











ORIGINAL ARTICLE

Inter-pathologist agreement on diagnosis, classification and grading of canine glioma

Gregory A. Krane^{1,2,3}  | Keith R. Shockley⁴  | David E. Malarkey¹  |
 Andrew D. Miller⁵  | C. Ryan Miller⁶  | Debra A. Tokarz^{7,8}  |
 Heather L. Jensen¹  | Kyathanahalli S. Janardhan⁹  | Matthew Breen²  |
 Christopher L. Mariani^{3,10} 

Correspondence

Christopher L. Mariani, 1060 William Moore Drive, Raleigh, NC 27607, USA.
 Email: clmarian@ncsu.edu

Present address

Gregory A. Krane, Moderna, Cambridge, Massachusetts, USA

Kyathanahalli S. Janardhan, Merck, West Point, Pennsylvania, USA.

Funding information

Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University; National Institute of Environmental Health Sciences

Abstract

Histopathological evaluation of tumours is a subjective process, but studies of inter-pathologist agreement are uncommon in veterinary medicine. The Comparative Brain Tumour Consortium (CBTC) recently published diagnostic criteria for canine gliomas. Our objective was to assess the degree of inter-pathologist agreement on intracranial canine gliomas, utilising the CBTC diagnostic criteria in a cohort of eighty-five samples from dogs with an archival diagnosis of intracranial glioma. Five pathologists independently reviewed H&E and immunohistochemistry sections and provided a diagnosis and grade. Percentage agreement and kappa statistics were calculated to measure inter-pathologist agreement between pairs and amongst the entire group. A consensus diagnosis of glioma subtype and grade was achieved for 71/85 (84%) cases. For these cases, percentage agreement on combined diagnosis (subtype and grade), subtype only and grade only were 66%, 80% and 82%, respectively. Kappa statistics for the same were 0.466, 0.542 and 0.516, respectively. Kappa statistics for oligodendroglioma, astrocytoma and undefined glioma were 0.585, 0.566 and 0.280 and were 0.516 for both low-grade and high-grade tumours. Kappa statistics amongst pairs of pathologists for combined diagnosis varied from 0.352 to 0.839. 8% of archival oligodendrogliomas and 61% of archival astrocytomas were reclassified as another entity after review. Inter-pathologist agreement utilising CBTC guidelines for canine glioma was moderate overall but varied from fair to almost perfect

Abbreviations: ACVP, American College of Veterinary Pathologists; BCa, adjusted bootstrap percentile; CBTC, Comparative Brain Tumour Consortium; CNPase, 2',3'-cyclic-nucleotide 3'-phosphodiesterase; GFAP, glial fibrillary acidic protein; H&E, haematoxylin and eosin; MRI, magnetic resonance imaging; Olig2, oligodendrocyte transcription factor 2.

Gregory A. Krane and Keith R. Shockley are co-first authors for this manuscript.

Portions of these data were presented in oral presentations delivered at the 2021 National Toxicology Program's Satellite Symposium at the Annual Meeting of the Society of Toxicologic Pathology (Virtual) and at the 2021 Annual Research Day Symposium at North Carolina State University's College of Veterinary Medicine (Virtual), as well as at the 2020 American College of Veterinary Pathologists' Annual Meeting in abstract/poster form (Virtual). General statements from the Society of Toxicologic Pathology presentation are briefly shared in a publication (Elmore SA, Choudhary S, Krane GA, Plumlee Q, Quist EM, Suttie AW, Tokarz DA, Ward JM, Cora M. Proceedings of the 2021 National Toxicology Program Satellite Symposium. *Toxicol Pathol.* 2021;49(8):1344–1367), though no quantitative data are shared in that manuscript, the goal of which was to serve as general conference presentation proceedings.

For affiliation refer to page 888

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Veterinary and Comparative Oncology* published by John Wiley & Sons Ltd.

between pairs of pathologists. Agreement was similar for oligodendrogliomas and astrocytomas but lower for undefined gliomas. These results are similar to pathologist agreement in human glioma studies and with other tumour entities in veterinary medicine.

KEYWORDS

brain, cancer, dog, histopathology, immunohistochemistry

1 | INTRODUCTION

Gliomas are common in dogs, comprising approximately 30%–40% of intracranial tumours, and are particularly frequent in brachycephalic breeds such as the Boxer, Boston terrier and English and French bulldogs.^{1,2} Though MRI can document the presence of an intracranial mass, histopathologic examination is necessary to confirm the diagnosis of a glioma, including its subtype (oligodendroglioma, astrocytoma or undefined glioma) and grade (low-grade or high-grade).¹ Due to the inherent subjectivity of microscopic evaluation, morphologic assessment is subject to both intra- and inter-observer variability. Therefore, assessment of inter-pathologist diagnostic agreement is necessary to evaluate the utility of diagnostic criteria.

Historically, the diagnosis of canine glioma has been based on World Health Organization guidelines.³ That system, which is now over two decades old, splits canine gliomas into thirteen different subtypes. Newer consensus recommendations have recently been provided by the Comparative Brain Tumour Consortium (CBTC), a panel of veterinary and physician pathologists whose members consolidated potential diagnoses into three broad categories of oligodendroglioma, astrocytoma and undefined glioma, each of which are further classified into low-grade and high-grade tumours.⁴ However, there are few studies examining diagnostic agreement for brain tumours amongst veterinary pathologists.^{4,5}

Understanding the degree of pathologist agreement for canine glioma utilising the current guidelines is critical not only for diagnostic pathology, but also for clinical research, for which diagnostic criteria must be harmonised in order for multi-institutional studies to be successful. Percentage agreement and kappa statistics are common methods of assessing such agreement. The kappa statistic estimates the degree to which different raters agree compared to what would be predicted by chance alone.⁶ Kappa statistics can range from +1 (complete agreement) to 0 (random chance) to –1 (complete disagreement). The agreement levels for the kappa statistic are defined as poor (0), slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), almost perfect (0.81 ≤ 1.0) and perfect (1.0) agreement.⁷

In humans, differences in pathologist assessment can lead to substantial effects on patient diagnosis and case management. Approximately 20%–40% of human gliomas are reclassified after independent subspecialist review, largely due to the subjective nature of diagnostic criteria.^{8,9} The goal of this study was to evaluate inter-pathologist percentage agreement and kappa statistics for the diagnosis of tumour

subtype and grade in a canine glioma cohort, using the CBTC guidelines.⁴

2 | MATERIALS AND METHODS

2.1 | Case identification and sample preparation

A retrospective search utilising the terms glioma, oligodendroglioma, astrocytoma, mixed glioma, oligoastrocytoma, glioblastoma and gliomatosis cerebri was performed to identify canine samples diagnosed as intracranial glioma from the pathology archives at North Carolina State University's College of Veterinary Medicine (2006–2018), yielding 85 cases for review. Brains had been fixed whole in 10% neutral buffered formalin for 48–72 hours, with samples routinely processed for histology, and embedded in paraffin. In order to maintain histologic and immunohistochemical consistency between cases, new samples were sectioned at 5 µm, mounted on charged glass slides, stained with haematoxylin and eosin (H&E) or prepared for immunohistochemical detection of oligodendrocyte transcription factor 2 (Olig2), glial fibrillary acidic protein (GFAP), 2',3'-cyclic-nucleotide 3'-phosphodiesterase (CNPase) and Ki-67 (Table S1). Slides were then digitally scanned (40X magnification) to an electronic database for diagnostic review (Aperio AT Turbo, Leica Biosystems, Buffalo Grove, IL, United States). The choice of immunohistochemical markers was based on CBTC guidelines for the diagnosis of canine glioma.⁴

2.2 | Diagnostic review

A panel of four board-certified (ACVP) veterinary anatomic pathologists (GAK, DEM, ADM, DAT) and one physician neuropathologist (CRM) independently reviewed all 85 specimens, which included the same H&E-stained sections, as well as sections immunohistochemically labelled with Olig2, GFAP, CNPase and Ki-67. Pathologists were instructed to utilise the CBTC diagnostic criteria to diagnose each case as oligodendroglioma, astrocytoma, undefined glioma, or other (non-glioma) and for cases diagnosed as glioma, to provide a grade (low- or high-grade).⁴ Pathologists, who were blinded to each other's analyses, provided their assessments to a single pathologist (GAK), who collated the data after recording his own diagnoses (so as not to be biased by others).

2.3 | Statistical analyses

Diagnoses from the pathologists were recorded as low-grade oligodendroglioma, high-grade oligodendroglioma, low-grade astrocytoma, high-grade astrocytoma, low-grade undefined glioma, high-grade undefined glioma or other. Cases for which three or more pathologists agreed on combined diagnosis and grade qualified as achieving consensus. Concordance rates (5/5 pathologists (100%); 4/5 (80%) or 3/5 (60%)) were calculated for cases achieving consensus.

The proportion of agreeing pairs out of all possible rating pairs can be computed for each case. The overall percentage agreement of pathologists' diagnoses was calculated as the mean of the proportion of agreeing pairs across all raters.¹⁰ This calculation was performed on the cases achieving consensus for assessments of glioma subtype, glioma grade and combined glioma subtype and grade; cases including a response of 'other' (non-glioma) were not included in calculations involving grade.

Cohen's kappa¹¹ was used to assess inter-pathologist agreement beyond chance for individual pairs of pathologists and Fleiss' kappa¹² was used to assess inter-pathologist agreement beyond chance across the entire group of pathologists. Kappa analyses were based on all cases achieving consensus for combined subtype and grade, but cases including a response of 'other' (non-glioma) were not included in calculations involving grade. Bootstrap 95% confidence intervals for kappa were calculated based on the adjusted bootstrap percentile (BCa) method with 1,000 bootstrap replicates using the *R/boot* package (R Core Team 2019, Vienna, Austria). The BCa approach adjusts for bias and skewness in the bootstrap distribution.¹³ All calculations were conducted in the R software environment (R Core Team 2019, Vienna, Austria).

Archival diagnoses that were changed after review in this study were identified and reclassification rates for archival diagnostic categories were calculated. For these calculations, archival diagnoses of glioblastoma were considered to be analogous to high-grade astrocytoma and diagnoses of mixed glioma or oligoastrocytoma were considered to be analogous to undefined glioma.

3 | RESULTS

The glioma subtype and grade diagnoses made by each pathologist, original archival diagnoses and resulting consensus diagnoses are shown in Figure 1. Pathologist diagnoses considering only subtype or only grade are shown in Figures S1 and S2, respectively. Although there was variability amongst pathologists, a consensus diagnosis of glioma subtype and grade was achieved in 71/85 (84%) cases (Figure 1). Consensus glioma diagnoses were low-grade oligodendroglioma ($n = 9$, 13%), high-grade oligodendroglioma ($n = 43$, 61%), low-grade astrocytoma ($n = 8$, 11%), high-grade astrocytoma ($n = 10$, 14%) and high-grade undefined glioma ($n = 1$, 1%). For the fourteen remaining cases, five had a consensus diagnosis of glioma but not a consensus on subtype or grade, two had a consensus diagnosis of histiocytic sarcoma, two had a consensus diagnosis of encephalitis, one

had a consensus diagnosis of meningioangiomatosis, one was a poorly differentiated neoplastic mass with no diagnostic consensus, and three did not have any tumour in available sections. Considering archival diagnoses, 19 astrocytomas, 3 oligodendrogliomas, 8 mixed gliomas and 7 samples diagnosed as gliomatosis were reclassified after review by the pathology panel. These archival diagnoses and their reclassifications are shown in Table 1.

Percentage agreement amongst all pathologists on combined glioma subtype and grade diagnosis for the 71 cases that reached consensus was 66% and for subtype alone and grade alone were 80% and 82%, respectively. Concordance rates for each subtype and grade for the cases that reached consensus are shown in Table 2. Four cases that included a diagnosis of 'other' (non-glioma) were not included in calculations involving grade, leaving 67 cases available for these calculations.

Kappa analyses were based on the 71 cases achieving a consensus diagnosis for combined subtype and grade, although four cases that included a diagnosis of 'other' (non-glioma) were not included in calculations involving grade, leaving 67 cases available for these calculations. Overall pathologist agreement for glioma subtype, grade and combined diagnosis as measured by Fleiss' kappa statistic was moderate and similar between subtype and grade (Table 3), while agreement between pairs of pathologists as measured by Cohen's kappa statistic ranged from fair to almost perfect (Table 4).

4 | DISCUSSION

In this study, a consensus diagnosis of glioma subtype and grade was achieved in 71/85 (84%) cases by a panel of veterinary and physician pathologists utilising the CBTC classification criteria. Considering original archival diagnoses, the reclassification rate was higher for archival diagnoses of mixed gliomas (100%) and astrocytomas (61%) than for oligodendrogliomas (8%). Overall inter-pathologist agreement for canine glioma subtype and grade was moderate (kappa = 0.542 and 0.516, respectively), while individual inter-pathologist agreements ranged from fair (kappa = 0.310) to almost perfect (kappa = 0.905). Percentage agreements for inter-pathologist pairs were consistent with their corresponding Cohen's kappa statistics. Agreement for combined subtype and grade was also moderate (kappa = 0.466). Agreement was similar between oligodendroglioma (kappa = 0.585, moderate) and astrocytoma (kappa = 0.566, moderate) but greater than that for undefined gliomas (kappa = 0.280, fair). Agreement for low-grade and high-grade tumours was identical (kappa = 0.516, moderate).

Factors shown to affect inter-observer diagnostic agreement for human gliomas include experience level of the pathologist, institution type and glioma subtype. Inter-observer agreement is higher between pathologists working at tertiary care centres (universities and referral centres) and neuropathology specialists than it is between neuropathologists and generalist pathologists working at local hospitals.^{14,15} In another study, when separate panels of neuropathologists and generalist pathologists reviewed the same astrocytoma case series,

Archival Diagnosis	Pathologist 1	Pathologist 2	Pathologist 3	Pathologist 4	Pathologist 5	Consensus Diagnosis
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Astrocytoma						*
Astrocytoma						*
Mixed Glioma		NT				*
Mixed Glioma						*
Gliomatosis Cerebri						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Oligodendroglioma						
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Glioblastoma Multiforme						*
Glioblastoma Multiforme						*
Glioblastoma Multiforme						*
Glioblastoma Multiforme						*
Mixed Glioma						*
Mixed Glioma						*
Mixed Glioma			EMB			*
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Glioblastoma Multiforme						*
Gliomatosis Cerebri		NT				*
Gliomatosis Cerebri		NT				*
Astrocytoma						*
Astrocytoma						*
Astrocytoma						*
Glioblastoma Multiforme						*
Glioblastoma Multiforme						*
Glioblastoma Multiforme						*
Glioblastoma Multiforme						*
Oligodendroglioma						*
Mixed Glioma						*
Gliomatosis Cerebri						*
Astrocytoma						*
Oligodendroglioma	INF	INF	INF			* Key
Oligodendroglioma	NOS	LSA	NOS		RC	*
Astrocytoma						*
Astrocytoma	EMB		EMB			*
Astrocytoma	EMB		EMB			*
Astrocytoma	HS	HS	HS			*
Astrocytoma	HS		HS			*
Astrocytoma	MGA		MGA		MGA	*
Glioblastoma Multiforme	HS	HS	HS			*
Mixed Glioma						*
Mixed Glioma	NT	NT	NT	NT	NT	*
Gliomatosis Cerebri	INF	INF	INF			*
Gliomatosis Cerebri	NT	NT	NT	NT	NT	*
Gliomatosis Cerebri	NT	NT	NT	NT		*

FIGURE 1 Archival diagnoses, individual diagnoses by each pathologist and consensus diagnoses for 85 cases from the original archival cohort. Cases where the archival diagnosis differed from the diagnosis after pathologist review are marked by an asterisk (*). Abbreviations: EMB, embryonal tumour; HS, histiocytic sarcoma; INF, inflammatory lesion (encephalitis); LSA, lymphosarcoma; MGA, meningioangiomatosis; NOS, tumour not otherwise specified; NT, no tumour in section; RC, round cell tumour (not otherwise specified)

neuropathologists tended to agree more amongst themselves (kappa = 0.63) than they did with generalists (kappa = 0.36), showing the effect of subspecialty training and experience on evaluation of these tumours.¹⁶

Different subtypes of glioma present varying levels of diagnostic difficulty, which also affects inter-observer diagnostic variability. In humans, glioblastoma has high levels of inter-observer agreement compared to other tumour subtypes such as low-grade astrocytoma.¹⁷

TABLE 1 Archival diagnoses and reclassification after pathologist panel review

Reclassified diagnosis	Archival diagnosis			
	Oligodendroglioma n = 39	Astrocytoma n = 31	Mixed glioma n = 8	Gliomatosis cerebri n = 7
Low-grade oligodendroglioma		2	2	1
High-grade oligodendroglioma		8	3	
Low-grade astrocytoma		1 ^c		2
High-grade astrocytoma	1		1	1
High-grade undefined glioma		1		
Oligodendroglioma not otherwise specified ^a		1		
Glioma not otherwise specified ^b		3	1	
Histiocytic sarcoma		2		
Tumour not otherwise specified	1			
Inflammatory lesion	1			1
Meningioangiomatosis		1		
No tumour in section			1	2
Reclassification total	3 (8%)	19 (61%)	8 (100%)	7 (100%)

^aThree of five pathologists agreed on a subtype of oligodendroglioma, but consensus was not reached on grade.

^bAt least three of five pathologists diagnosed the case as some subtype of glioma, but consensus was not achieved on subtype or grade.

^cArchival diagnosis of glioblastoma (analogous to high-grade astrocytoma) was reclassified to low-grade astrocytoma.

TABLE 2 Diagnostic concordance rates for glioma cases achieving consensus on subtype and grade

	Number of cases	Diagnostic concordance ^a		
		100%	80%	60%
Combined diagnosis ^b	71	35%	31%	34%
Oligodendroglioma overall	52	71%	17%	12%
Astrocytoma overall	18	28%	44%	28%
Low-grade glioma overall	17	41%	41%	18%
High-grade glioma overall	54	67%	22%	9%
Low-grade oligodendroglioma	9	22%	44%	33%
High-grade oligodendroglioma	43	47%	28%	26%
Low-grade astrocytoma	8	13%	25%	63%
High-grade astrocytoma	10	20%	30%	50%
High-grade undefined glioma	1	0%	100%	0%

^aNumber of pathologists in agreement (100% = 5/5 pathologists, 80% = 4/5 pathologists, 60% = 3/5 pathologists).

^bConsensus diagnosis on both glioma subtype and grade.

This can be attributed to the dramatic and straightforward histologic features of this neoplasm, such as vast regions of necrosis and overt microvascular proliferation. Diagnosis of other tumour subtypes are often based on subtle histologic features that may result in more variability in individual pathologist interpretation. Gliomas in human patients with mixed morphology (previously referred to oligoastrocytomas and analogous to undefined gliomas in dogs) present a particular diagnostic challenge, as they exhibit both oligodendroglial and astrocytic features. Their diagnostic subjectivity is reflected by low inter-observer agreement compared to other subtypes.^{18,19} Although there were very few undefined gliomas in our case series, our study is in agreement with

such data, as undefined glioma had the lowest kappa statistic of any subtype in our cohort. Cases of recurrent glioma also present diagnostic challenges that can confound inter-observer agreement. Prior administration of radiation therapy and chemotherapy can introduce histologic changes that can complicate assessment of both tumour subtype and grade.^{20,21} For example, it can be difficult to discern whether necrosis in the setting of recurrent tumours is caused by properties inherent to the tumour or by the radiation therapy.²¹

Pathologists evaluate many histologic features during microscopic evaluation. Some of these features are more amenable to agreement than others and include high cellularity, mitoses, endothelial

TABLE 3 Agreement amongst all pathologists on tumour subtype, grade and combined diagnoses considering cases achieving a consensus diagnosis

	Number of cases	Fleiss' kappa statistic (95% confidence interval)	Percentage agreement
Combined overall ^a	71	0.466 (0.422–0.509)	66%
Subtype overall	71	0.542 (0.477–0.606)	80%
Grade overall ^b	67	0.516 (0.440–0.591)	82%
Oligodendroglioma	52	0.585 (0.511–0.659)	
Astrocytoma	18	0.566 (0.492–0.640)	
Undefined glioma	1	0.280 (0.206–0.354)	
Low-grade glioma ^b	14	0.516 (0.440–0.592)	
High-grade glioma ^b	53	0.516 (0.440–0.592)	
Low-grade oligodendroglioma	8	0.380 (0.219–0.548)	
High-grade oligodendroglioma ^b	42	0.545 (0.471–0.619)	
Low-grade astrocytoma	6	0.554 (0.480–0.628)	
High-grade astrocytoma	10	0.430 (0.356–0.504)	
High-grade undefined glioma	1	0.361 ^c	

^aConsensus diagnosis on both glioma subtype and grade.

^bCases with any pathologist response of 'other' (non-glioma) were excluded from kappa analysis and percentage agreement calculation for grade.

^cLow sample size precludes calculation of confidence interval.

proliferation and necrosis.^{22,23} Histologic features of human gliomas frequently subject to inter-pathologist disagreement include pleomorphism, anaplasia, increased vascularity and the presence of subpial or leptomeningeal infiltration.^{22,23,24} Ultimately, it is imperative that diagnostic criteria prioritise features with high levels of agreement in order to result in reproducible diagnoses regardless of the examiner. Assessment of specific histologic criteria was not evaluated in our study, although the CBTC study investigated such factors in canine gliomas and showed that there was substantial agreement by pathologists on the presence of necrosis and microvascular proliferation.⁴

Twenty to forty percent of human gliomas are reclassified after review by neuropathology specialists or a central review board, and although some of these changes are minor, others can result in significant clinical impacts to the patient.^{8,9,15} In our study, we investigated reclassification of archival diagnoses after review by our pathology panel (Table 1). This analysis is limited by the fact that diagnostic criteria were not standardised between pathologists making the archival diagnoses, whereas our pathology review panel used the CBTC glioma diagnostic criteria. Regardless, in our study, cases with an archival diagnosis of mixed glioma or astrocytoma had higher rates of reclassification than did cases with an archival diagnosis of oligodendroglioma (100%, 61% and 8%, respectively). Given that Fleiss' kappa statistic was similar in our study between oligodendroglioma and astrocytoma, these data support the hypothesis that the CBTC initiative may have improved concordance for diagnosis of astrocytoma in particular.

To the authors' knowledge, assessment of diagnostic agreement for brain tumours amongst veterinary pathologists is limited to three studies.^{4,5,25} One study was limited to meningiomas.²⁵ Another study examined multiple types of nervous system tumours in cats and dogs and reported substantial diagnostic agreement based on assessment of H&E sections (kappa = 0.66), with improvement after utilisation of a combination of immunohistochemical and special histochemical

stains (kappa = 0.76).⁵ The high level of agreement in that study may relate in part to utilisation of broad diagnoses across different tumour entities (e.g., glioma, meningioma), as opposed to our study, which further split glioma into subtype and grade.⁵ In the CBTC study, assessment of gliomas by a panel of veterinary and physician neuropathologists showed moderate agreement for oligodendroglioma and astrocytoma subtypes, slight agreement for the undefined glioma subtype and moderate agreement for grade, consistent with our study.⁴ The CBTC study also found that veterinary pathologists are more likely than physician pathologists to make a diagnosis of oligodendroglioma, with physician pathologists showing a bias toward astrocytoma.⁴ Our study did not include enough pathologists to adequately investigate differences in diagnoses between physician and veterinary pathologists.

There are several studies assessing inter-pathologist agreement for other disorders in veterinary medicine. One study investigated agreement between a review panel of two pathologists and the original examiner in amputated canine and feline digits and found complete agreement in 80% of cases, though kappa statistics were not reported.²⁶ Another study, utilising a panel of three observers, investigated intra- and inter-observer agreement in the setting of canine soft tissue sarcoma and reported inter-observer kappa statistics of 0.60 and 0.43 for diagnosis and grade, respectively.²⁷ These results are similar to our findings of 66%, 80% and 82% overall agreement for combined diagnosis, tumour subtype and tumour grade, respectively and kappa statistics of 0.542 and 0.516 for tumour subtype and grade, respectively. Several other veterinary studies investigated inter-pathologist diagnostic agreement for mast cell tumours, and found the agreement is variable depending on the diagnostic system utilised (kappa = 0.503 with no standardised diagnostic criteria and kappa = 0.621 with the Patnaik system).^{28–30} Inter-pathologist agreement as determined by calculation of a kappa statistic for the more

TABLE 4 Agreement between pairs of pathologists on tumour subtype, grade, and combined diagnoses considering cases achieving a consensus diagnosis

	Pathologist pair	Percentage agreement	Cohen's kappa statistic (95% confidence interval)
Combined diagnosis ^a	1-2	69%	0.497 (0.337–0.661)
	1-3	90%	0.839 (0.690–0.932)
	1-4	58%	0.355 (0.196–0.520)
	1-5	66%	0.459 (0.299–0.626)
	2-3	69%	0.508 (0.352–0.680)
	2-4	58%	0.352 (0.209–0.516)
	2-5	62%	0.410 (0.258–0.559)
	3-4	62%	0.428 (0.265–0.592)
	3-5	66%	0.470 (0.318–0.622)
	4-5	59%	0.392 (0.245–0.541)
Subtype only	1-2	85%	0.643 (0.464–0.821)
	1-3	93%	0.843 (0.691–0.965)
	1-4	76%	0.382 (0.138–0.586)
	1-5	80%	0.570 (0.360–0.752)
	2-3	82%	0.589 (0.408–0.756)
	2-4	75%	0.310 (0.102–0.526)
	2-5	80%	0.567 (0.372–0.747)
	3-4	78%	0.439 (0.228–0.656)
	3-5	80%	0.582 (0.372–0.749)
	4-5	76%	0.420 (0.204–0.622)
Grade only ^b	1-2	82%	0.513 (0.215–0.740)
	1-3	97%	0.905 (0.633–1.000)
	1-4	78%	0.481 (0.285–0.690)
	1-5	85%	0.493 (0.191–0.754)
	2-3	85%	0.594 (0.347–0.786)
	2-4	78%	0.504 (0.236–0.686)
	2-5	76%	0.327 (0.061–0.571)
	3-4	81%	0.550 (0.354–0.733)
	3-5	85%	0.493 (0.176–0.743)
	4-5	75%	0.403 (0.199–0.621)

^aConsensus diagnosis on both glioma subtype and grade.

^bCases with any pathologist response of 'other' (non-glioma) were excluded from kappa analysis and percentage agreement calculation for grade.

recent Kiupel classification system³¹ has not been reported, to the authors' knowledge. However, inter-pathologist agreement using a different statistic (Cronbach's alpha) showed agreements ranging from 63.0% to 74.6% (depending on grade) and 96.8% for the Patnaik and Kiupel systems (respectively), although the number of raters was different between these groups (28 vs. 6 pathologists, respectively).³¹ A more recent study comparing these two classification systems showed percent agreement calculations of 72.9% and 77.0% for the Patnaik and Kiupel systems, respectively.³²

There are several limitations in the present study. Agreement between pathologists on individual diagnostic criteria (e.g., microvascular proliferation, necrosis, detection of mitoses) was not assessed, and such an analysis might have provided further insight into the factors driving pathologist disagreements. Agreement before and after evaluation of immunohistochemistry sections was not incorporated

into the study design and might have been useful to assess the effect of the availability of immunohistochemistry on diagnostic agreement. In our experimental design, each pathologist evaluated the slides once. If each pathologist had read each slide multiple times, in different orders, with an interim break between assessments, intra-observer variability could have been assessed. Finally, two of the pathologists in our review panel were also members of the CBTC glioma panel (ADM, CRM), which may have introduced some bias into our study.

Given the moderate level of agreement as defined by kappa statistics, assessment of canine gliomas by a single pathologist is likely sufficient for routine diagnostic pathology, provided the CBTC guidelines are utilised. Our study reports levels of agreement similar to those reported for human gliomas and also similar to agreement reported for other well-established diagnostic systems in veterinary

medicine (canine soft tissue sarcoma and mast cell tumours),^{4,5,8,9,14-16,26-30} Pathologists are still encouraged to consult with other pathologists, including neuropathology subspecialists, as needed, depending on the case. In certain research settings, such as preclinical drug development, where diagnostic precision potentially impacts the health of numerous animals or humans, standard peer-review is regularly utilised and similar approaches are recommended for veterinary clinical trials and other research studies. In this scenario, which is well described in the toxicologic pathology literature, a second pathologist or a panel of pathologists reviews the diagnoses and reconciles disagreements with the original pathologist prior to finalisation of study data.^{33,34} When these differences cannot be reconciled, a larger panel of pathologists (pathology working group) reviews the data and votes on the ultimate diagnoses prior to finalisation of the study.^{33,34}

There are limited data regarding the impact of glioma subtype and grade on survival in dogs. One study of dogs with intracranial gliomas treated with surgical resection and an immunotherapeutic tumour vaccine concluded that those with astrocytomas had longer survival than those with oligodendrogliomas or undefined gliomas and that dogs with low-grade gliomas lived longer than dogs with high-grade tumours.³⁵ If this differential response to therapy and survival of dogs with varied glioma subtypes and grades is supported by additional studies, it would reinforce the importance of providing an accurate glioma diagnosis and consistent agreement between pathologists. Additionally, molecular pathology will likely provide further assistance with regards to providing a diagnosis, prognosis and treatment plan for canine patients with gliomas, as it has for humans,³⁶⁻³⁹ and may even supersede histologic assessment in the future.

ACKNOWLEDGEMENTS

The authors thank Dr. Luke Borst, Sandra Horton, and Joanna Barton (North Carolina State University, Raleigh, NC, United States) for help with sample identification and retrieval, Dr. Michael Nolan (North Carolina State University, Raleigh, NC, United States) for scientific review of earlier phases of the study, Eli Ney (National Institutes of Environmental Health Sciences, Research Triangle Park, NC, United States) for assistance with figure preparation and Dr Cynthia Willson (Integrated Laboratory Systems, Research Triangle Park, NC, United States) and Ms Caroll Co (Social and Scientific Systems, Durham, NC, United States) for internal review of the manuscript.

AFFILIATIONS

¹Cellular and Molecular Pathology Branch, National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

²Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, North Carolina, USA

³Comparative Neuroimmunology and Neuro-Oncology Laboratory, College of Veterinary Medicine, North Carolina State University, Raleigh, North Carolina, USA

⁴Biostatistics and Computational Biology Branch, Division of Intramural Research, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

⁵Section of Anatomic Pathology, Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

⁶Comprehensive Neuroscience Center, O'Neal Comprehensive Cancer Center, Division of Neuropathology, Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, USA

⁷Experimental Pathology Laboratories, Research Triangle Park, North Carolina, USA

⁸Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, North Carolina, USA

⁹Integrated Laboratory Systems, Research Triangle Park, North Carolina, USA

¹⁰Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, North Carolina, USA

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare with regards to the work of this manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE APPROVAL

This study was conducted on samples previously obtained and archived during the course of routine diagnostic evaluations (including necropsies), during which clients gave permission for patient information and archived case material to be used in future research studies; IACUC approval is not required for such studies at our institution. [Correction added on 24 August 2022, after first online publication: The preceding sentence was corrected in this version.]

ORCID

Gregory A. Krane  <https://orcid.org/0000-0002-8325-2435>

Keith R. Shockley  <https://orcid.org/0000-0002-4924-954X>

David E. Malarkey  <https://orcid.org/0000-0002-2573-8533>

Andrew D. Miller  <https://orcid.org/0000-0001-6350-5581>

C. Ryan Miller  <https://orcid.org/0000-0002-0096-8762>

Debra A. Tokarz  <https://orcid.org/0000-0002-1802-7346>

Heather L. Jensen  <https://orcid.org/0000-0002-1712-9682>

Kyathanahalli S. Janardhan  <https://orcid.org/0000-0002-5774-9317>

Matthew Breen  <https://orcid.org/0000-0002-8901-4155>

Christopher L. Mariani  <https://orcid.org/0000-0003-3228-6536>

REFERENCES

1. Miller AD, Miller CR, Rossmeisl JH. Canine primary intracranial cancer: a clinicopathologic and comparative review of glioma, meningioma, and choroid plexus tumors. *Front Oncol.* 2019;9:1151.

2. Truve K, Dickinson P, Xiong A, et al. Utilizing the dog genome in the search for novel candidate genes involved in glioma development-genome wide association mapping followed by targeted massive parallel sequencing identifies a strongly associated locus. *PLoS Genet*. 2016;12:e1006000.
3. Koestner ABT, Fatzner R, Schulman FY, Summers BA, Van Winkle TJ. *Histological Classification of Tumors of the Nervous System of Domestic Animals*. Armed Forces Institute of Pathology; 1999.
4. Koehler JW, Miller AD, Miller CR, et al. A revised diagnostic classification of canine glioma: towards validation of the canine glioma patient as a naturally occurring preclinical model for human glioma. *J Neuro-pathol Exp Neurol*. 2018;77:1039-1054.
5. Belluco S, Avallone G, Di Palma S, et al. Inter- and intraobserver agreement of canine and feline nervous system tumors. *Vet Pathol*. 2019;56:342-349.
6. Funkhouser WK Jr, Hayes DN, Moore DT, et al. Interpathologist diagnostic agreement for non-small cell lung carcinomas using current and recent classifications. *Arch Pathol Lab Med*. 2018;142:1537-1548.
7. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
8. van den Bent MJ. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. *Acta Neuropathol*. 2010;120:297-304.
9. Bruner JM, Inouye L, Fuller GN, Langford LA. Diagnostic discrepancies and their clinical impact in a neuropathology referral practice. *Cancer*. 1997;79:796-803.
10. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. 2008;61:29-48.
11. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37-46.
12. Fleiss JL, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. *Psychol Bull*. 1969;72:323-327.
13. Efron BE. Better bootstrap confidence intervals. *J Am Stat Assoc*. 1987;82:171-185.
14. Gupta T, Nair V, Epari S, Pietsch T, Jalali R. Concordance between local, institutional, and central pathology review in glioblastoma: implications for research and practice: a pilot study. *Neurol India*. 2012;60:61-65.
15. Aldape K, Simmons ML, Davis RL, et al. Discrepancies in diagnoses of neuroepithelial neoplasms: the San Francisco Bay Area Adult Glioma Study. *Cancer*. 2000;88:2342-2349.
16. Prayson RA, Agamanolis DP, Cohen ML, et al. Interobserver reproducibility among neuropathologists and surgical pathologists in fibrillary astrocytoma grading. *J Neurol Sci*. 2000;175:33-39.
17. Scott CB, Nelson JS, Farnan NC, et al. Central pathology review in clinical trials for patients with malignant glioma. A report of Radiation Therapy Oncology Group 83-02. *Cancer*. 1995;76:307-313.
18. Kros JM, Gorlia T, Kouwenhoven MC, et al. Panel review of anaplastic oligodendroglioma from European Organization for Research and Treatment of Cancer Trial 26951: assessment of consensus in diagnosis, influence of 1p/19q loss, and correlations with outcome. *J Neuro-pathol Exp Neurol*. 2007;66:545-551.
19. Krouwer HG, van Duinen SG, Kamphorst W, et al. Oligoastrocytomas: a clinicopathological study of 52 cases. *J Neurooncol*. 1997;33:223-238.
20. Holdhoff M, Ye X, Piotrowski AF, et al. The consistency of neuropathological diagnoses in patients undergoing surgery for suspected recurrence of glioblastoma. *J Neurooncol*. 2019;141:347-354.
21. Kim JH, Bae Kim Y, Han JH, et al. Pathologic diagnosis of recurrent glioblastoma: morphologic, immunohistochemical, and molecular analysis of 20 paired cases. *Am J Surg Pathol*. 2012;36:620-628.
22. Giannini C, Scheithauer BW, Weaver AL, et al. Oligodendrogliomas: reproducibility and prognostic value of histologic diagnosis and grading. *J Neuropathol Exp Neurol*. 2001;60:248-262.
23. Gilles FH, Tavaré CJ, Becker LE, et al. Pathologist interobserver variability of histologic features in childhood brain tumors: results from the CCG-945 study. *Pediatr Dev Pathol*. 2008;11:108-117.
24. Preusser M, Heinzl H, Gelpi E, et al. Histopathologic assessment of hot-spot microvessel density and vascular patterns in glioblastoma: poor observer agreement limits clinical utility as prognostic factors: a translational research project of the European Organization for Research and Treatment of Cancer Brain Tumor Group. *Cancer*. 2006;107:162-170.
25. Belluco S, Marano G, Baiker K, et al. Standardisation of canine meningioma grading: inter-observer agreement and recommendations for reproducible histopathologic criteria. *Vet Comp Oncol*. 2022;20:509-520.
26. Wobeser BK, Kidney BA, Powers BE, et al. Agreement among surgical pathologists evaluating routine histologic sections of digits amputated from cats and dogs. *J Vet Diagn Invest*. 2007;19:439-443.
27. Yap FW, Rasotto R, Priestnall SL, Parsons KJ, Stewart J. Intra- and inter-observer agreement in histological assessment of canine soft tissue sarcoma. *Vet Comp Oncol*. 2017;15:1553-1557.
28. Kiser PK, Lohr CV, Meritet D, et al. Histologic processing artifacts and inter-pathologist variation in measurement of inked margins of canine mast cell tumors. *J Vet Diagn Invest*. 2018;30:377-385.
29. Northrup NC, Harmon BG, Gieger TL, et al. Variation among pathologists in histologic grading of canine cutaneous mast cell tumors. *J Vet Diagn Invest*. 2005;17:245-248.
30. Northrup NC, Howerth EW, Harmon BG, et al. Variation among pathologists in the histologic grading of canine cutaneous mast cell tumors with uniform use of a single grading reference. *J Vet Diagn Invest*. 2005;17:561-564.
31. Kiupel M, Webster JD, Bailey KL, et al. Proposal of a 2-tier histologic grading system for canine cutaneous mast cell tumors to more accurately predict biological behavior. *Vet Pathol*. 2011;48:147-155.
32. Camus MS, Priest HL, Koehler JW, et al. Cytologic criteria for mast cell tumor grading in dogs with evaluation of clinical outcome. *Vet Pathol*. 2016;53:1117-1123.
33. Mann PC, Hardisty JH. Pathology working groups. *Toxicol Pathol*. 2014;42:283-284.
34. Morton D, Sellers RS, Barale-Thomas E, et al. Recommendations for pathology peer review. *Toxicol Pathol*. 2010;38:1118-1127.
35. Merickel JL, Pluhar GE, Rendahl A, et al. Prognostic histopathologic features of canine glial tumors. *Vet Pathol*. 2021;58:3009858211025795.
36. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol*. 2016;131:803-820.
37. Louis DN, Perry A, Burger P, et al. International Society of Neuropathology-Haarlem Consensus Guidelines for nervous system tumor classification and grading. *Brain Pathol*. 2014;24:429-435.
38. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science*. 2008;321:1807-1812.
39. Brat DJ, Verhaak RG, Aldape KD, et al. Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N Engl J Med*. 2015;372:2481-2498.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Krane GA, Shockley KR, Malarkey DE, et al. Inter-pathologist agreement on diagnosis, classification and grading of canine glioma. *Vet Comp Oncol*. 2022;20(4):881-889. doi:10.1111/vco.12853