glioma studies are used frequently, although traditional paradigms have limited construct validity.
- Better tumor models are available, but their uptake has been very limited.
- Improvements in modeling should improve the relevance of animal glioma research.

Glioblastoma is an aggressive primary brain tumor with a prognosis among the worst across all cancer types. Despite significant investment into the improvement of surgical and oncological therapies, and fundamental change in our understanding of glioma genetics and cell biology, there have been no major developments in treatment in the last 2 decades.[1–6]

Animal modeling of cancer has been a mainstay of drug development for half a century, and these studies often inform clinical trials. However, the translation of research from lab to clinic is known to be particularly inefficient, and the design and conduct of animal studies contribute. In cancer research, over 90% of drugs apparently effective in vivo are ineffective

## Importance of the Study

Our understanding of glioma biology has improved radically in the last 2 decades, although this has not led to a major improvement in outcomes. Animal research is a vital part of the bench-to-bedside process, and many candidate therapies are tested in animals before application to clinical trials. There are a range of tools available to model glioma in animals, although traditional models have significant limitations in disease recapitulation and reproducibility; newer models have clear advantages in terms of precise disease recapitulation in the molecular age. However, we show that despite their appearance in recent publications, there has been a corresponding marked increase in the use of traditional models, meaning that the vast majority of studies continue to model glioma poorly. We offer some practical suggestions that should improve the validity and reproducibility of future animal glioma research.

in human phases, and there is compelling evidence suggesting that the construct validity, design, and conduct of the preclinical study have a contributory role.[7–11]

The most commonly published glioblastoma paradigms use immortalized cell lines derived from animal or human tumors induced in the 1960s and 1970s.[12–16] Concerns have been raised as to these models' validity, given a lack of precise disease recapitulation[17] and a poor correlation between xenograft cell line use in vivo and clinical outcomes.[11] In previous preclinical glioma meta-analyses, we have observed that tumor model selection often accounts for substantial between-study heterogeneity.[14,15,18,19] If data obtained using these paradigms are neither valid nor reproducible, serious questions arise as to whether they are more harmful than informative overall.

Other model types have been described with advantages: with an ability to recapitulate key molecular, genetic, and microenvironment features and address the common issue of immunocompromised hosts.[20–22] Similarly, newer cell line models have been reported that more closely resemble human disease and have a clearer delineation of key clinicopathological characteristics.[23] However, it is unclear to what degree these models have been adopted.

Systematic review is a tool to conduct a thorough and unbiased appraisal of all evidence pertaining to a particular research question. The recent addition of machine learning tools in systematic review procedures has allowed the inclusion of larger volumes of studies, previously not possible with manual techniques.[24,25] We, therefore, aimed to use this technique to summarize preclinical glioma modeling practices without restrictions or bias, with a particular focus on whether a change has occurred during the last 20 years.

## Materials and Methods

The protocol for this study was published on March 8, 2022.[26]

### Search Strategy

We searched Pubmed and Embase on February 28, 2022, and updated on November 18, 2023, using the following terms:

(Glioblastoma OR glioma OR astrocytoma OR gliosarcoma OR brain tumor)
AND
(In vivo OR orthotopic OR heterotopic)

Searches were limited to animals.

### Study Selection

We aimed to include all studies that had tested any glioma model in an animal, regardless of the focus or outcomes tested. As such, our inclusion criteria were broad:

1) Original research article—all publication types considered
2) Any animal model
3) Any glial tumor: glioblastoma, glioma, astrocytoma, oligodendroglioma, low-grade glioma, diffuse midline glioma, and other pediatric glial tumors
4) Any study focus, treatment, or outcome
5) No comparator or control required

We excluded non-glial tumors (eg, medulloblastoma, pituitary adenoma, metastases, ependymoma, etc.), or tumors due to a germline mutation (eg, neurofibromatosis). There were no limitations on the date of publication or language. The vast majority (>99%) were published in English, with the remainder in Chinese, generally with abstracts in English. Where required to aid screening decisions, key information was translated using Google Translate.

The primary search returned 23 959 unique publications, of which 1000 were selected at random in Excel for manual screening. Two investigators (T.H./D.B.) independently screened the titles, abstracts, and full texts where necessary. Disagreements were resolved by direct discussion. Binary classification (include or exclude) screening decisions were then used alongside the title and abstract text to train a machine learning algorithm based at the EPPI center at University College London for automated screening. The algorithm uses a tri-gram "bag-of-words" model for feature selection and implements a linear support vector machine with stochastic gradient descent.[27] This algorithm was used only to aid study screening and not with data extraction later in the study.

Of the 1000 publications, 463 were labeled as included and 537 were labeled as excluded through screening. Eight hundred (80%) labeled publications were used to train the algorithm and 200 (20%) were retained for performance

validation. We applied a threshold score of 0.56, which indicated the automated screening had a 96.88% sensitivity (recall), 76.92% specificity, 79.49% positive predictive value (precision), and 87.33 F1 score (see Supplementary Materail 1).

After training, we submitted the initial search of 23 959 publications for the automated screening process, with 12 850 publications subsequently labeled as included.

### Publication Sample Full-Text Review

Next, we divided publications into 5 groups, according to date of publication (pre-2001, 2001–2005, 2006–2010, 2011–2015, 2016-present), and randomly selected 200 from each group to give a working cohort of 1000 articles for manual full-text review.

For each article, we extracted bibliographic data (year, authors, journal); data pertaining to the host animal (species and immune status); the tumor model (site of tumor, tumor model category, cell type, number of models used per article); whether there was genetic confirmation of cell type or key molecular features relevant to clinical practice (eg, IDH, MGMT, EGFR, TERT, PTEN, p53, ATRX, 1p/19q, H3.3 histone mutations, CDKN2A/B); and the overall focus of the study.

We categorized tumors into "cell line," "PDX," "GEMM/oncomice," "induced tumors," and "other." Cell lines were defined as any transplantable cells that have been maintained in culture or neurospheres in vitro; "PDX," patient-derived tumors that were processed and implanted in vivo directly, or maintained in vivo until their use. We encountered instances where tumors were described as "patient-derived" but had been maintained in in vitro culture—these were assigned to the "cell line" category. Cells described as U373 were included in the U251 group because of a well-known contamination issue.[28,29] GEMM/oncomice were any animals with a genetic construct that leads individuals to develop tumors spontaneously or after induction of a non-carcinogenic agent. Induced tumors were defined as those whose development was caused by the delivery of a carcinogen or other agent in otherwise healthy animals.

The review was conducted in accordance with PRISMA guidlines (Supplemetary Materail 4). Data were extracted using the SyRF online platform, available at https://syrf.org.uk.[30]

### Full Dataset Abstract Review

In November 2023, we updated the search and, after submitting the additional 1208 studies, automatic screening decisions identified a further 921 studies for inclusion. Finally, during the publication review process, a further 12 relevant GEMM studies were highlighted to us that had not been included in the search, which we included at this stage. As such, the total dataset included 13 783 publications.

We categorized each abstract if it included one of the critical terms pertaining to each of the study design features identified during the first stage, as described in Supplementary Material 3. This process was completed using Microsoft Excel.

Systematic searches were saved via the export functions in each search website and imported into Endnote X9, for further management. Abstract search functions were performed in Microsoft Excel, and Statistical Analyses were performed using PASW Statistics 18. For the first dataset, the cutoff *P*-value was Bonferroni adjusted to $P = .004$ (11 comparisons), and for the second, $P = .025$ (2 comparisons).

## Results

In total, our searches identified 25 168 unique publications (23 959 at primary search and 1209 further at update). During the manual screening process, we included 463 and excluded 537, and the machine learning algorithm identified a further 12 387, leaving 12 850 included after the first search and 933 more after the update, leading to a total of 13 783 included studies (Figure 1). Of these, 150 were conference abstracts, and the remainder were journal articles.

The yearly volume of studies included increased substantially over time, with an average of 5.4 publications/year between 1975 and 1980 rising to a peak of 1143/year in 2020–2021. In 2022 and 2023, there were 1051 included articles, equivalent to 560/year (Figure 2).

### Full-Text Review

We selected 1000 studies to represent the full range of dates over which preclinical glioma research has been published. We excluded 285 articles, leaving 715 for a full review. Of those excluded, 91 did not use a glioma model, 90 reported glioma in vitro only, the model was unclear in 40, 30 were reviews, 24 clinical studies, 4 retracted, 2 technical notes, 2 neurofibromatosis studies, and we were unable to retrieve 2 manuscripts. A greater proportion of older studies were excluded (39.7% published before 2001 compared with 24.4% 2016–2022). To ensure a more even distribution across groups, and to allow a better representation of current practices in the most recent group, we recategorized studies into those up to and including 1998, 1999–2002, 2003–2006, 2007–2010, 2011–2014, 2015–2018, and 2019–2022. Raw data are available in Supplementary Material 2.

### Animal Hosts

Across the 715 publications, we identified 983 models, a mean of 1.37 per study (range 1–25). There was a trend for the number of models to increase with time, with the earliest articles reporting a mean of 1.19 publications, rising to a peak mean of 1.50 in 2015–2018 (Figure 3A). The majority (682/983, 69.4%) used mice, with 296 (30.1%) using rats, 3 dogs and 1 each of hamsters and zebrafish. Most animals had immune dysfunction, with almost half of all models using athymic animals (486/983, 49.4%). We also observed the use of animals severe combined immune deficient (SCID) animals (78, 7.9%), BDIX mice (11, 1.1%), and NOD-SCID gamma (7, 0.71%). Only 1 model reported the use of animals with a humanized immune system. Immunocompetent animals were used in 367 (37.3%) models. Rats were favored in earlier studies although there

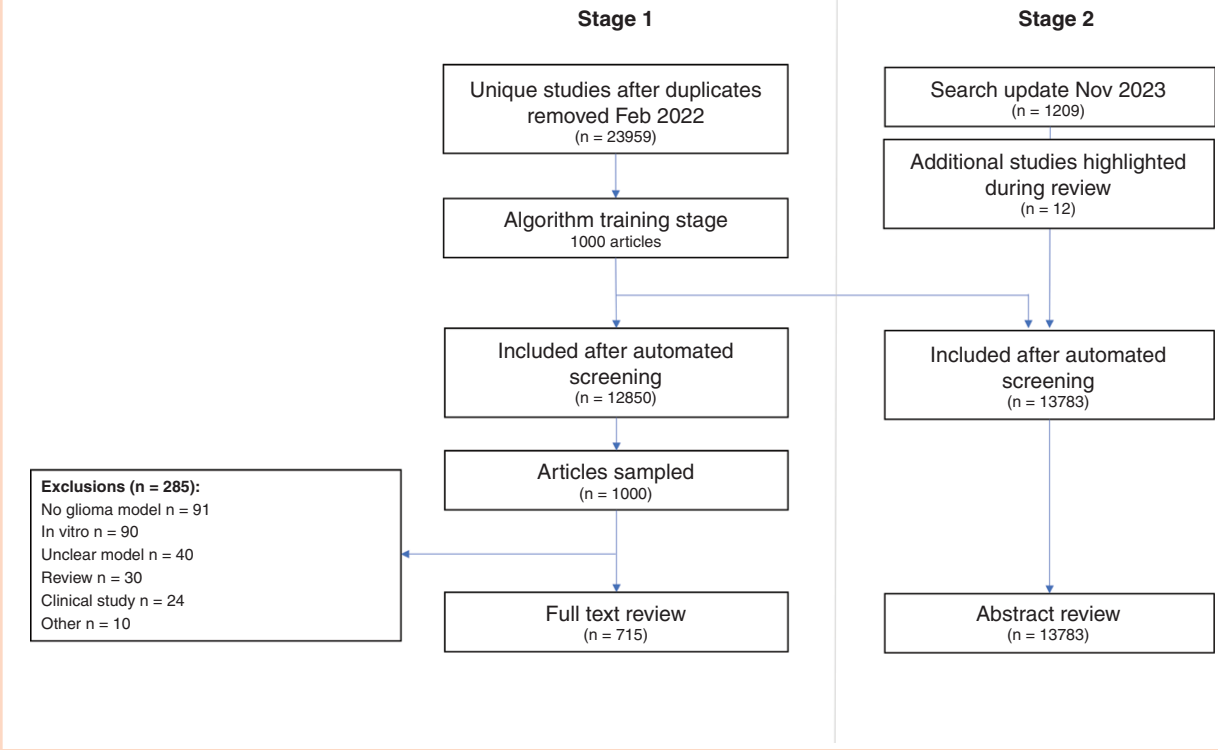


**Figure 1.** PRISMA chart for study inclusion.

was a clear move toward mice (Figure 3B; $\chi^2 = 188$, df = 12, $P < .001$) and athymic and SCID (Figure 3C; $\chi^2 = 138$, df = 42, $P < .001$) animals with time.

Overall, the majority of models used orthotopic tumor implantation (661/983, 67.2%). All but one of the remainder used subcutaneous tumors. In the earliest group, subcutaneous tumors were used in 42.6% of models, and there was a trend for orthotopic models to be favored over time, described in 73.9% of those published in 2019–2022 (Figure 3D; $\chi^2 = 18.3$, df = 6, $P = .006$).

## Glioma Paradigms

The vast majority of publications (641/715, 89.7%; 799/983 models, 81.3%) used xenograft or syngenic cell lines. PDXs were the main model reported in 46 publications (6.4%) and used in 137 models (13.9%). GEMMs were described in 14 (2.0%) and induced tumors 13 (1.9%) of publications. Almost all (699/715, 97.8%) used 1 model type, with 15 describing the use of both cell lines and PDXs, and 1 using cell lines and a GEMM. Cell lines were the prominent paradigm throughout, although PDX and GEMM models were used in a minority of recent publications (Figure 4A; $\chi^2 = 51.6$, df = 24, $P = .001$).

We observed the inoculation of 110 different cell lines in vivo, although 1 of 10 lines was used in 631/799 models (79.0%). The most frequently described was U87 (210/799 cell line models, 26.3%), then C6 (149/799, 18.6%), 9L (80/799, 10.0%), U251 (70/799, 8.76%), and GL261 (43/799, 5.38%). There was a clear trend for syngenic rat models C6 and 9L to be favored in earlier studies, replaced by U87 and GL261 more recently. U87 was used in 47.6% of all models during the 2015–2018 period and is by far the most frequently used model currently (Figure 4B). The commonly described cell lines were generally acquired from commercial sources and were developed between 1968 (U87, C6) and 1994 (CNS-1).

Studies that primarily used cell lines used a single cell line in vivo in 80.7% of cases (517/641 publications): 98 used 2 cell lines (15.28%) and 26 (4.06%) used 3 or more. One publication used 6 lines. The was no change in the number of tumor models used over time, with the majority (73.5%) of publications from 2019 to 2022 only using 1 line in vivo (Figure 4C, common cell lines summarized in Table 1; $\chi^2 = 42.7$, df = 48, $P = .691$).

Of the cell line paradigms, the host animal features understandably differed according to whether xenograft or syngenic model design. Selecting studies reporting the use of 1 of the 10 most frequently used cell lines (see Table 1), xenograft models were transplanted into mice in 291/304 (95.7%) of cases, with alterations of immune function in 285/204 (93.8%), mostly athymic (78.6%). As expected, the majority (237/327, 72.5%) of syngenic models used animals with an intact immune system, although a prominent minority (74/327, 22.9%) used athymic animals (Figure 4D).

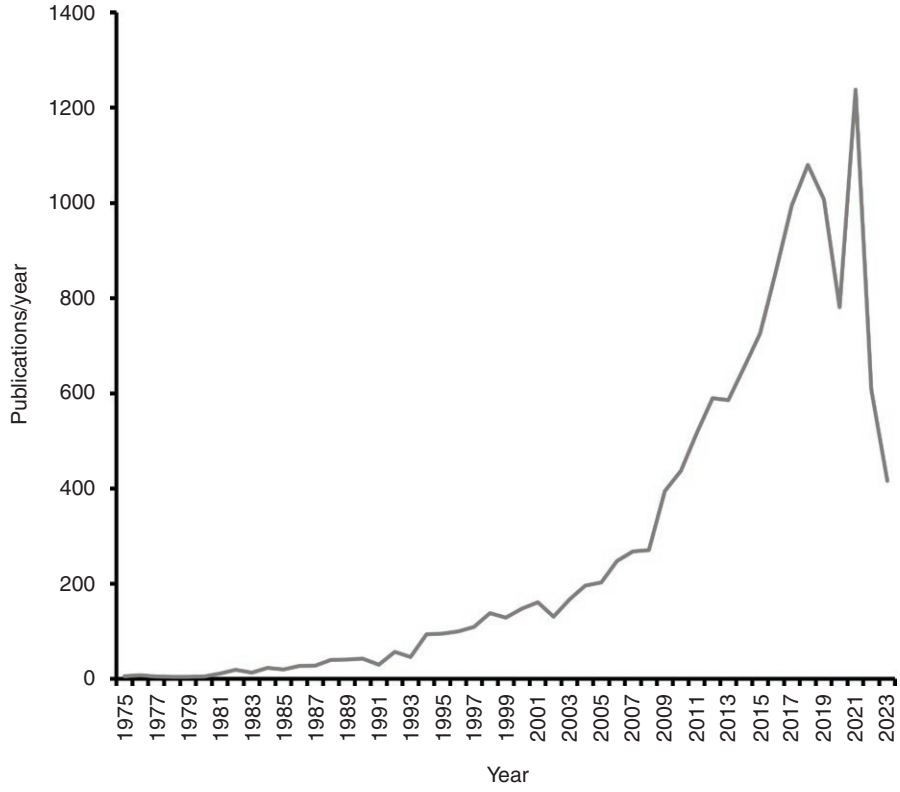


**Figure 2.** Frequency of publications included after automated screening stage.

One publication (1.18%) reported the use of PDXs in the pre-1998 group, and their use was infrequent until around 2011. Since then, they have been consistently used in around 10% of publications. The median number of PDX models per study was 1 (interquartile range 1–3), and 55.4% of PDX studies used a single model, although some publications used many—25 tumors in 1 and between 5 and 7 in another 5. There was no trend for increasing the use of multiple PDX models with time; rather, studies using PDXs before 2007 always used more than 1 model, whereas 50% (4/8) of 2019–2022 publications using PDXs only used a single model. The majority of PDX models used mice (124/137, 90.5%), and most of those using rats were reported in 1999–2002. Athymic animals were used most frequently (77/137, 56.2%), followed by SCID (26/137, 19.0%). A minority reported the use of immunocompetent animals (28/137, 20.4%).

GEMMs were reported in 38/983 models (3.87%), first described in a publication from 1998. although mostly represented between 2010 and 2022. They were all in mice and immunocompetent in 37/38 (97.4%). Genetic engineering targeted 1 or multiple genes, including Ink4a/ARF, p53, CDKN2A, PTEN, ATRX, and EGFR. The animals included in experiments all developed intracranial tumors. The preferred targets did not obviously change over time, with Ink4a/ARF reported most frequently in both 2003–2006 and 2019–2022.

Induced tumors were used in 8 mostly earlier models, with the agent most commonly used being ethylnitrosourea. Two models used viruses (eg, avian sarcoma virus) infusion to induce tumors. The models using ethylnitrosourea were all historical, 1986–1997, and tumors were induced in both rats and mice, athymic or without comorbidity.

## Additional In Vitro Modeling

During the data extraction phase, we noticed that many publications tested their hypothesis on tumor tissues in vitro, in addition to those tested in vivo. As such, we recorded if additional tumor models had been tested in vitro as we felt this was relevant, particularly when a publication had only used a single in vivo model. We observed this in 144/715 (20.1%) of publications, and this was described more frequently with time (34.6% in 2019–2022, 3.5% in pre-1998; $\chi^2 = 40.0$, df = 12, $P < .001$, see Figure 4C). Multiple models were tested in vitro in the majority of instances and there was a range between commercial cell lines and primary patient-derived tumor samples.

## Study Focus

There was unsurprisingly a wide range of investigations and interventions described in the 715 included publications. We categorized each by its main focus, although many publications addressed more than 1.

Treatments were tested in 444/715 publications (61.2%), most frequently gene therapies (135/715, 18.9%),

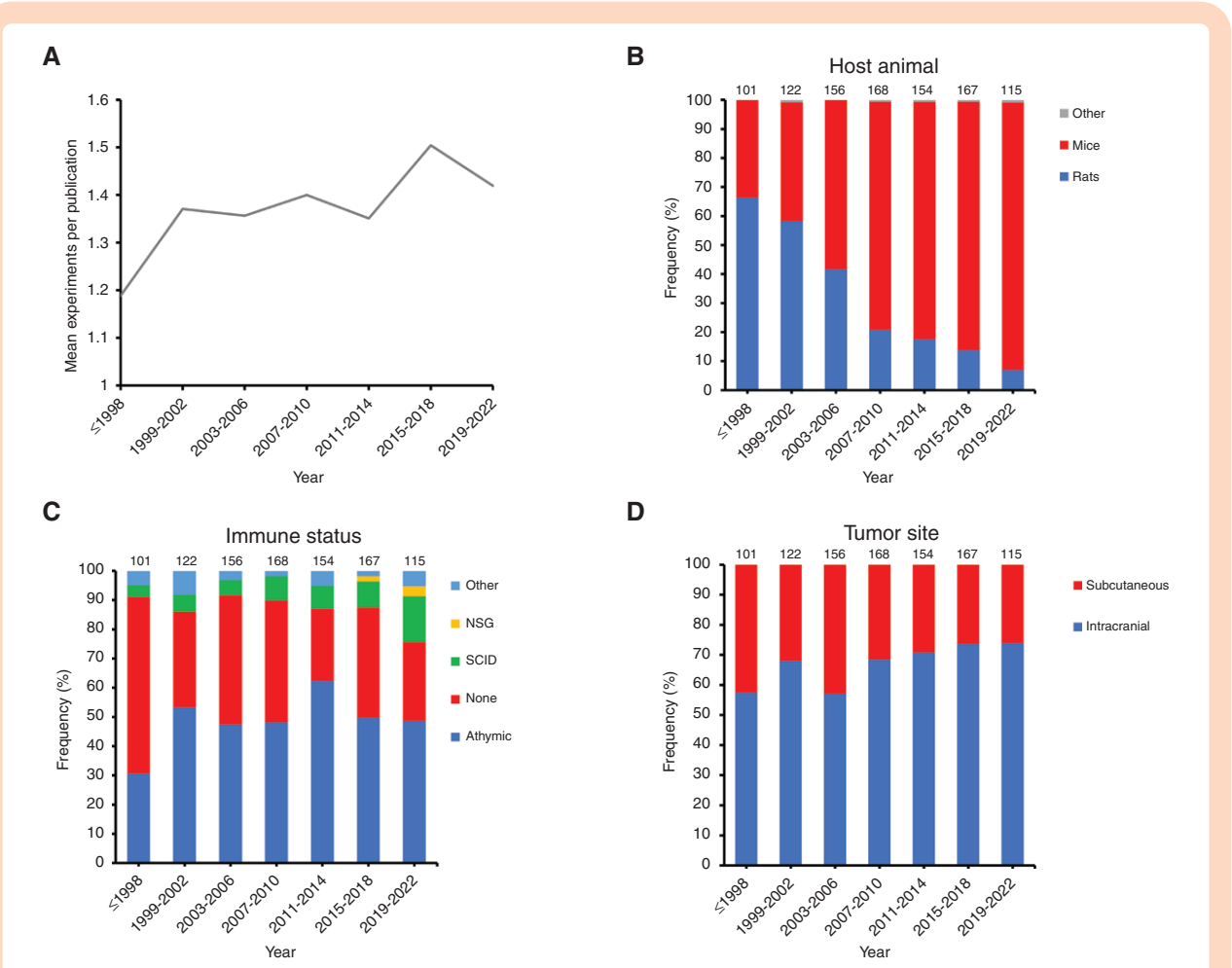


**Figure 3.** Summary of host animals. (A) Mean experiment paradigms per article versus year of publication. (B) Stacked bar chart of host animal grouped by year of publication, showing a clear trend for mice to become favored over time; $\chi^2 = 188$, df = 12, $P < .001$. (C) Stacked bar chart of comorbidity grouped by year of publication, showing an increase in the use of athymic and severe combined immune deficient (SCID) mice; $\chi^2 = 138$, df = 42, $P < .001$. (D) Stacked bar chart of tumor location grouped by year of publication, showing a slight reduction in subcutaneous tumor use with time; $\chi^2 = 18.3$, df = 6, $P = .006$. Data from publications ($n = 715$) are summarized in A, and experiments ($n = 983$) in A–C. The numbers at the top of each stacked bar represent the total number in each group.

chemotherapy/small-molecule therapy (120/715, 16.8%), immunotherapies (59/715, 8.3%), and radiation-based treatments (53/715, 7.4%). Tumor biology was the main focus in 20.4% of publications (148/715), where studies focused on gliomagenesis or glioma cell biology (77/715, 10.5%), tumor–brain interactions (26/715, 4.6%), or tumor genetics (25/715, 3.5%). Imaging (mostly advanced MRI and PET sequences) was used in 86 publications (12.0%) and new models were the main focus of 19 publications (2.7%, see Figure 5A).

There were trends for gene therapies to be described less frequently with time, with a peak of 31.3% of articles published between 2003 and 2006 reducing to 4.9% in 2019–2022; and small-molecule therapies were represented with increasing frequency over time, consistently around 20% since 2011. Similarly, tumorigenesis/glioma cell biology and glioma genetics were investigated with increasing frequency, representing 13.5% and 9.9% of 2019–2022 studies, respectively ($\chi^2 = 165$, df = 72, $P < .001$; Figure 5A).

Unsurprisingly, some tumor model types were favored for different study types ($\chi^2 = 91.8$, df = 48, $P < .001$; Figure 5B). That is, 52.6% of experiments using GEMMs investigated tumorigenesis and glioma cell biology, particularly in a tumor's early stages. We did not observe any GEMM studies focusing on tumor genetics—appreciable given the predetermined genetic status quo of the tumors—and only 5/19 publications (26.3%) tested a treatment. PDX models were most frequently deployed to investigate glioma cell biology, in 10/46 (21.7%), and a greater proportion of PDX models were used to investigate tumor–brain interactions than any other tumor model types (10.9%)—again unsurprising given accurate recapitulation of the brain–tumor microenvironment is one of the main advantages of PDX. Gene therapies, imaging, antiangiogenic treatments, advanced delivery techniques, and radiation were most frequently investigated using cell line models.

Overall, when study focuses were dichotomized into "treatment" and "other," we observed that a greater

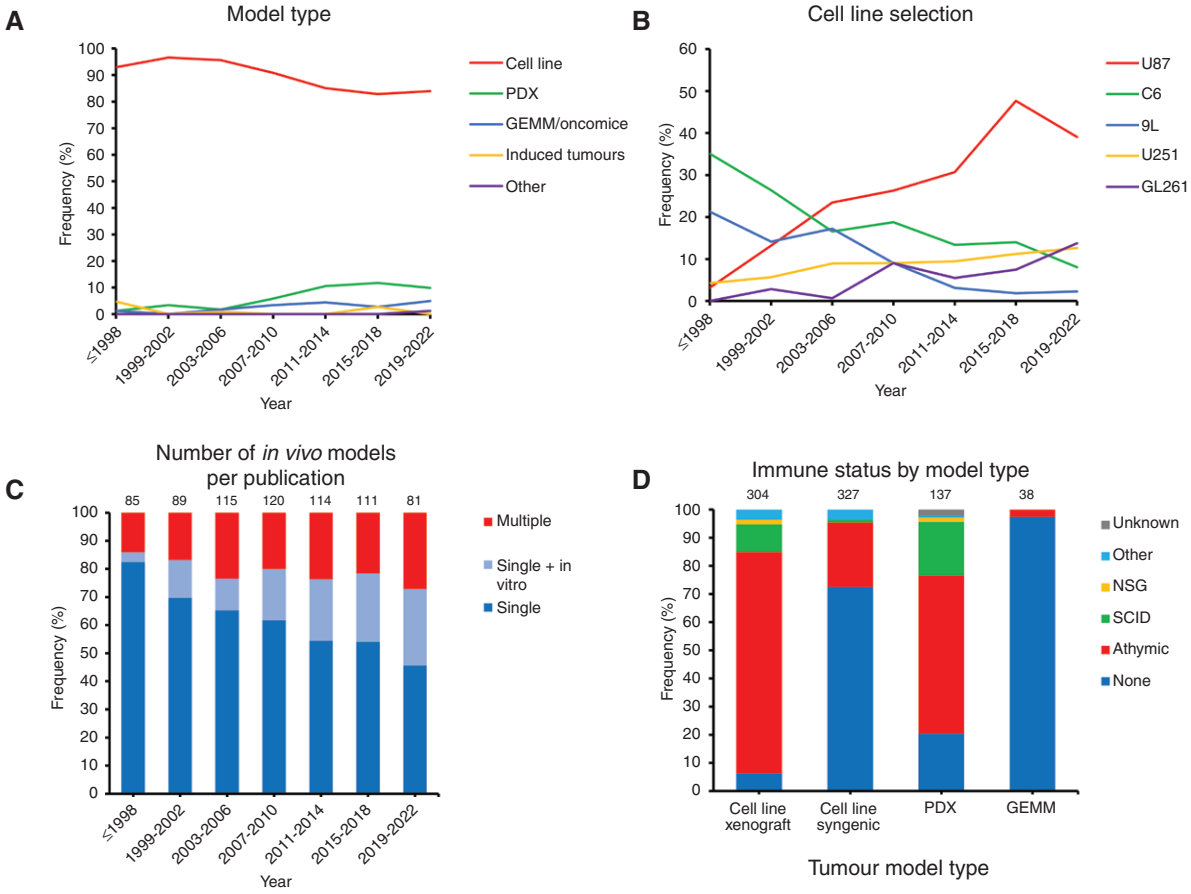


**Figure 4.** Summary of tumor paradigms. (A) Frequency of tumor model type over time, showing emergence of PDX model since 2011 but ongoing predominance of cell line models; $\chi^2 = 51.6$, df = 24, $P = .001$. (B) Frequency of cell line selection over time, including the 5 most commonly reported lines as a proportion of all cell lines. There was an increasing prevalence of U87 use and decreasing use of C6 and 9L models. (C) Stacked bar chart of the number of models used grouped by year of publication, showing no change in the proportion of publications using more than 1 in vivo tumor model; $\chi^2 = 42.7$, df = 48, $P = .691$. However, when articles using a single in vivo model were divided according to whether they used further models in vitro, there was an increasing frequency of the representation of multiple models with time; $\chi^2 = 40.0$, df = 12, $P < .001$. (D) Stacked bar chart of immune status grouped by tumor type, showing predominance for use of athymic and severe combined immune deficient (SCID) animals for xenograft models, and animals with intact immune systems for most but not all syngenic and GEMM models. Data are expressed as a percentage of publications ($n = 715$) in A and C; as a percentage of all cell line experiments ($n = 799$) in B; and all experiments ($n = 983$) in D. The numbers at the top of each stacked bar represent the total number in each group.

**Table 1.** Summary of Common Cell Lines

| Model | Origin | Description | Count in Sample | Count in Abstract Review |
|---|---|---|---|---|
| U87 | Human | 1968 | 207 | 2082 |
| C6 | Rat | 1968 | 146 | 1303 |
| 9L | Rat | 1971 | 79 | 450 |
| U251 | Human | 1973 | 69 | 850 |
| GL261 | Mouse | 1970 | 43 | 571 |
| F98 | Rat | 1980 | 30 | 179 |
| T98 | Human | 1979 | 13 | 193 |
| RG2 | Rat | 1980 | 12 | 97 |
| BT4C | Rat | 1974 | 11 | 52 |
| CNS-1 | Rat | 1994 | 8 | 15 |

proportion of publications using cell lines tested a treatment (64%) when compared with PDX (47.8%) or GEMM (26.3%). Finally, we considered whether investigators might take further steps to introduce tumor heterogeneity in studies that test a treatment. Overall, the number of articles using multiple models was less in this group and there was no change in this difference with time ($\chi^2 = 2.15$, df = 1, $P = .142$; Figure 5C).

## Model Verification and Confirmation of Key Genetic/Molecular Features

Of the 715 publications, 71 (9.9%) described confirmation of relevant genetic or molecular tumor features analogous to those of established value in clinical practice. Of these, the vast majority focused on a single gene that was of relevance to the study (48/71, 67.6%, mostly EGFR). A further

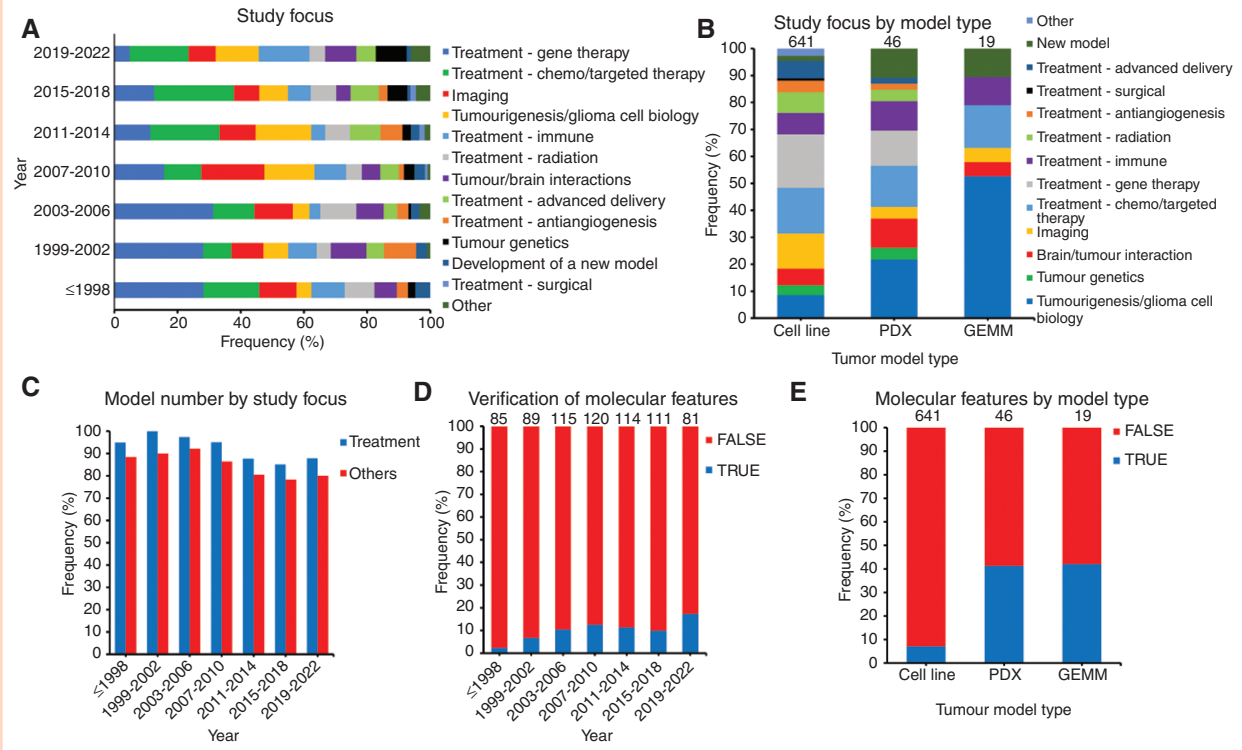


**Figure 5.** Summary of publication focus and model selection. (A) Frequency of article focus over time, showing predominance of gene therapies moving toward targeted therapies and tumorigenesis; $\chi^2 = 165$, df = 72, $P < .001$. (B) Stacked bar chart of article focus grouped by tumor type, showing predominance for use of GEMMs for tumor biology experimentation, and a more even distribution of article focuses for studies using cell lines and PDXs; $\chi^2 = 91.8$, df = 48, $P < .001$. (C) Clustered bar chart showing the proportion of publications using a single tumor model, when categorized according to whether they assessed a treatment or had another focus. Those testing treatments were more likely to use a single model, and was no change in this difference over time; $\chi^2 = 2.15$, df = 1, $P = .142$. (D) Stacked bar chart showing the proportion of publications describing key genetic features or verifying tumor identity grouped by year of publication. There was a nonsignificant trend for increasing frequency although this was only reported in 17.3% of recent publications; $\chi^2 = 12.2$, df = 6, $P = .058$. (E) Stacked bar chart showing the proportion of publications describing key genetic features or verifying tumor identity grouped by main model type. Publications using mainly PDX and GEMM models validated key molecular features more frequently than those using cell lines; $\chi^2 = 77.1$, df = 4, $P < .001$. Data in A–E are expressed as a percentage of publications ($n = 715$). The numbers at the top of each stacked bar represent the total number in each group.

10 (14.1%) described 2 genes only. The status of IDH and MGMT, vitally informative of clinical glioma management, were confirmed in only 6 and 5 publications, respectively. There were only 3 publications that verified the identity of the cell lines using short tandem repeat analysis. There was a nonsignificant trend for genetic features to be described more frequently over time, although in 2019–2022, this was only 17.3% of publications ($\chi^2 = 12.2$, df = 6, $P = .058$; Figure 5D). Similarly, description of key molecular features was more frequently reported for PDX models (41.3%, 19/46) and GEMM (42.1%, 8/19) than for cell lines (7.17%, 46/641; $\chi^2 = 77.1$, df = 4, $P < .001$; Figure 5E).

## Abstract Screen

To provide an estimate of the representation of model types reported in all identified literature, we searched the abstracts of each of the 13 783 included studies for terms pertaining to the different tumor model types, as highlighted in the methods. Overall, we found that at least one of our terms used was included in 8189/13 783 (59.4%) of abstracts, compared with 519/715 (72.6%) of the publications subjected to full-text review. Of the 519, the model indicated was the same as that identified on full-text data extraction in 494/519 (95.1%) of instances. While inappropriate to infer the absolute number of publications using each paradigm (because of the limited specificity of the automated screening and abstract term searches), we felt it justifiable to use this approach to analyze the relative representation of each model type, as the limitations above appeared to apply to each subset equally.

Of the 8189 abstracts, cell lines were by far the most commonly reported model type, with these terms included in 7497 (91.5%), compared with 494/519 (95.1%) of the sample. PDX terms were included in 637/8189 (7.7%) and 18/519 (3.5%); GEMMs in 519/8189 (6.3%) and 37/519 (7.1%); and induced tumors 162/8189 (1.9%) and 20/519 (3.9%) of the whole dataset and sample, respectively. Using this technique, the observed distributions of the different model types changed over time in a manner comparable to that observed in the primary dataset ($\chi^2 = 304$, df = 18, $P < .001$; Figure 6A and C).

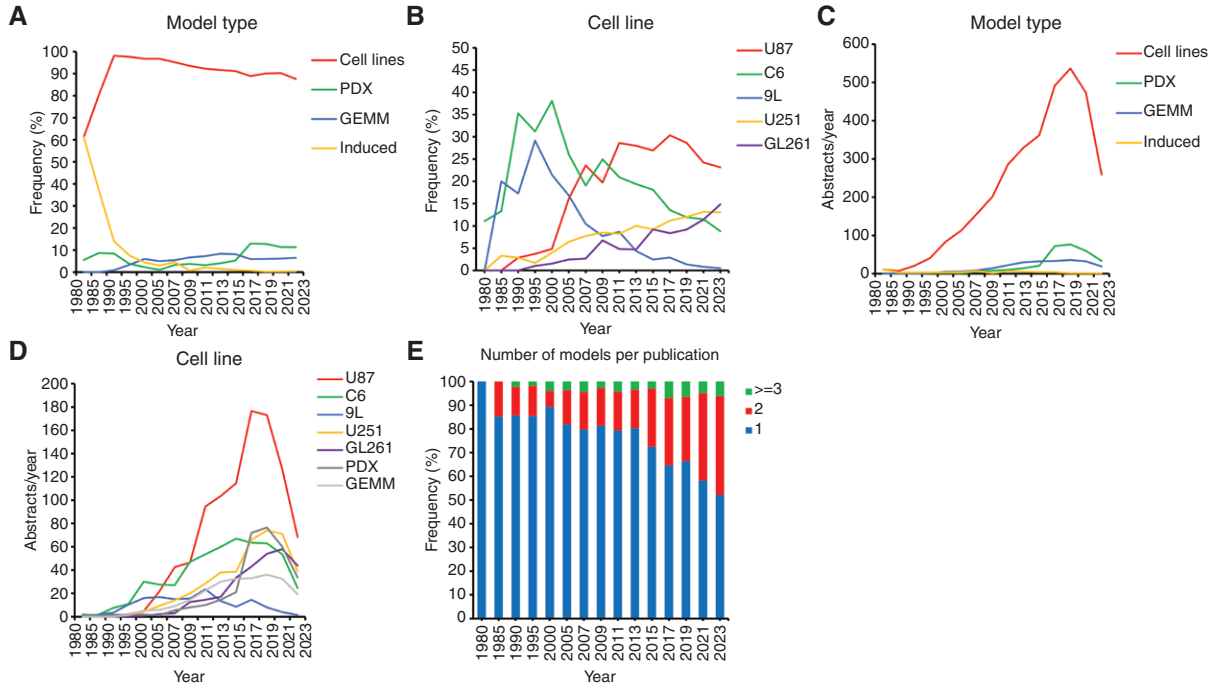


**Figure 6.** Summary of tumor models identified in abstract screen. (A) Frequency of model type over time for publications with a model identifiable from abstract text ($n = 8189$). Induced tumors were common in early studies, with appearance of PDX and GEMM recently, but cell lines predominate throughout; $\chi^2 = 304$, df = 18, $P < .001$. (B) Frequency of common cell lines over time, expressed as a percentage of those with an identifiable cell line ($n = 7497$), showing preference for C6 and 9L in early studies, replaced by U87 since around 2010. (C) Volume of abstracts with identifiable model category versus time: appearance of PDX and GEMM model has occurred alongside a large increase in the volume of cell line publications. (D) Volume of individual cell line models compared with PDX and GEMM over time. The volume of U87 literature has increased in volume to a much greater extent than PDX and GEMM models. (E) Stacked bar chart of the number of models identified in abstract, grouped by year of publication, showing an increase in those describing more than one tumor model; $\chi^2 = 310$, df = 6, $P < .001$, but it is unclear whether all are used in vivo.

We observed that, of the cell lines, U87 was mentioned in 2082 abstracts and C6 1303. U87 remains the predominant model in modern literature, included in 27.8% (1396/5009) of abstracts published in the last 10 years, compared with 11.1% (556) PDX and 6.5% (324) GEMM (see Figure 6B and D). Of the abstracts naming cell lines, 71.5% (3654/5105) described 1 of the 20 most commonly used lines, 23.6% (1205) 2 lines, and 4.8% (246) 3 or more. Multiple cell lines were mentioned with increasing frequency over time, rising from 11.9% (46/386) in pre-1998 to 51.7% (156/302) in 2022–2023 ($\chi^2 = 310$, df = 6, $P < 0.001$; Figure 6E), although we were unable to differentiate between their use in vivo and in vitro.

# Discussion

## Principal Findings

In this review, we have summarized the practices of animal glioblastoma modeling since their first application in the 1960s and described how glioma has been modeled in animals since. Newer cell lines, PDX and GEMM models, offer clear recapitulation advantages, but their adoption has been somewhat muted as flawed cell lines continue to dominate the preclinical glioma literature. While traditional cell lines clearly have advantages—easy to acquire and generate tumors, with established experience of their biology and behavior, and available in large numbers—we express concerns about the way in which they are being applied and the relevance of data obtained from testing them. There remains a proclivity to use a single cell line in vivo, usually without any confirmation of tumor identity or relevant genetic features, and those cell lines that are favored have major limitations which might not exist for newer alternatives.

## Validity of Methodological Approach

We adopted a simple and broad search strategy and inclusion/exclusion criteria to maximize the sensitivity of the search We used an automated screening process, which allowed us to generate a dataset that includes almost all studies reporting the use of an animal glioma model, rather than having to narrow the search question or use samples of a search because of limitations of human resource. While the specificity of the process was limited, at 72%, we have confidence that almost all relevant studies are captured.

While we are confident in the validity of our systematic review, we acknowledge the possibility of omission of

publications not including our search terms in their title or abstract. During the publication review process, a number of GEMM studies were highlighted to us that were not included in our search—this is likely a result of a search strategy that we adopted leading to the underrepresentation of this type of experiment. This, of course, reflects a weakness in this review, but we believe that the overall validity has been maintained, as firstly, this experiment type forms a small subset of the dataset (in our study and other published reviews), and secondly, we included these publications in the secondary abstract analyses. We did not change the literature search or main analysis as these had already been prespecified in our protocol.

We chose a sample of 1000 studies, anticipating that around a quarter of them would be excluded. While the volume of publications has increased dramatically over time, we chose to include publications distributed evenly over time. As such, the older literature is relatively overrepresented—the main focus of this review was to appraise how glioma modeling has changed over time. To address the underrepresentation of modern data, we developed a technique to identify and categorize studies by the screening of abstract text in Excel, validating this against the 1000-study sample. While sensitivity was limited, as many publications do not include modeling specifics in the abstract text, the screen resulted in the correct categorization of abstracts in 95% of instances—as such allowing for the reliable description of trends in practices within the entire dataset, including those published as recently as November 2023.

We did not extract data from all excluded publications as this was not feasible manually. Future developments in language processing algorithms should facilitate more complete data extraction from large datasets,[24] although we have not yet validated this approach.

During full-text review, we extracted a limited number of features, in particular with respect to study design, aims, and focus. We designed the study in this way as questions pertaining to specific treatments were beyond the scope of this review. That being said, practices of animal modeling may differ according to the specific focus—it is known that some interventions require certain experimental conditions, which may not have been appreciated in this analysis.[17] We have not captured the design of the study, which also may be relevant: those studies with a focus on drug discovery, basic cell biology, or the first in vivo testing of a new agent, are more likely to be applied to logistically favorable models, such as subcutaneous cell lines, than confirmatory studies whose main focus is to directly inform a clinical trial.[17]

## Glioma Model Applications

We observed, in both analyses, a dominance of traditional cell lines, often a single line used in isolation. While newer paradigms have been used in the last decade and there appears to be enthusiasm for them, their introduction has been eclipsed by a large parallel increase in the use of traditional cell lines.

The ongoing dominance of U87 has been particularly surprising: clearly favored because of its availability, familiarity, reliable tumor seeding, and range of available subtypes.[31] U87 cells have a number of well-known disadvantages that we feel seriously undermine the validity of conclusions made when using them in vivo. Specifically, the precise histological diagnosis of the original U87 tumor is unknown. It is known to show, as with many other immortalized lines, a drift of karyotype over the years. Comparison of commercially available U87 cells with those deep-frozen in Uppsala in 1968 has revealed a large discrepancy in genotype and mitochondrial DNA, implying the former has either changed substantially or originated from a different, unidentifiable, source due to contamination[32] and commercially available cells are more cytogenetically aberrant than other comparable cells.[33] U87 tumors barely invade brain parenchyma, a cardinal feature of human GBM.[34–37]

Many of these issues persist for other cell lines—contamination is a known issue for all cells maintained in vitro.[28,29,38] The models we encountered most commonly were all developed decades ago, following which the diagnostic criteria for glioblastoma have changed, critically with the integration of key genetic and molecular tumor features.[39] While these have been inferred or analyzed recently,[40–42] it cannot be assumed that these features were present in the primary samples given the genetic instability known to affect these models.[43] Cell lines typically display low intratumoral heterogeneity, unlike human disease.[17] The culture conditions have been shown to affect the genotype of tumor cells in vitro, and older lines predate developments in this area.[16,44,45]

We observed very few publications validating their models—while cell identity is assured by suppliers, misidentification remains a major problem for cell lines.[38] Overall, only 10.2% of publications offered validation of the cell line or determination of key molecular features (see Figure 5D). The Nature Publishing Group now requires STR validation of cell lines and this represents the first institutional step to combat this issue.

Xenograft cell lines have become favored over syngenic models, with U87 replacing C6 as the model of choice in most experiments; the latter was initially favored because of its favorable growth, histology, and genotype.[46] While advantageous in terms of cellular disease recapitulation, the requirement for the use of immunocompromised animals renders xenograft cells limited for the study of brain–tumor interactions and immune-based treatments. Many other syngenic lines lack the histological features of GBM.[47] The GL261 model has become the model of choice for immunocompetent experiments, presumably for logistical reasons, although there are still difficulties when testing human immune interventions on mice without a humanized immune system.[16,48]

We anecdotally noted that recent articles published in influential or reputable journals were more likely to use advanced models—and, conversely, those reporting traditional cell lines only were often in low-impact journals. Unfortunately, we were unable to quantify this by, for example, correlation with journal impact factor as we were unable to automate the extraction of this data.

We were surprised to see that, despite much attention in recent cancer literature, the uptake of PDX models has been somewhat muted, with only around 10% of modern experiments using them. However, this is comparable to other cancer areas.[49] These models have advantages, including the preservation of tumor microenvironment, better recapitulation of tumor growth, heterogeneity, and

invasion,[50–53] and, critically, a modern correlation with the sample donor's disease—clinicopathological features, multiomic analyses, prognosis.[17] With current advances in understanding cancer biology, novel treatments are often tailored to tumors with specific genomic, transcriptomic, or metabolomic profiles.[21,51,53] Correspondingly, valid models should have clear descriptions of these parameters. While traditionally laborious to develop, and difficult to exchange commercially, PDX procedure standardizations appear to have allayed these issues,[54–56] but access to fresh tissue remains a limitation. Other problems include the possibility for genetic drift (although minimal if passage numbers are limited[57,58]) and an observation that xenograft tumors arise from an aggressive subpopulation of cells from the tumor specimen.[59,60] Another key advantage is the relative ease of using tumors from multiple patients in parallel, allowing for the introduction of intertumoral heterogeneity. We were surprised to see that the majority of recent publications (21/33 since 2010, 51.5%) used PDX tumors derived from a single specimen.

Similarly, GEMM models have clear advantages given the innate tailoring of key genetic features. Tumors arising in each individual display heterogeneity, and the host's intact immune system remains advantageous for the study of both tumorigenesis and brain–tumor interactions. Despite encountering a description of their use in the 1990s, and an apparent acknowledgment of their advantages throughout the literature, we were surprised to observe limited uptake.

While we did not focus on the specifics of study design, we extracted basic data pertaining to article focus, with a hypothesis that those more likely to inform clinical trials—that is, testing treatments—might be more likely to model glioblastoma with greater precision. Rather, we observed that those testing a treatment used cell lines more frequently. Similarly, there has been no clear trend in the way models have been applied according to study focus, with single cell lines used in a comparable number of treatments (77.9%) and other (78.7%) studies since 2010.

### How Might We Improve?

Ultimately, it appears that the construct validity of much of the preclinical glioma literature is poor: most current studies still use what we argue are now outdated cell lines that lack many key behaviors and features of human glioblastoma, are of uncertain origin and have shifting karyotype and behavior. Many newer cell lines have been described that are advantageous in each of these ways, but also with a clear characterization of their clinicopathological and genomic features at source.[17,23,61,62] A shift to the use of modern, relevant models—maintained appropriately and with transparent characterization—would improve the relevance and reproducibility of future animal research. The development of cell line repositories should help in this regard.[63–65] It is our view that traditional xenograft models like U87 should no longer have a place in this field.

Another key concern is the lack of heterogeneity in these studies: human glioblastomas are known to have a high degree of variability both within a tumor and between individuals. This is not captured in studies using a single cell line in vivo—where genetically identical tumors are tested in homogenous host and environmental conditions in a fashion known to have poor reproducibility. In our sample, 517/715 (72.3%) of publications conducted their experiments in this manner and we believe that conclusions pertaining to the treatment or biology of glioblastoma in general cannot be reliably made from a study of this design. There has been a tendency for recent experiments to include multiple models in vitro, which we would consider an improvement but still lack the vital features that in vivo modeling provides.

We envisage the optimal preclinical study should adopt the structure of the so-called "preclinical trial"—with measures to introduce heterogeneity and minimize bias in a manner comparable to a clinical trial. That is, using high-quality tumor modeling, measures should be taken to introduce tumor heterogeneity. Preclinical and co-clinical trials have been used to great effect in other cancer areas and show great promise in the application of animal research to personalized cancer therapy.[49,51] While the use of multiple PDX or GEMM constructs is favorable, we accept the associated practical challenges, especially for earlier-stage preclinical experiments. As such, the use of more than 1 modern cell line in a justified manner seems a reasonable approach and should yield improvements in the validity and reproducibility of results.[66,67]

## Conclusions

We have demonstrated that the volume of preclinical experimentation has increased dramatically in recent years. While advanced models are described and available for recapitulation of glioblastoma in animals, the majority of publications model glioblastoma poorly in vivo. The resultant output therefore inevitably has severe limitations with validity and reproducibility. We have provided some practical suggestions that should improve future rodent modeling and would support the withdrawal of the older glioma xenograft lines.

## Supplementary Material

Supplementary material is available online at *Neuro-Oncology Advances* (https://academic.oup.com/noa).


## Keywords

cell lines | glioblastoma | preclinical models | reproducibility | systematic review


## Conflict of interest statement

None of the authors have any actual or potential conflicts of interest.

## Funding

E.S. is funded by grants entitled: "Improving Reproducibility In SciencE; iRISE" - Horizon Europe, and "Improving The Credibility And Utility Of In Vitro Research" – Medical Research Council.

## Authorship statement

Conception (T.H., E.S.), design (T.H.), literature search (T.H., E.W.), study screening (T.H., E.W., D.B.), data extraction (T.H., D.B.), data analysis (T.H.), interpretation (T.H.), write up (T.H., E.W., D.B., E.S.).

## Data availability

All raw data used for this study are included in the Supplementary Material.

## Affiliations

Patrick G Johnson Centre for Cancer Research, Queens University Belfast, Belfast, UK (T.C.H.); Department of Neurosurgery, Royal Victoria Hospital, Belfast, UK (T.C.H., D.B.); Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK (E.W., E.S.S.)

## References

1. Yuan X, Curtin J, Xiong Y, et al. Isolation of cancer stem cells from adult glioblastoma multiforme. *Oncogene.* 2004;23(58):9392–9400.
2. Singh SK, Hawkins C, Clarke ID, et al. Identification of human brain tumour initiating cells. *Nature.* 2004;432(7015):396–401.
3. Kondo T. Brain cancer stem-like cells. *Eur J Cancer.* 2006;42(9):1237–1242.
4. Brennan CW, Verhaak RG, McKenna A, et al.; TCGA Research Network. The somatic genomic landscape of glioblastoma. *Cell.* 2013;155(2):462–477.
5. Aldape K, Zadeh G, Mansouri S, Reifenberger G, von Deimling A. Glioblastoma: pathology, molecular mechanisms and markers. *Acta Neuropathol.* 2015;129(6):829–848.
6. Stupp R, Mason WP, van den Bent MJ, et al.; European Organisation for Research and Treatment of Cancer Brain Tumor and Radiotherapy Groups. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987–996.
7. Vesterinen HM, Sena ES, ffrench-Constant C, et al. Improving the translational hit of experimental treatments in multiple sclerosis. *Mult Scler.* 2010;16(9):1044–1055.
8. Sena ES, Currie GL, McCann SK, Macleod MR, Howells DW. Systematic reviews and meta-analysis of preclinical studies: why perform them and how to appraise them critically. *J Cereb Blood Flow Metab.* 2014;34(5):737–742.
9. Macleod MR, Lawson McLean A, Kyriakopoulou A, et al. Risk of bias in reports of in vivo research: a focus for improvement. *PLoS Biol.* 2015;13(10):e1002273.
10. Ellis LM, Fidler IJ. Finding the tumor copycat. Therapy fails, patients don't. *Nat Med.* 2010;16(9):974–975.
11. Johnson JI, Decker S, Zaharevitz D, et al. Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *Br J Cancer.* 2001;84(10):1424–1431.
12. Benda P, Lightbody J, Sato G, Levine L, Sweet W. Differentiated rat glial cell strain in tissue culture. *Science.* 1968;161(3839):370–371.
13. Ponten J, Macintyre EH. Long term culture of normal and neoplastic human glia. *Acta Pathol Microbiol Scand.* 1968;74(4):465–486.
14. Hirst TC, Vesterinen HM, Conlin S, et al. A systematic review and meta-analysis of gene therapy in animal models of cerebral glioma: why did promise not translate to human therapy? 2015.
15. Hirst TC, Vesterinen HM, Sena ES, et al. Systematic review and meta-analysis of temozolomide in animal models of glioma: was clinical efficacy predicted? *Br J Cancer.* 2013;108(1):64–71.
16. Lenting K, Verhaak R, Ter Laan M, Wesseling P, Leenders W. Glioma: experimental models and reality. *Acta Neuropathol.* 2017;133(2):263–282.
17. Gómez-Oliva R, Domínguez-García S, Carrascal L, et al. Evolution of experimental models in the study of glioblastoma: toward finding efficient treatments. *Front Oncol.* 2020;10(1):614295.
18. Amarasingh S, Macleod MR, Whittle IR. What is the translational efficacy of chemotherapeutic drug research in neuro-oncology? A systematic review and meta-analysis of the efficacy of BCNU and CCNU in animal models of glioma. *J Neurooncol.* 2009;91(2):117–125.
19. Jue TR, Sena ES, Macleod MR, McDonald KL, Hirst TC. A systematic review and meta-analysis of topoisomerase inhibition in pre-clinical glioma models. *Oncotarget.* 2018;9(13):11387–11401.
20. Oldrini B, Curiel-García A, Marques C, et al. Somatic genome editing with the RCAS-TVA-CRISPR-Cas9 system for precision tumor modeling. *Nat Commun.* 2018;9(1):1466.
21. Howley R, Kinsella P, Buckley PG, et al. Comparative genomic and proteomic analysis of high grade glioma primary cultures and matched tumor in situ. *Exp Cell Res.* 2012;318(17):2245–2256.
22. Pollard SM, Yoshikawa K, Clarke ID, et al. Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. *Cell Stem Cell.* 2009;4(6):568–580.
23. Xie Y, Bergström T, Jiang Y, et al. The human glioblastoma cell culture resource: validated cell models representing all molecular subtypes. *EBioMedicine.* 2015;2(10):1351–1363.
24. van de Schoot R, de Bruin J, Schram R, et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nat Mach Intell.* 2021;3(2):125–133.
25. Currie GL, Angel-Scott HN, Colvin L, et al. Animal models of chemotherapy-induced peripheral neuropathy: a machine-assisted systematic review and meta-analysis. *PLoS Biol.* 2019;17(5):e3000243.
26. Hirst TC, Emma W, Browne D. Systematic review of animal glioblastoma modelling. 2022. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022279761. Accessed December 10, 2024.
27. Bannach-Brown A, Przybyła P, Thomas J, et al. Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. *Syst Rev.* 2019;8(1):23.
28. Capes-Davis A, Theodosopoulos G, Atkin I, et al. Check your cultures! A list of cross-contaminated or misidentified cell lines. *Int J Cancer.* 2010;127(1):1–8.

29. Bady P, Diserens AC, Castella V, et al. DNA fingerprinting of glioma cell lines and considerations on similarity measurements. *Neuro Oncol.* 2012;14(6):701–711.
30. Bahor Z, Liao J, Currie G, et al. Development and uptake of an online systematic review platform: the early years of the CAMARADES Systematic Review Facility (SyRF). *BMJ Open Sci.* 2021;5(1):e100103.
31. Bairoch A. The cellosaurus, a cell-line knowledge resource. *J Biomol Tech.* 2018;29(2):25–38.
32. Allen M, Bjerke M, Edlund H, Nelander S, Westermark B. Origin of the U87MG glioma cell line: good news and bad news. *Sci Transl Med.* 2016;8(354):354re353.
33. Clark MJ, Homer N, O'Connor BD, et al. U87MG decoded: the genomic sequence of a cytogenetically aberrant human cancer cell line. *PLoS Genet.* 2010;6(1):e1000832.
34. Engebraaten O, Hjortland GO, Hirschberg H, Fodstad O. Growth of precultured human glioma specimens in nude rat brain. *J Neurosurg.* 1999;90(1):125–132.
35. Halle B, Thisgaard H, Hvidsten S, et al. Estimation of tumor volumes by 11C-MeAIB and 18F-FDG PET in an orthotopic glioblastoma rat model. *J Nucl Med.* 2015;56(10):1562–1568.
36. Candolfi M, Curtin JF, Nichols WS, et al. Intracranial glioblastoma models in preclinical neuro-oncology: neuropathological characterization and tumor progression. *J Neurooncol.* 2007;85(2):133–148.
37. Schulz JA, Rodgers LT, Kryscio RJ, Hartz AMS, Bauer B. Characterization and comparison of human glioblastoma models. *BMC Cancer.* 2022;22(1):844.
38. Hughes P, Marshall D, Reid Y, Parkes H, Gelber C. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? *Biotechniques.* 2007;43(5):575, 577, 577–578, 581–572 passim. 8, 581.
39. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol.* 2021;23(8):1231–1251.
40. Ichimura K, Pearson DM, Kocialkowski S, et al. IDH1 mutations are present in the majority of common adult gliomas but rare in primary glioblastomas. *Neuro Oncol.* 2009;11(4):341–347.
41. Ryu CH, Yoon WS, Park KY, et al. Valproic acid downregulates the expression of MGMT and sensitizes temozolomide-resistant glioma cells. *J Biomed Biotechnol.* 2012;2012(1):1–9.
42. Nusinow DP, Szpyt J, Ghandi M, et al. Quantitative proteomics of the cancer cell line encyclopedia. *Cell.* 2020;180(2):387–402.e16.
43. Torsvik A, Stieber D, Enger P, et al. U-251 revisited: genetic drift and phenotypic consequences of long-term cultures of glioblastoma cells. *Cancer Med.* 2014;3(4):812–824.
44. Reynolds BA, Weiss S. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science.* 1992;255(5052):1707–1710.
45. Lee J, Kotliarova S, Kotliarov Y, et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell.* 2006;9(5):391–403.
46. Giakoumettis D, Kritis A, Foroglou N. C6 cell line: the gold standard in glioma research. *Hippokratia.* 2018;22(3):105–112.
47. Barth RF, Kaur B. Rat brain tumor models in experimental neuro-oncology: the C6, 9L, T9, RG2, F98, BT4C, RT-2 and CNS-1 gliomas. *J Neurooncol.* 2009;94(3):299–312.
48. Szatmári T, Lumniczky K, Désaknai S, et al. Detailed characterization of the mouse glioma 261 tumor model for experimental glioblastoma therapy. *Cancer Sci.* 2006;97(6):546–553.
49. Abdolahi S, Ghazvinian Z, Muhammadnejad S, et al. Patient-derived xenograft (PDX) models, applications and challenges in cancer research. *J Transl Med.* 2022;20(1):206.
50. Hidalgo M, Amant F, Biankin AV, et al. Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov.* 2014;4(9):998–1013.
51. Gao H, Korn JM, Ferretti S, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nat Med.* 2015;21(11):1318–1325.
52. Vaubel RA, Tian S, Remonde D, et al. Genomic and phenotypic characterization of a broad panel of patient-derived xenografts reflects the diversity of glioblastoma. *Clin Cancer Res.* 2020;26(5):1094–1104.
53. Liu Y, Wu W, Cai C, et al. Patient-derived xenograft models in cancer therapy: technologies and applications. *Signal Transduct Target Ther.* 2023;8(1):160.
54. Gürsel DB, Beyene RT, Hofstetter C, et al. Optimization of glioblastoma multiforme stem cell isolation, transfection, and transduction. *J Neurooncol.* 2011;104(2):509–522.
55. Evrard YA, Srivastava A, Randjelovic J, et al.; NCI PDXNet Consortium. Systematic establishment of robustness and standards in patient-derived xenograft experiments and analysis. *Cancer Res.* 2020;80(11):2286–2297.
56. Antunes L, Angioi-Duprez KS, Bracard SR, et al. Analysis of tissue chimerism in nude mouse brain and abdominal xenograft models of human glioblastoma multiforme: what does it tell us about the models and about glioblastoma biology and therapy? *J Histochem Cytochem.* 2000;48(6):847–858.
57. Goplen D, Wang J, Enger P, et al. Protein disulfide isomerase expression is related to the invasive properties of malignant glioma. *Cancer Res.* 2006;66(20):9895–9902.
58. Kogiso M, Qi L, Braun F, et al. In vivo evolution of a novel patient-derived orthotopic xenograft (PDOX) mouse models of pediatric pleomorphic xanthoastrocytoma harboring BRAF V600E mutation. *Neuro Oncol.* 2016;18(Suppl 3):iii83.1–iiiii83.
59. Laks DR, Masterman-Smith M, Visnyei K, et al. Neurosphere formation is an independent predictor of clinical outcome in malignant glioma. *Stem Cells.* 2009;27(4):980–987.
60. Beier D, Hau P, Proescholdt M, et al. CD133(+) and CD133(-) glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer Res.* 2007;67(9):4010–4015.
61. Binder ZA, Wilson KM, Salmasi V, et al. Establishment and biological characterization of a panel of glioblastoma multiforme (GBM) and GBM variant oncosphere cell lines. *PLoS One.* 2016;11(3):e0150271.
62. Cruvinel Carloni A, José da Silva Oliveira R, Oliveira Silva VA, et al. Omics profile of two immortalized Brazilian glioblastoma cell cultures. *Clin Cancer Res.* 2018;24(1_Supplement):A29.
63. Zhang L, Yu H, Yuan Y, et al. The necessity for standardization of glioma stem cell culture: a systematic review. *Stem Cell Res Ther.* 2020;11(1):84.
64. Ghandi M, Huang FW, Jané-Valbuena J, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature.* 2019;569(7757):503–508.
65. Gonçalves E, Poulos RC, Cai Z, et al. Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell.* 2022;40(8):835–849.e8.
66. Kane PB, Kimmelman J. Is preclinical research in cancer biology reproducible enough? *Elife.* 2021;10(1):e67527.
67. Hirst JA, Howick J, Aronson JK, et al. The need for randomization in animal trials: an overview of systematic reviews. *PLoS One.* 2014;9(6):e98856.