Receiver operating characteristic curve analysis of the performance of various radiographic protocols when screening dogs for pulmonary metastases

Paul Y. Barthez, Dr Med Vet; William J. Hornof, DVM, MS;
Alain P. Théon, Dr Med Vet, MS; Thérèse J. Craychee, DVM;
Joe P. Morgan, DVM, Vet med dr

Summary: *Five radiographic protocols for detecting pulmonary metastases in dogs were compared by analyzing receiver operating characteristic curves for the protocols. Protocols compared were a right lateral view only, a left lateral view only, right lateral and dorsoventral views, both lateral views, and all 3 views. Three radiologists used each of the protocols to evaluate 99 sets of thoracic radiographs. Fifty-two sets of radiographs were from dogs confirmed histologically to have pulmonary metastases and 47 were from dogs proven at necropsy to be free of pulmonary metastases. Results of the 5 protocols were not statistically different. We concluded that a third view is not necessary when routinely screening dogs with cancer for pulmonary metastases and that the standard 2-view thoracic examination should be adequate. However, in individual cases, a third view may be the determining factor in establishing a radiographic diagnosis and should be obtained if any suspicious areas are seen.*

Thoracic radiology is currently used in small animal oncology to detect pulmonary metastases.[1] However, when a lateral view of the thorax is obtained, the dependent lung is compressed by gravitational force and is less inflated than the nondependent lung.[1-5] A lesion in the dependent lung can be masked, especially if it is in an area superimposed on the cardiac silhouette.[6] As a result, it has been recommended that a 3-view protocol, including right lateral, left lateral, and ventrodorsal views, be used routinely when screening dogs and cats for pulmonary metastases. Yet, this recommendation was made on the basis of individual clinical case observations[2,4,6] and sensitivity studies.[3]

Sensitivity and specificity are terms used to describe the discriminating power of a diagnostic test.[7] Sensitivity represents the true-positive fraction, or likelihood of a positive test result for an animal that truly has the disease in question; whereas, specificity represents the true-negative fraction, or likelihood of a negative test result for an animal that truly does not have the disease. A relation between sensitivity and specificity that is independent of the imaging process exists and can be expressed by a receiver operating characteristic (ROC) curve.[8] Diagnostic tests have areas under the ROC curve ranging between 0.5 and 1. If the test is perfect, the area under the curve is 1. If the test is not useful in discriminating healthy from diseased animals, the area under the curve is 0.5.[8] The area under the ROC curve and its SE can be used to evaluate the performance of different imaging modalities, protocols, or techniques.[8-13]

The purpose of the study reported here was to compare, by means of ROC curve analysis, 5 radiographic protocols for screening dogs suspected of having pulmonary metastases.

Materials and Methods

Case selection—All necropsy reports and reports of histologic examination of lung biopsy specimens from dogs that underwent a 3-view radiographic study of the thorax for any reason, over a 3-year period, were reviewed. No effort was made to exclude reports from dogs with underlying pulmonary diseases, but reports from dogs with lymphoma were excluded. This review allowed us to establish 2 groups of dogs: dogs confirmed, on the basis of necropsy or examination of biopsy specimens, to have pulmonary metastasis (true positive) and those determined to be free of metastasis at necropsy (true negative). All sets of thoracic radiographs for these dogs were collected, and dogs were further subdivided, by examining reports of radiographic examinations, into dogs with radiographically visible metastasis and those without radiographically visible metastasis. If results of the radiographic study were negative for metastasis but metastasis was seen at necropsy, radiographs for that dog were included only if necropsy was per-

From the Veterinary Medical Teaching Hospital (Barthez, Craychee) and the Department of Radiological Sciences (Hornof, Théon, Morgan), School of Veterinary Medicine, University of California, Davis, CA 95616.

Address reprint requests to Dr. Hornof.

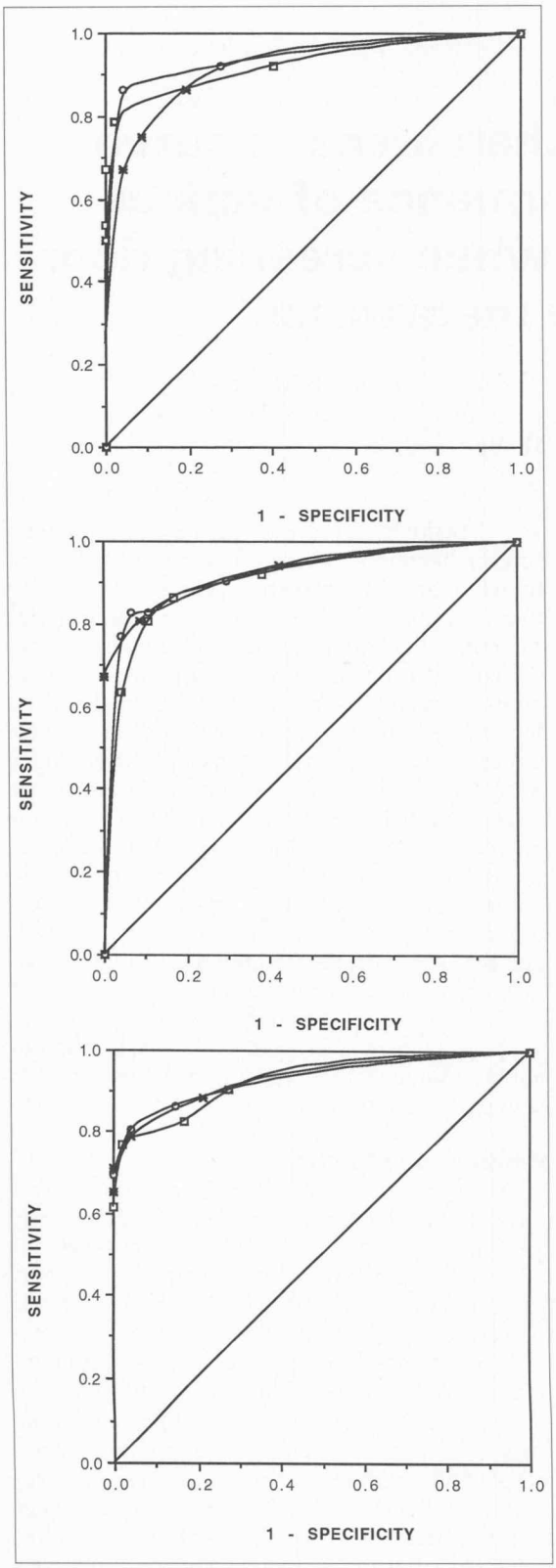

Figure 1—Receiver operating characteristic curves for radiologist 1 (top), radiologist 2 (middle), and radiologist 3 (bottom) evaluating thoracic radiographs for evidence of pulmonary metastases. For simplicity, only 3 of the 5 protocols used are shown. ○ = right lateral, left lateral, and dorsoventral (or ventrodorsal) views were examined; □ = right lateral and dorsoventral (or ventrodorsal) views were examined; and x = only the right lateral view was examined.

formed within 1 week of the radiographic study. In instances in which results of the radiographic study and the necropsy or histologic examination were both positive for metastasis and multiple sets of radiographs where available, the earliest radiographic study interpreted as positive was used.

This selection process yielded 99 independent sets of thoracic radiographs. Each set included a right lateral (RL), left lateral (LL), and ventrodorsal or dorsoventral (DV) view. Of the 99 sets of radiographs, 52 were from dogs known to have pulmonary metastases and 47 were from dogs that did not have any evidence of pulmonary metastases.

Method—Five radiographic protocols were compared: (1) RL view only, (2) LL view only, (3) RL and DV views, (4) RL and LL views, and (5) all 3 views (RL, LL, DV). Three veterinary radiologists, who had not been involved in selection of radiographs for the study, were the readers. The readers independently evaluated the 99 sets of radiographs during 5 sessions, using the 5 protocols in random order. Radiographs were evaluated using 1 protocol in each session. At each session, readers were given the 99 sets of radiographs in random order. Readers were told that radiographs were from dogs with a history of primary neoplasia; however, readers were not told the type of tumor or its location. Readers were asked to evaluate the radiographs for pulmonary metastases and to give each set of radiographs a score from 1 to 5 (Appendix). For each reader, at least 7 days separated one session from the next.

Data analysis—Data from the 5 evaluations were analyzed to produce ROC curves for each reader. Sensitivity and specificity were calculated by dichotomizing the 5-point rating scale at each of the possible cutoff points,[9] and ROC curves were constructed by ploting sensitivity vs (1 − specificity) and including the origin (0,0) and the point of unity (1,1). The area under the curve, A(z), and SE were calculated by use of nonparametric statistics (ie, Wilcoxon test).[13] Mean A(z) ± SE was calculated for each protocol as described.[14] Stepwise comparison of the protocols was performed, using a two-tailed *t*-test.[13] Because ROC curves were derived from the same radiographs, we were able to take into account the correlation between the areas under the curve when comparing the protocols.[15] Correlation coefficients between the areas under the curve were determined by calculating Kendall τ values as described elsewhere.[12,14,15]

Results

The overall shape of the ROC curves was similar for all readers and all protocols (Fig 1). Mean A(z) was not significantly different ($P > 0.1$) among the 5 protocols (Table 1). Although mean A(z) for the protocol in which only the LL view was examined was lower than that for the other protocols, the difference was not significant ($P > 0.1$).

Table 1—Area under the receiver operating characteristic curve, A(z), and SE for 3 radiologists using 5 protocols to screen 99 sets of thoracic radiographs for evidence of pulmonary metastases

Protocols	Radiologist 1		Radiologist 2		Radiologist 3		Mean	
	A(z)	SE	A(z)	SE	A(z)	SE	A(z)	SE
RL view	0.895	0.03	0.923	0.03	0.916	0.03	0.911	0.03
LL view	0.866	0.04	0.845	0.04	0.936	0.03	0.882	0.04
RL and LL views	0.915	0.03	0.902	0.03	0.920	0.03	0.912	0.03
RL and DV views	0.916	0.03	0.896	0.03	0.914	0.03	0.909	0.03
RL, LL, and DV views	0.936	0.03	0.903	0.03	0.922	0.03	0.920	0.03

RL = right lateral, LL = left lateral, and DV = dorsoventral or ventrodorsal.

Discussion

In this study, none of the radiographic protocols was found to be better than the others when used to screen for pulmonary metastases in dogs. This contradicts the results of previous studies[3,16]; however, we believe that our results are more trustworthy because we used a balanced, adequate number of positive and negative cases, we only selected cases in which a histologic diagnosis had been made, and we used ROC curves to analyze the data.

The number of cases that could be included in this study was limited by the ability of the readers to concentrate on the study for long periods. Therefore, the possibility of a type-II error (ie, one protocol is really better for detecting pulmonary metastases, but the difference between this protocol and the other protocols is not large enough to be detected with the number of cases used) should still be considered.[13] Because the SE for A(z) varies inversely with the square root of the number of cases,[13,14] we would have had to use a much larger number of cases to detect such small differences between protocols. From a practical point of view, such small differences, even if they were found to exist, would have little or no clinical significance

In this study, some radiographs were clearly more difficult to correctly interpret than were others. The sample was representative of radiographs obtained from dogs that would typically be screened for pulmonary metastases and was not composed of radiographs from dogs with end-stage cancer and obvious pulmonary metastases. Radiographs from dogs without pulmonary metastases were included so that specificity could be determined. Surprisingly, some radiographs from dogs without pulmonary metastases were given high scores, and some were even given a score of 5 (definitely has pulmonary metastases).

The standard of reference is an important issue when designing studies comparing imaging techniques.[9] The true status of each individual included in the study must be independently defined.[14] In this study, histologic examination was used as the standard because it is the most reliable method for diagnosing pulmonary metastases.

The ROC curves show how sensitivity varies inversely with specificity. Analysis of ROC curves has been used in the field of radiology for the past 20 years and is considered to be the best method for comparing different imaging techniques.[8,9,14,17,19] The advantage of this method is that it minimizes the influence of underreading and overreading when evaluating different protocols or techniques. Curves can be fitted to the data by use of parametric (binormal assumption)[20] or nonparametric methods.[13] In this study, a nonparametric method was used because this method is well-accepted and convenient and because it does not require that any assumptions be made concerning the underlying distribution of the test results.[8] With this method, the areas under the curve may have been underestimated; however, this is acceptable because we were concerned only with making comparison between protocols.[13]

The lack of difference between the results of the standard 2-view (RL and DV views) and the 3-view protocol indicates that the need for 3 views when screening for pulmonary metastases in dogs with cancer may have been overemphasized and the additional cost of routinely obtaining all 3 views can probably not be justified. However, in particular instances, a third view may be useful in determining the radiographic diagnosis and should be obtained if any suspicious areas are seen. Although in one study[3] there was a difference between the LL and RL views, we did not find a statistical difference between the protocols in which one or the other lateral view was the only view examined. We also did not find a significant difference when comparing 1- and 2-view protocols. Therefore, a single lateral view of the thorax may be adequate for detection of pulmonary metastases in dogs. However, we recommend that 2 orthogonal views be obtained so that other potential sites for metastases (eg, ribs, mediastinum) can be assessed and the thorax can be evaluated to determine whether there is any increased risk for complications if anesthesia is needed for diagnostic or therapeutic procedures.

Appendix

Rating scale used for evaluating thoracic radiographs for evidence of pulmonary metastases

Score	Interpretation
1	Definitely no evidence of pulmonary metastases
2	Probably no pulmonary metastases
3	Possible pulmonary metastases
4	Probable pulmonary metastases
5	Definitely positive for pulmonary metastases

References

1. Miles KG, Lattimer JC, Jergens AE, et al. A retrospective evaluation of the radiographic evidence of pulmonary metastatic disease on initial presentation in the dog. *Vet Radiol* 1990;31(2):79–82.

2. Biller DS, Myer CW. Case examples demonstrating the clinical utility of obtaining both right and left lateral thoracic radiographs in small animals. *J Am Anim Hosp Assoc* 1987; 23:381–386.

3. Lang J, Wortman JA, Glickman LT, et al. Sensitivity of radiographic detection of lung metastases in the dog. *Vet Radiol* 1986;27(3):74–78.

4. Spencer CP, Ackerman N, Burt JK. The canine lateral thoracic radiograph. *Vet Radiol* 1981;22(6):262–266.

5. Ahlberg NE, Hoppe F, Kelter U, et al. A computed tomographic study of volume and x-ray attenuation of the lungs of Beagles in various body positions. *Vet Radiol* 1985;26(2):43–47.

6. Pechman RD. Effect of dependency versus nondependency on lung lesion visualization. *Vet Radiol* 1987;28(5):185–190.

7. Phillips WC, Scott JA. Medical decision making: practical points for practicing radiologists. *Am J Radiol* 1990; 154:1149–1155.

8. Centor RM. Signal detectability: the use of ROC curves and their analyses. *Med Decis Making* 1991;11:102–106.

9. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. *Invest Radiol* 1989;24:234–245.

10. Koblik PD, Hornof W, Harnagel SH, et al. A comparison of pulmonary angiography, digital subtraction angiography, and ^{99m}Tc-DTPA/MAA ventilation-perfusion scintigraphy for detection of experimental pulmonary emboli in the dog. *Vet Radiol* 1989;30(4):159–168.

11. Wortman JA, Chang-Winterkorn P, Knight DH, et al. Receiver operating characteristic curve analysis of the utility of thoracic radiographs for the diagnosis of feline cardiac disease. *Vet Radiol* 1987;28(4):121–126.

12. McNeil BJ, Hanley JA. Statistical approaches to the analysis of receiver operating characteristic curves. *Med Decis Making* 1984;4:137–150.

13. Hanley JA, McNeil J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.

14. Swets JA, Pickett RM. *Evaluation of diagnostic systems.* New York: Academic Press Inc, 1982.

15. Hanley JA, McNeil J. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839–843.

16. Holt D, Van Winkle T, Schelling C, et al. Correlation between thoracic radiographs and postmortem findings in dogs with hemangiosarcoma: 77 cases (1984-1989). *J Am Vet Med Assoc* 1992;200:1535–1539.

17. Brismar J. Understanding receiver-operating-characteristic curves: a graphic approach. *Am J Roentgenol* 1991; 157:1119–1121.

18. Rummeny EJ, Wernecke K, Saini S, et al. Comparison between high-field strength MR imaging and CT for screening of hepatic metastases: a receiver operating characteristic analysis. *Radiology* 1992;182:879–886.

19. Kelsey CA, Moseley RD, Mettler FA, et al. Cost-effectiveness of stereoscopic radiographs in detection of lung nodules. *Radiology* 1982;142:611–613.

20. Hanley JA. The robustness of the binormal assumptions used in fitting ROC curves. *Med Decis Making* 1988;8: 197–203.

Altered platelet indices in dogs with hypothyroidism and cats with hyperthyroidism

Changes in platelet indices (platelet count and platelet size) and PCV associated with thyroid disease were studied in 7 dogs with hypothyroidism and 21 cats with hyperthyroidism that were admitted to the veterinary teaching hospital. Compared with control (euthyroid) dogs, dogs with hypothyroidism had higher platelet count ($P = 0.003$), smaller platelet size ($P = 0.01$), and lower PCV ($P = 0.02$). Comparison of the group of hyperthyroid cats with a group of similarly aged, clinically normal cats with normal thyroxine values indicated that the group of hyperthyroid cats had significantly ($P = 0.03$) higher mean platelet size than did control cats, but differences were not found in mean platelet count or PCV. Results of this investigation indicate that the changes in platelet size reported in human beings with thyroid endocrinopathies also are found in animals so affected.—*P. Sullivan, R. Gompf, L. Schmeitzel, et al in Am J Vet Res 54 (December 1993).*