Received: 11 March 2023 | Revised: 23 April 2023 | Accepted: 8 May 2023

DOI: 10.1111/vru.13259

ORIGINAL INVESTIGATION

WILEY

Veterinary radiologic error rate as determined by necropsy

Jonathan Cohen[1] | Anthony J. Fischetti[2] | Heather Daverio[3]

[1]Department of Radiology, MedVet Medical and Cancer Centers for Pets, Fairfax, Ohio, USA

[2]Department of Diagnostic Imaging, Schwarzman Animal Medical Center, New York City, New York, USA

[3]Department of Anatomic Pathology, Schwarzman Animal Medical Center, New York City, New York, USA

Correspondence
Dr. Anthony Fischetti, Department of Diagnostic Imaging, Schwarzman Animal Medical Center, 510 East 62nd Street, New York, NY 10065, USA.
Email: anthony.fischetti@amcny.org

Abstract

A large-scale postmortem auditing of antemortem imaging diagnoses has yet to be accomplished in veterinary medicine. For this retrospective, observational, single-center, diagnostic accuracy study, necropsy reports for patients of The Schwarzman Animal Medical Center were collected over a 1-year period. Each necropsy diagnosis was determined to be either correctly diagnosed or discrepant with its corresponding antemortem diagnostic imaging, and discrepancies were categorized. The radiologic error rate was calculated to include only clinically significant missed diagnoses (lesion was not reported but was retrospectively visible on the image) and misinterpretations (lesion was noted but was incorrectly diagnosed). Nonerror discrepancies, such as temporal indeterminacy, microscopic limitations, sensitivity limitations, and study-type limitations were not included in the error rate. A total of 1099 necropsy diagnoses had corresponding antemortem imaging; 440 diagnoses were classified as major diagnoses, of which 176 were discrepant, for a major discrepancy rate of 40%, similar to reports in people. Seventeen major discrepancies were diagnoses that were missed or misinterpreted by the radiologist, for a calculated radiologic error rate of 4.6%, comparable with error rates of 3%–5% reported in people. From 2020 to 2021, nearly half of all clinically significant abnormalities noted at necropsy went undetected by antemortem imaging, though most discrepancies owed to factors other than radiologic error. Identifying common patterns of misdiagnosis and discrepancy will help radiologists refine their analysis of imaging studies to potentially reduce interpretive error.

KEYWORDS
autopsy, diagnostic error, discrepancy, interstitial lung disease, misdiagnosis, reader bias

1 | INTRODUCTION

While modern medical advances have revolutionized our antemortem diagnostic capabilities, postmortem evaluation remains a valuable tool for medical education and hospital quality control.[1] Clinical diagnoses are based on all case factors, including history, clinical signs, physical examination, diagnostics, and response to treatment. Most studies in human medicine[2–4] and veterinary medicine[5,6] evaluate discrepancies between postmortem autopsy diagnoses and antemortem clinical diagnoses. By contrast, only one study in people has evaluated the rate of discrepancy between autopsy and antemortem radiologic diagnoses specifically.[7] A similar study for veterinary diagnostic imaging has not been performed.

Radiologic–pathologic discrepancies can be separated into diagnostic (radiologic) error and nonerror discrepancies that preclude

Abbreviations: ACVP, American College of Veterinary Pathologists; ACVR, American College of Veterinary Radiology; ILD, interstitial lung disease; PTE, pulmonary thromboembolism.

TABLE 1 Descriptive categories.[7]

Descriptive category	Definition of category
Missed diagnosis	Abnormality was not reported on initial imaging review but was identified retrospectively and was identified by necropsy.
Misinterpretation	Abnormality was reported on initial imaging review but the radiologic diagnosis/interpretation was incorrect.
Temporally indeterminate	Abnormality was not present on antemortem imaging but was identified at necropsy (or vice versa), due to its development or resolution during the interval between imaging and the patient's death.
Microscopic limitation	Abnormality was not present on antemortem imaging, and was not reported at gross necropsy, but was visible at the microscopic level.
Sensitivity limitation	Abnormality was not present on antemortem imaging but was reported at gross necropsy. The imaging modality was appropriate, yet the lesion was too small to be detected.
Study-type limitation	Abnormality was not present on antemortem imaging but would more likely have been detected had a different imaging modality been used.
Comorbidity limitation	Abnormality was reported at necropsy but was not present on antemortem imaging because it was obscured by a separate lesion in the same or nearby organ.
Size–volume discrepancy	The subjective size or volume described radiologically differed from that reported at necropsy.
Technical limitation of radiology	Abnormality was not present on antemortem imaging but was reported at necropsy, due to technical factors of the individual study, including image rotation, motion artifact, poor resolution, etc.
Description without diagnosis	Abnormality was reported in the radiologic findings but no radiologic diagnosis was reported, making the significance of the described abnormality uncertain.
Indeterminate discrepancy	Abnormality was identified on antemortem imaging and at necropsy, though interpretation of the abnormality differed, and no definitive diagnosis was reached at necropsy.
Missed diagnosis (necropsy)	Abnormality was definitively present on antemortem imaging but was not reported at necropsy.
Autopsy limitation	Abnormality was reported on antemortem imaging but was not identified at necropsy due to limiting factors associated with necropsy, such as postmortem changes (autolysis, pulmonary congestion, fixation artifact) and exclusion of certain structures on basic necropsy.
Multifactorial	The discrepancy could have been caused by two or more of the above categories (other than missed diagnosis and misinterpretation).

visualization on an imaging study. Some notable nonerror discrepancies included lesion development or resolution in the imaging-death time interval (temporal indeterminacy), lesion obscuration by a concurrent disease (comorbidity limitation), and lesions that are too small to be detected by imaging. Some pathology can only be detected microscopically, regardless of the modality chosen. In contrast to radiologic discrepancy, radiologic error is more relevant to radiologists since it omits the types of errors that are impossible to detect. In people, radiologic error rates range from 3% to 5%, calculated by extrapolating interobserver disagreement[8,9] or autopsy.[7] A large-scale survey to evaluate radiologic error rate has yet to be accomplished in veterinary medicine.

TABLE 2 Criteria for defining major versus minor diagnoses and discrepancies.

Class	Definition	Type	Example
I	Pathology directly related or significantly contributed to death or clinical syndrome prompting euthanasia, which, if detected, may have significantly altered case outcomes.	Major	Pulmonary thromboembolism in a dog presenting in respiratory distress
II	Pathology was a component, or sequela, of the clinical syndrome causing death or prompting euthanasia, but by itself would not have significantly altered case outcome.	Minor	Splenic infarction in a dog that died from systemic complications from correctly diagnosed disseminated neoplasia.
III	Pathology incidentally noted and unrelated to cause of death or clinical syndrome prompting euthanasia.	Minor	Chronic kidney disease in a patient that was euthanized for a disease unrelated to renal dysfunction.

TABLE 3 Imaging modalities.

Combination	Number
CXR	37
CXR, AUS	30
AUS	10
CXR, AXR	8
CXR, AXR, AUS	7
CT	7
CXR, AUS, CT	6
CXR, CT	5
CXR, AUS, US (oth)	4
MRI	3
AXR, AUS	3
AUS, CT	3
CXR, MRI	2
CXR, AUS, MRI	2
CXR, AXR, CT	1
CXR, MXR	1
CXR, CT, MRI	1
CXR, AUS, MXR	1

Abbreviations: AUS, abdominal ultrasound; AXR, abdominal radiographs; CT, computed tomography; CXR, thoracic radiographs; MRI, magnetic resonance imaging; MXR, musculoskeletal radiographs; US (oth), nonabdominal ultrasound.

The primary objective of this study is to quantify the radiologic error rate over a 1-year period in our veterinary teaching hospital. The secondary objectives of this study were to identify the most common sources of clinically relevant missed diagnoses, misinterpretations, and nonerror discrepancies. Our hypotheses were that radiologic error rate would be similar to rates reported in people.

2 | MATERIALS AND METHODS

2.1 | Selection and description of subjects

The study was a retrospective, observational, single-center, diagnostic accuracy study. The use of patient data was approved by the hospital's director of research in accordance with our Institutional Animal Care and Use Committee.

Medical records (to include pathology reports) and imaging studies from The Schwarzman Animal Medical Center between August 2020 and August 2021 were searched to identify patients with the following inclusion criteria: (1) a finalized necropsy report, (2) a corresponding imaging study performed within 1 month of the patient's death, (3) the imaging study included anatomic information described in the necropsy report, and (4) a finalized imaging report was made for the imaging study. Decisions for the inclusion or exclusion of patients were made by a small animal rotating intern veterinarian (J.C.), under the direction of an American College of Veterinary Radiology (ACVR)-certified veterinary radiologist (A.J.F.). All necropsies and necropsy reports were performed by one of two in-house American College of Veterinary Pathologists (ACVP)-certified veterinary pathologists (5–15 years of experience). All imaging studies were performed in-house to include radiography (Quantum HF, Radiographic Imaging System, Ronkonkoma, NY), ultrasound (Aplio i700, Canon Medical Systems, Tustin, CA), computed tomography (Aquilion-64, Canon Medical Systems), and magnetic resonance imaging (Achieva 1.5T, MR Philips, Andover, MA). All imaging reports were finalized by one of

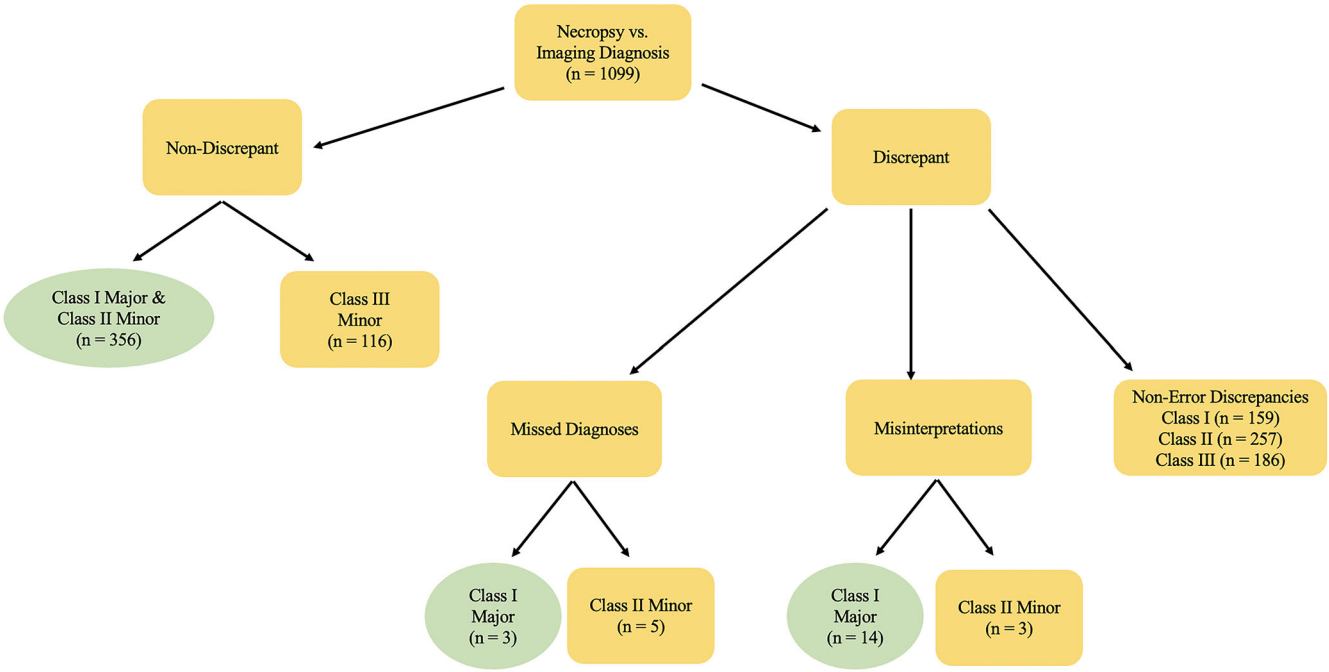

FIGURE 1 Flowchart of radiologic–pathologic diagnoses, discrepancies, and errors by class. The data within ovals are used in the calculation of radiologic error. [Color figure can be viewed at wileyonlinelibrary.com]

three in-house ACVR-certified veterinary radiologists (2–16 years of experience).

2.2 | Data recording and analysis

The electronic imaging database was queried to tabulate patient information. Necropsy diagnoses were categorized by J.C., with consultation from A.J.F. and an ACVP-certified veterinary pathologist (H.D.) as (1) correctly diagnosed by antemortem imaging and (2) discrepant with antemortem imaging findings. All imaging was re-evaluated to identify additional sources of discrepancies. Discrepancies were defined as discordance between autopsy diagnosis and antemortem radiologic diagnosis or absence of a corresponding radiologic diagnosis. One of 13 descriptive categories, defined in Table 1, and similar to Murken's categories,[7] was subsequently assigned to each discrepant diagnosis. Only "missed diagnosis" and "misinterpretation" discrepancies constituted radiologic errors.

For cases with multiple studies of the same modality, studies performed closest to the time of death were used to evaluate for acute conditions. Older studies were used to evaluate more chronic diseases and to assess for temporal change. When multiple imaging modalities evaluated the same anatomic region, pathologic diagnoses detected by one modality but not the other were categorized as correct. Determination of correct versus discrepant was made only once for each diagnosis.

Radiologists' interpretation of findings did not always yield a definitive radiologic diagnosis, making the objective determination of discrepancy sometimes difficult. If the radiologist utilized a broad category of disease to interpret a finding (e.g., hepatopathy, lymphadenopathy, neoplastic etiology), and the corresponding necropsy diagnosis fit under that broad descriptor, then the radiologic diagnosis was nondiscrepant. If the scope of differential diagnoses was narrow, then discrepancy versus nondiscrepancy was determined by the radiologist's prioritization. Radiologic differentials described as "most likely" or "less likely but not excluded" were nondiscrepant if confirmed at necropsy. However, if the same radiologic differential was described as "unlikely", then it was discrepant.

Each diagnosis was classified by J.C. as major (class I) or minor (class II or III), determined by its clinical significance toward the patient's morbidity or mortality (Table 2), using a derivation of the Goldman et al. criteria.[10] Classification decisions were made by J.C., based on the final diagnostic assessment and comments in the pathologist's report, with consultation from H.D. when clinical significance of a diagnosis was ambiguous. We did not parse each patient's clinical record for treatment decisions; therefore, we modified Goldman's class definitions by classifying all diagnoses, deemed at necropsy to have directly related or significantly contributed to the patient's death (or clinical syndrome prompting euthanasia), as class I diagnoses, as we were unable to determine whether antemortem detection would have altered case management. Major discrepancies involved class I diagnoses only. Discrepancies involving an organ where the relevant pathology was part of a widespread disease process, such as disseminated neoplastic disease, were categorized as major if the diseased organ in question resulted in organ-specific clinical signs, or if the widespread disease was not already documented radiologically elsewhere in the body. However, all organs correctly diagnosed radiologically as infiltrated by such a widespread disease process were categorized as major.

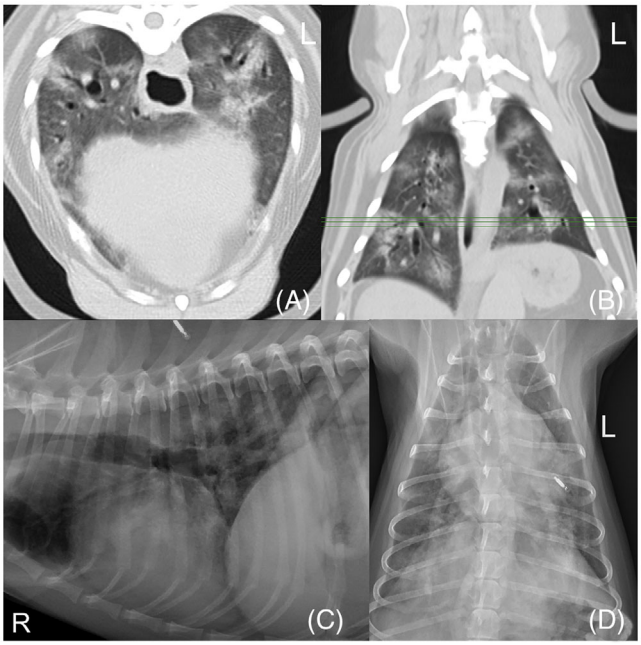

FIGURE 2 Nine-year-old female, spayed Chihuahua, acute on chronic tachypnea; history stated "suspect PTE". The imaging and history can be used as an example of major misinterpretation due to framing bias; the imaging also illustrates the challenges of diagnosing ILD, even with the addition of CT. Transverse (A) and dorsal (B) CT images show multifocal unstructured interstitial to alveolar lung pattern that is somewhat wedge-shaped as it extends to the lung periphery. Right lateral (C) and ventrodorsal (D) radiographs show multifocal lung disease as well as enlarged pulmonary arteries caudodorsally. PTE was prioritized for the combination of findings. Histopathologic diagnosis revealed ILD, to include cryptogenic organizing pneumonia and secondary diffuse alveolar damage. Embolic processes were not identified. Image acquisition parameters for CT: standard reconstruction kernel, WW:1500, WL:−600, 2 mm slice thickness. ILD, interstitial lung disease. [Color figure can be viewed at wileyonlinelibrary.com]

All necropsy diagnoses initially characterized as correct or discrepant radiologically were assessed by H.D., who reviewed all finalized necropsy reports and photographs, when available, and provided an assessment regarding the presence or absence of discrepancy and error. The pathologist was not blinded to the categories of correct versus discrepant. All discrepancies, for which the intern or pathologist suspected radiologic error, were then reviewed by A.J.F. After confirming the discrepancies reflective of true radiologic error, each error was categorized as major or minor by consensus of all authors. Diagnoses were categorized by organ system, and further by diagnosis.

2.3 | Statistics

Contributing perceptual and cognitive biases were determined for each error by two authors (J.C. and A.J.F.). Statistical analysis was performed by one of the authors (A.J.F.) with graduate-level training in statistics using commercially available software (Microsoft Excel, Microsoft Corp, 2018). Discrepancy rates for major diagnoses, and the radiologic error rate, were calculated using the following formulas derived from Murken's study[7]:

Major discrepancy rate = Class I discrepancies/Class I diagnoses,

Radiologic error rate = (MD + MI)/(C + MD + MI),

where MD is major (class I) missed diagnoses, MI is major (class I) misinterpretations, and C is major (class I) and minor (class II) correct diagnoses. We included class II correct diagnoses in this calculation because they remained relevant toward the patient's morbidity/mortality, and an accurate radiologic diagnosis benefitted the overall antemortem understanding of the case. We excluded class III correct diagnoses because they were incidental and unrelated findings, and we excluded minor missed diagnoses and misinterpretations because the case outcome would not have changed had the lesion been correctly diagnosed.

3 | RESULTS

A total of 174 necropsies were performed between August 1, 2020, and August 21, 2021. A total of 131 of 174 cases had at least one diagnostic imaging study within 1 month of death or euthanasia, whereas 43 of 174 did not. A total of 117 of 131 necropsies with antemortem imaging included both gross and histopathological diagnoses, whereas 14 of 131 were limited to gross evaluation, plus cytology in two cases.

The study population included 69 dogs, 47 cats, 10 small mammals (4 rabbits, 3 guinea pigs, and 3 rats), one miniature pig, two birds, and two turtles. The median ages for dogs, cats, and small mammals were 9 years (2 weeks to 15 years), 10 years (4 weeks to 17 years), and 2 years (9 weeks to 10 years), respectively. The two birds, two turtles, and miniature pigs were 10 weeks and 15 years, 7 and 35 years, and 4 years old, respectively.

The 131 cases with antemortem imaging included 105 thoracic radiographic studies, 19 abdominal radiographic studies, two musculoskeletal radiographic studies, 66 abdominal ultrasounds, three cervical ultrasounds, one musculoskeletal ultrasound, 23 CT studies, and eight MRI studies. Most cases utilized multiple radiologic modalities; modality distribution is organized in Table 3.

There were 1356 total necropsy diagnoses of which 1099 had a corresponding antemortem radiologic diagnosis. See Figure 1 for the numeric breakdown of nondiscrepant (correct) and discrepant major (class I) and minor (class II and III) diagnoses. A total of 176 of 440 major diagnoses were discrepant, for a major discrepancy rate of 40.0%.

Of the 1099 total pathologic–radiologic diagnoses, 163 diagnoses (15%) were determined to have most likely developed or resolved after the radiologic study had been performed (temporally indeterminate). The modality distribution for the remaining nontemporally indeterminate discrepancies is listed in Table 4. Thoracic radiography and abdominal ultrasound were the most common modalities utilized in this study. Diagnoses most commonly implicated in radiologic discrepancy are listed in Table 5. The organ system and diagnosis most commonly associated with major discrepancy were the respiratory tract and interstitial lung disease (ILD), respectively.

TABLE 4 Modality distribution for diagnoses and discrepancies involving lesions present at the time of imaging.

Modality	Diagnoses	All discrepancies	Major discrepancies
CXR	335 (35.8%)	182 (39.2%)	70 (53.4%)
AUS	355 (37.9%)	151 (32.5%)	36 (27.5%)
CT	167 (17.8%)	97 (20.9%)	17 (13.0%)
MRI	35 (3.7%)	14 (3.0%)	4 (3.1%)
AXR	33 (3.5%)	19 (4.1%)	4 (3.1%)
MXR	5 (0.5%)	1 (0.2%)	0 (0.0%)
US (oth)	6 (0.6%)	0 (0.0%)	0 (0.0%)
Total	936	464	131

TABLE 5 Common diagnoses with more than one major discrepancy.

Diagnosis	# All diagnoses	# Class I major diagnoses	# Class I major discrepancies
Pulmonary thromboembolism	23	17	15
Pancreatitis—acute	17	15	11
Cardiac inflammation	16	11	10
ILD secondary	25	12	10
ILD primary	20	17	9
Acute kidney injury	17	12	9
Infectious pneumonia	24	23	8
Pleural effusion	41	20	7
Acute chronic kidney disease	11	8	6
Pericardial effusion	29	7	6
Cardiomegaly	58	23	5
Gastrointestinal inflammation	45	15	5
CNS neoplasia	7	7	4
Cardiac neoplasia	15	4	4
Pulmonary neoplasia	17	16	3
Hepatopathy inflammatory	29	13	3
Thrombus—large vessel	7	5	3
Pulmonary hemorrhage	6	5	3
Gastrointestinal diffuse neoplasia	7	4	3
Pancreatitis—chronic	10	3	2
Pulmonary edema	22	15	2
Abdominal effusion	39	9	2
Splenic neoplasia	7	6	2
Vascular neoplasia	5	5	2
Prostatitis	4	4	2
Prostatic neoplasia	4	3	2
Portal vein hypoplasia	3	3	2
Liver shunt	4	2	2
Cardiac fibrosis	3	2	2

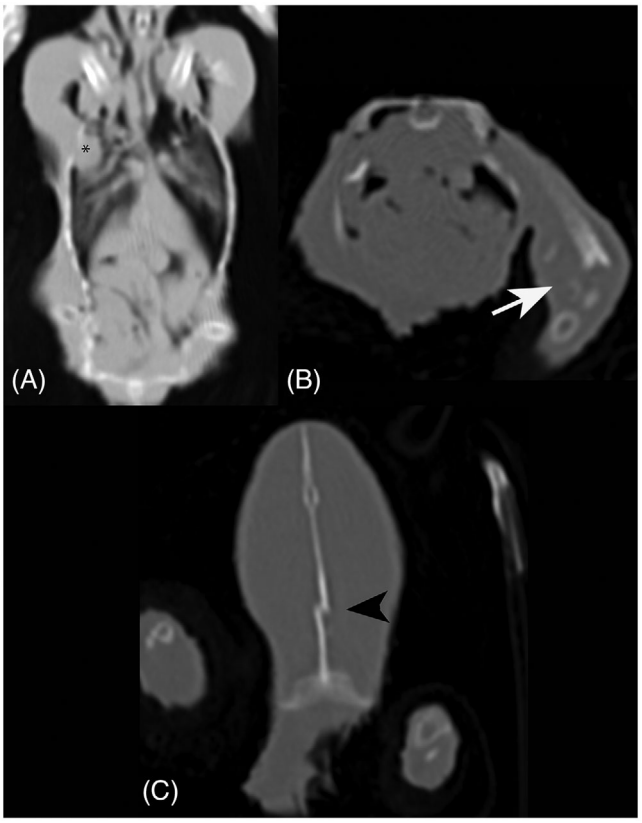

FIGURE 3 One-year-old female Conure bird, recently attacked by two dogs. Right is to the left of the images. Computed tomography examination is an example of missed diagnosis due to satisfaction of search error. (A), A right cranial lung soft tissue opacity, correctly diagnosed as a pulmonary contusion (*). (B), A comminuted left femoral fracture (arrow) that was correctly diagnosed, along with other fractures of the appendicular skeleton. An obvious keel fracture was missed (black arrow in (C)). Image acquisition parameters: standard reconstruction kernel, WW:1500, WL:−600 for lungs; WW:1000, WL:250 for bone; both with 1 mm slice thickness.

Major missed diagnoses and misinterpretations comprised 3 and 14, respectively, of all discrepancies. The number of correct class I and II diagnoses was 264 and 92, respectively (356 total). In this study, the radiologic error rate was (3 + 14)/(356 + 3 + 14) = 4.6%. Representative images and descriptions of each example of error in this study can be found in Supporting Information S1 and Figures 2–7.

4 | DISCUSSION

The common major discrepancies in our study highlight some of the inherent limitations of diagnostic imaging. Pulmonary thromboembolism (PTE) comprised 9% of all major discrepancies, which is comparable to the 7%–24% in human studies.[2–4] Only 2 of 17 cases of PTE (12%) were detected on antemortem imaging, whereas most were only detected microscopically. Most discrepancies involving acute kidney injury and acute pancreatitis were due to the sensitivity limitation of ultrasound in detecting gross pathological features, such as cortical streaking and pinpoint discolored foci, respectively. Study-type limitations, most frequently encountered with cardiac disease missed by radiography (myo/endocarditis, cardiomegaly, pericardial effusion), highlight the importance of choosing diagnostic tests that most accurately detect lesions suspected clinically.

4.1 | Interstitial lung disease

In our study, the organ system most associated with discrepancy was the respiratory tract. An interesting observation was the prevalence of ILD in our study population. Primary ILD (without concurrent pulmonary pathology) comprised 17 of 109 (15.6%) of all major diagnoses involving the lower respiratory tract, nine cases (5.1%) of all major discrepant diagnoses, and two cases (11.8%) of radiologic error. Secondary ILD (which occurred as a result of underlying lung pathology) comprised another 12 of 109 (11%) lower respiratory tract diseases and 10 (5.7%) of all major discrepancies.

ILD represents a heterogeneous group of diseases involving alterations to the distal pulmonary parenchyma, across which gas exchange occurs.[11–14] Although the true pulmonary interstitium is anatomically bordered by the alveolar epithelium and capillary endothelium and is comprised of mesenchymal stromal cells, loose connective tissue, and tissue-specific immune cells, ILDs can also be associated with alterations in the terminal airways and spaces, blood vessels, and/or pleura.[11–14] ILD represents an amplified repair response, where after initial insult, immune cells are recruited and activated, creating a proinflammatory and/or profibrotic environment.[13,14] Fibrosis and altered alveoli structure with loss of functional gas exchange are common manifestations of end-stage disease, regardless of the initiating insult.[11,13,14] Pulmonary hypertension is a known sequel of ILDs.[13–16]

ILDs are responsible for 15%–20% of all lung diseases in people,[11,17] classified into several subtypes encompassing those of known underlying etiologies and idiopathic diseases (e.g., idiopathic pulmonary fibrosis [IPF], miscellaneous diseases), with extensive literature describing their histopathologic and imaging characteristics.[11,12,18–20] CT is more sensitive and specific in diagnosing ILD than radiography.[21] But radiologic diagnosis is often complicated by shared CT features amongst ILDs. In people, a multidisciplinary approach, involving the clinician, radiologist, and pathologist, is essential to reaching an accurate diagnosis, which has crucial treatment and prognostic implications.[12–14]

Comparatively, ILDs are less well-characterized in veterinary medicine, with a smaller body of literature primarily focused on IPF.[13–15,22,23] Other subtypes are described more rarely in the veterinary literature and are reviewed by Reinero.[13,14] Of the 20 primary ILD cases in our study, only one had a CT scan. This particular case involved multiple ILDs and was misinterpreted (Figure 2). ILD was correctly identified radiographically as a differential diagnosis in all cases where pulmonary fibrosis was reported histologically. With fibrotic disease, radiographic changes were nonspecific, typically involving a diffuse or patchy bronchointerstitial to alveolar pattern. Radiography could not reliably differentiate into a specific ILD subtype (e.g., usual interstitial pneumonia, nonspecific interstitial pneumonia,

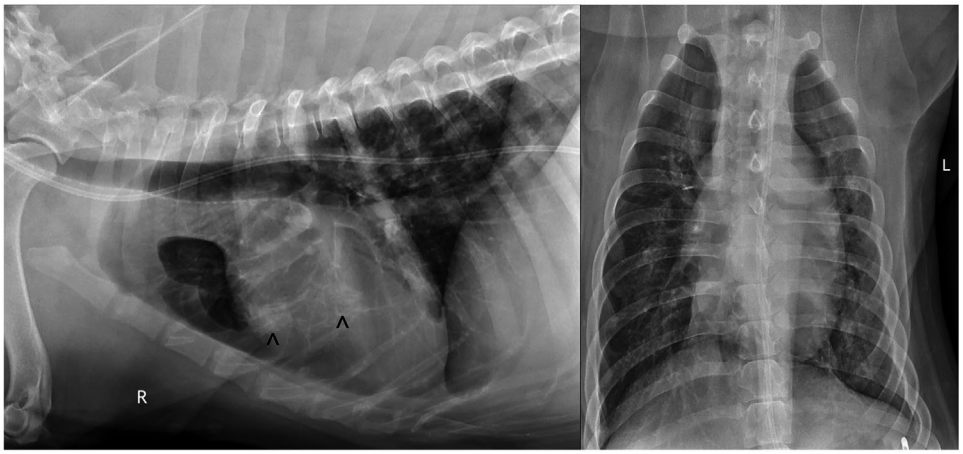

FIGURE 4 Seven-year-old, male neutered, Bernese Mountain Dog; febrile, hypoglycemia, had recent anesthesia for neuter. Right lateral and ventrodorsal thoracic radiographs as an example of major misinterpretation due to framing bias (a history highly suggestive of aspiration pneumonia). Diffuse bronchointerstitial to patchy alveolar lung pattern is most notable ventrally, summating over the heart on the lateral projection (^). A gastroesophageal tube is in place. Histopathological evaluation of the lung showed acute lung injury associated with interstitial lung disease (ILD), microthrombosis, and edema, all secondary to systemic sepsis from a prostatic abscess. Gross and microscopic evidence of aspiration pneumonia was lacking. This example also illustrates the challenges of diagnosing ILD.

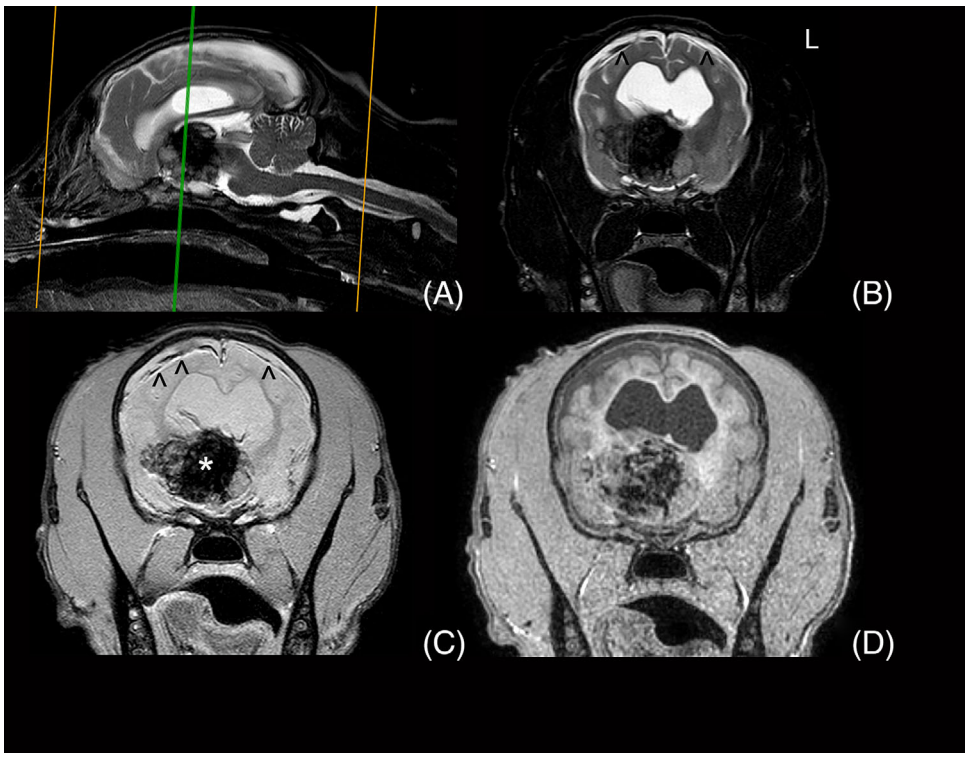

FIGURE 5 Five-month-old female bulldog with right supratentorial neurolocalization and thalamic signs; adipsia, thermoregulatory issues. History also includes possible prior head trauma. The MRI is an example of major misinterpretation due to attribution bias (certain differentials not included because of the patient's signalment) and framing bias (differentials prioritized because of historical information). Sagittal T2-weighted (A), transverse T2-weighted (B), gradient-echo (C), and postcontrast, fat-saturated T1-weighed (D) images show a very large right thalamic mass (*) with areas of signal void and mass effect causing obstructive hydrocephalus. There are also extensive subdural T2-hyperintensity and linear signal voids (^). Magnetic resonance imaging diagnosis concludes multifocal intra-axial and extra-axial hemorrhage due to coagulopathy, systemic inflammation/infection, or trauma. Histopathological diagnosis was pilocytic astrocytoma with associated extensive hemorrhage. Image acquisition parameters: slice thickness 3 mm; slice gap 3 mm; TR 2600 ms, TE 120 ms for T2-weighted images; TR 4000 TE 25, flip angle 15 for gradient echo; TR 700 TE 20 for T1-weighted image. [Color figure can be viewed at wileyonlinelibrary.com]

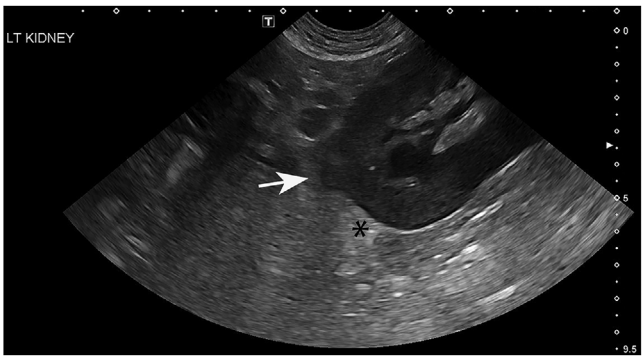

FIGURE 6 Ten-year-old female, spayed Labrador retriever with progressive azotemia, elevated liver enzymes, and leukocytosis. The abdominal ultrasound is an example of a major misinterpretation due to premature closure. The image is centered on the cranial pole of the right kidney, depicting an irregular focus of what was interpreted as echogenic subcapsular fluid (white arrow) and retroperitonitis (*) as a sign of acute nephropathy. Differentials included infectious, primary inflammatory, and toxin-induced nephritis. Additionally, mild chronic, nonspecific architectural changes were noted in the liver, and a few lymph nodes were mildly enlarged. Necropsy diagnosed high-grade lymphoma in the kidneys, liver, spleen, nodes, and gastrointestinal tract.

radiation-induced fibrosis, or organizing pneumonia). Based on autopsy findings, primary ILD occurred in our study population with a similar frequency to reports in humans. Less frequent antemortem ILD diagnoses may be attributed to infrequent utilization of CT and lung biopsy in our dyspneic veterinary patients.

4.2 | Etiology of error

Of particular interest in this study were the factors that caused radiologists to miss or misinterpret lesions that were present in a study. Radiologic errors can be broadly classified as errors of either perception (not identifying a lesion that is present in an imaging study after retrospective review) or cognition (identifying a lesion but not providing a correct interpretation). Supporting Information S1 lists the etiology for each case of an error in our study.

All three perceptual errors (missed diagnoses) in our study were the result of satisfaction of search error, including failure to identify the keel fracture in a polytraumatized bird (Figure 3). Satisfaction of search occurs when lesions are overlooked after the radiologist identifies one or more other abnormalities.[24,25] Detection rate of additional abnormalities, after one is identified, decreased to 40% in one study.[26] Interventions to reduce the satisfaction of search error include a consistent, systematic approach to interpretation with report templates designed as checklists. The checklists can remind the radiologist to make concerted efforts to identify certain "do not miss" diagnoses that are common to the body region and modality.[25]

A wider variety of biases comprised the cognitive errors in our study. Framing bias occurs when interpretation of an abnormality is biased by information reported in the case history,[25,27] which contributed to four errors in our study. Figure 4 is an example—"fever, recent anesthesia" reported in the history contributed to the misdiagnosis of aspiration pneumonia in a dog with ILD. Performing an initial analysis of a radiologic study prior to reading the patient's clinical history may help reduce errors from framing bias. Of course, this recommendation must be weighed against the knowledge that radiological interpretation is more accurate when using history to help formulate the diagnosis.[29] Seeking a more thorough clinical history from the electronic medical record or the ordering clinician may also help reduce framing bias as imaging requests are often incomplete.[25]

Attribution bias occurs when certain differentials for a noted abnormality are not included because of the patient's signalment.[24,25,27] An example of this bias in our study was the 5-month-old dog with a brain tumor (Figure 5). An initial review of an imaging study without signalment could help stem an attribution bias. In a more general sense, premature closure bias (a broader form of attribution bias) occurs when conclusions are reached and accepted without considering and/or reporting plausible alternative differentials,[28] such as the exclusion of lymphoma as a sonographic differential for bilateral renomegaly and retroperitonitis in a dog with progressive azotemia (Figure 4). Structured report templates that include uncommon manifestations of disease could help reduce premature closure.[25] While making these efforts to provide more differential diagnoses, radiologists must still be effective in communicating their primary concerns, without regurgitating a meaningless list of possibilities.

Alliterative bias (i.e., satisfaction of report bias) contributed to three errors in our study and occurs when the interpretation of an abnormality is biased by the judgments made in a prior report by another radiologist.[24,25,27] Formulating an initial impression from a study before reviewing prior studies may help reduce alliterative bias.

Benign interpretation bias, not previously described, contributed to three errors in our study. We define this as the failure to consider significant pathology (a form of premature closure) in organs that appear abnormal from benign etiologies. Prostatic neoplasia in intact (or recently neutered) male dogs was twice misattributed to benign prostatic change alone (Figure 7). Benign interpretation bias, a cause of radiologic error, was distinguished from nonerror comorbidity limitation discrepancies, such as cardiogenic edema obscuring interstitial pneumonia, postoperative peritonitis obscuring acute pancreatitis, and pleural effusion collapsing the lungs and obscuring bronchopneumonia. In these cases, the overlapping disease processes were not necessarily benign and contributed to the patient's decline. In the dogs with prostatic neoplasia, prior reports describing benign prostatic hyperplasia or prostatitis may have also contributed to an alliterative bias.

Categorizing cognitive and perceptual errors can help identify trends in a radiologist's thought processes and training techniques. However, to truly minimize errors, radiologists must also consider systematic and cultural causes. Systematic causes of errors include high workload, short staffing, suboptimal protocols, and computer-related failures. Interestingly, this study evaluated cases during the peak of the COVID-19 epidemic, a time when systematic causes of error were potentially higher. Cultural sources of error include frequent interrup-

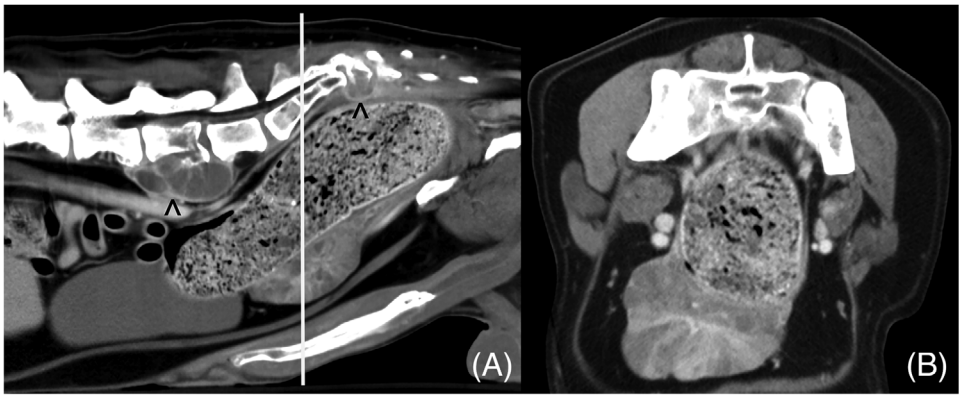

FIGURE 7 Eight-year-old male Briard dog with progressive pelvic limb lameness. This study is an example of benign interpretation bias. Serial ultrasound exams over 3 years showed static prostatic enlargement, heterogenous internal architecture, and retention cysts, suspect as benign for an intact dog. After 3 years of this assumption, abdominal CT (lateral reformatted (A) and transverse (B) images) showed polyostotic aggressive bone lesions and associated rim-enhancing soft tissue masses of L6 and S1 (^) and heterogenous contrast enhancement of an otherwise normally sized and shaped prostate. Histopathological diagnosis at necropsy revealed prostatic carcinoma with disseminated bony metastasis. Prostatic neoplasia was not given as a differential because of the false assumption that benign prostatic hyperplasia and retention cysts were responsible for the appearance of the prostate over the years. Image acquisition parameters for CT: standard reconstruction kernel, WW:350, WL50, 3 mm slice thickness.

tions that can break a train of thought. In our experience, a culture of reluctance to share/discuss mistakes with our peers prevents radiologists from avoiding recurring pitfalls. Peer learning conferences (e.g. Tumor Boards, Morbidity/Mortality Rounds, Radiology-Pathology Correlation Rounds) arranged in a nonpunitive environment, focusing on the root cause of error rather than who made the error, are time-tested strategies for reducing error and educating all clinicians involved.[25]

Because this was a retrospective study, we relied on reported gross and histologic descriptions, as well as archived photographs and slides from routinely processed formalin-fixed, paraffin-embedded tissues, to compare with radiologic images. This made precise identification and categorization of discrepancies more difficult than if direct cadaver re-examination were possible. One complication of retrospectively analyzing radiologic error is the tendency to exaggerate the predictability of a diagnosis after such a diagnosis is confirmed, known as hindsight bias.[25,27] We had access to the complete case history and postmortem findings for every patient. This eliminates the factor of diagnostic uncertainty present during prospective imaging interpretation. Thus, it is important to recognize the educational purpose of evaluating errors, which is to help identify the etiology of such errors, discuss potential biases, and reduce future occurrences. Lastly, utilizing autopsy as the "gold standard" to audit antemortem diagnostic performance has limitations. Nonbiased case selection, standardized necropsy protocols, use of sensitivity and specificity as statistical indices, and accounting for pathologist uncertainty and error maximize the utility of autopsy for this purpose.[30] Many patients sent to necropsy at our hospital have relatively uncomplicated antemortem diagnoses. However, our necropsy rate is not 100%, and there is inevitably some selection bias toward challenging cases, which has been shown to result in higher discrepancy/error rates.[31] Pathologist error, estimated using interpathologist disagreement, could not be accounted for because independent case evaluation by multiple pathologists is not performed at our institution. However, gross photos and histopathology slides were available for review in cases where there was radiologist–pathologist disagreement, allowing for some degree of retrospective review.

In conclusion, findings from the current study indicate that many lesions identified at necropsy may go undetected by antemortem imaging and that veterinary radiologists are not immune from perceptual and cognitive error. This is the first large-scale veterinary study utilizing necropsy to determine the veterinary radiologic error rate, which, at our institution, was 4.6%. This is comparable to the 3%–5% range reported in human radiology literature.[7,8,9] The results from this study underscore the value of routine necropsy in the clinical audit of diagnostic imaging.

5LIST OF AUTHOR CONTRIBUTIONS

Category 1

(a) Conception and design: Cohen, Fischetti, Daverio.
(b) Acquisition of data: Cohen, Fischetti, Daverio.
(c) Analysis and interpretation of data: Cohen, Fischetti, Daverio.

Category 2

(a) Drafting the article: Cohen.
(b) Revising article for intellectual content: Cohen, Fischetti, Daverio.

Category 3

(a) Final approval of the completed article: Cohen, Fischetti, Daverio.

Category 4

(a) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the study are appropriately investigated and resolved: Cohen, Fischetti, Daverio.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PREVIOUS PRESENTATION OR PUBLICATION DISCLOSURE

The results of this study were presented as an oral presentation at the ACVR annual meeting in Reno, NV, on October 20, 2022.

REPORTING CHECKLIST DISCLOSURE

No EQUATOR or other reporting guidelines were followed for the preparation of this manuscript.

ORCID

Jonathan Cohen https://orcid.org/0009-0009-9095-533X

REFERENCES

1. Burton JL, Underwood J. Clinical, educational, and epidemiological value of autopsy. *Lancet*. 2007; 369(9571):1471-1480. doi:10.1016/S0140-6736(07)60376-6
2. Pastores SM, Dulu A, Voigt L, Raoof N, Alicea M, Halpern NA. Premortem clinical diagnoses and postmortem autopsy findings: discrepancies in critically ill cancer patients. *Crit Care*. 2007; 11(2):R48. doi:10.1186/cc5782
3. Maris C, Martin B, Creteur J, et al. Comparison of clinical and postmortem findings in intensive care unit patients. *Virchows Arch*. 2007; 450(3):329-333. doi:10.1007/s00428-006-0364-5
4. Marshall HS, Milikowski C. Comparison of clinical diagnoses and autopsy findings: six-year retrospective study. *Arch Pathol Lab Med*. 2017; 141(9):1262-1266. doi:10.5858/arpa.2016-0488-OA
5. Dank G, Segev G, Moshe D, Kent MS. Follow-up study comparing necropsy rates and discrepancies between clinical and pathologic diagnoses at a veterinary teaching hospital: 2009 versus 1989 and 1999. *J Small Anim Pract*. 2012; 53(12):679-683. doi:10.1111/j.1748-5827.2012.01296.x
6. Schertenleib TI, Pospischil A, Hässig M, Kircher PR, Hilbe M. Comparison of clinical and pathological diagnoses in cats and dogs. *J Comp Pathol*. 2017; 156(2-3):217-234. doi:10.1016/j.jcpa.2017.01.004
7. Murken DR, Ding M, Branstetter BF 4th, Nichols L. Autopsy as a quality control measure for radiology, and vice versa. *AJR Am J Roentgenol*. 2012; 199(2):394-401. doi:10.2214/AJR.11.8386
8. Siegle RL, Baram EM, Reuter SR, Clarke EA, Lancaster JL, McMahan CA. Rates of disagreement in imaging interpretation in a group of community hospitals. *Acad Radiol*. 1998; 5(3):148-154. doi:10.1016/s1076-6332(98)80277-8
9. Berlin L. Accuracy of diagnostic procedures: has it improved over the past five decades? *AJR Am J Roentgenol*. 2007; 188(5):1173-1178. doi:10.2214/AJR.06.1270
10. Goldman L, Sayson R, Robbins S, Cohn LH, Bettmann M, Weisberg M. The value of the autopsy in three medical eras. *N Engl J Med*. 1983; 308(17):1000-1005. doi:10.1056/NEJM198304283081704
11. Crystal RG, Gadek JE, Ferrans VJ, Fulmer JD, Line BR, Hunninghake GW. Interstitial lung disease: current concepts of pathogenesis, staging and therapy. *Am J Med*. 1981; 70(3):542-568. doi:10.1016/0002-9343(81)90577-5
12. Travis WD, Costabel U, Hansell DM, et al. An official American Thoracic Society/European Respiratory Society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med*. 2013; 188(6):733-748. doi:10.1164/rccm.201308-1483ST
13. Reinero C. Interstitial lung diseases in dogs and cats part I: the idiopathic interstitial pneumonias. *Vet J*. 2019; 243:48-54. doi:10.1016/j.tvjl.2018.11.010
14. Reinero C. Interstitial lung diseases in dogs and cats part II: known cause and other discrete forms. *Vet J*. 2019; 243:55-64. doi:10.1016/j.tvjl.2018.11.011
15. Evola MG, Edmondson EF, Reichle JK, Biller DS, Mitchell CW. Valdés-Martínez A. Radiographic and histopathologic characteristics of pulmonary fibrosis in nine cats. *Vet Radiol Ultrasound*. 2014; 55(2):133-140. doi:10.1111/vru.12106
16. Behr J, Ryu JH. Pulmonary hypertension in interstitial lung disease. *Eur Respir J*. 2008; 31(6):1357-1367. doi:10.1183/09031936.00171307
17. Meyer KC. Pulmonary fibrosis, part I: epidemiology, pathogenesis, and diagnosis. *Expert Rev Respir Med*. 2017; 11(5):343-359. doi:10.1080/17476348.2017.1312346
18. Mueller-Mang C, Grosse C, Schmid K, Stiebellehner L, Bankier AA. What every radiologist should know about idiopathic interstitial pneumonias. *Radiographics*. 2007; 27(3):595-615. doi:10.1148/rg.273065130
19. Lynch DA, Travis WD, Müller NL, et al. Idiopathic interstitial pneumonias: cT features. *Radiology*. 2005; 236(1):10-21. doi:10.1148/radiol.2361031674
20. Godwin JD, Müller NL, Takasugi JE. Pulmonary alveolar proteinosis: cT findings. *Radiology*. 1988; 169(3):609-613. doi:10.1148/radiology.169.3.3186983
21. Mathieson JR, Mayo JR, Staples CA, Müller NL. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography. *Radiology*. 1989; 171(1):111-116. doi:10.1148/radiology.171.1.2928513
22. Johnson VS, Corcoran BM, Wotton PR, Schwarz T, Sullivan M. Thoracic high-resolution computed tomographic findings in dogs with canine idiopathic pulmonary fibrosis. *J Small Anim Pract*. 2005; 46(8):381-388. doi:10.1111/j.1748-5827.2005.tb00334.x
23. Thierry F, Handel I, Hammond G, King LG, Corcoran BM, Schwarz T. Further characterization of computed tomographic and clinical features for staging and prognosis of idiopathic pulmonary fibrosis in West Highland white terriers. *Vet Radiol Ultrasound*. 2017; 58(4):381-388. doi:10.1111/vru.12491
24. Bruno MA, Walker EA, Abujudeh HH. Understanding and Confronting Our Mistakes: the Epidemiology of Error in Radiology and Strategies for Error Reduction. *Radiographics*. 2015; 35(6):1668-1676. doi:10.1148/rg.2015150023
25. Itri JN, Tappouni RR, McEachern RO, Pesch AJ, Patel SH. Fundamentals of Diagnostic Error in Imaging. *Radiographics*. 2018; 38(6):1845-1865. doi:10.1148/rg.2018180021
26. Ashman CJ, Yu JS, Wolfman D. Satisfaction of search in osteoradiology. *AJR Am J Roentgenol*. 2000; 175(2):541-544. doi:10.2214/ajr.175.2.1750541
27. Onder O, Yarasir Y, Azizova A, Durhan G, Onur MR, Ariyurek OM. Errors, discrepancies and underlying bias in radiology with case examples: a pictorial review. *Insights Imaging*. 2021; 12(1):51. doi:10.1186/s13244-021-00986-8
28. Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging*. 2017; 8(1):171-182. doi:10.1007/s13244-016-0534-1
29. Lamb CR. Applying the concept of major and minor findings: guidance for trainees and exam candidates. *Vet Radiol Ultrasound*. 2022; 63(5):649-652. doi:10.1111/vru.13146

30. Saracci R. Is necropsy a valid monitor of clinical diagnosis performance? *BMJ*. 1991; 303(6807):898-900. doi:10.1136/bmj.303.6807.898
31. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. *JAMA*. 2003; 289(21):2849-2856. doi:10.1001/jama.289.21.2849

How to cite this article: Cohen J, Fischetti AJ, Daverio H. Veterinary radiologic error rate as determined by necropsy. *Vet Radiol Ultrasound*. 2023;64:573–584. https://doi.org/10.1111/vru.13259

SUPPORTING INFORMATION

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