

Veterinary Pathology Online

<http://vet.sagepub.com/>

Recommended Guidelines for the Conduct and Evaluation of Prognostic Studies in Veterinary Oncology

J. D. Webster, M. M. Dennis, N. Dervisis, J. Heller, N. J. Bacon, P. J. Bergman, D. Bienzle, G. Cassali, M. Castagnaro, J. Cullen, D. G. Esplin, L. Peña, M. H. Goldschmidt, K. A. Hahn, C. J. Henry, E. Hellmén, D. Kamstock, J. Kirpensteijn, B. E. Kitchell, R. L. Amorim, S. D. Lenz, T. P. Lipscomb, M. McEntee, L. D. McGill, C. A. McKnight, P. M. McManus, A. S. Moore, P. F. Moore, S. D. Moroff, H. Nakayama, N. C. Northrup, G. Sarli, T. Scase, K. Sorenmo, F. Y. Schulman, A. M. Shoieb, R. C. Smedley, W. L. Spangler, E. Teske, D. H. Thamm, V. E. Valli, W. Vernau, H. von Euler, S. J. Withrow, S. E. Weisbrode, J. Yager and M. Kiupel

Vet Pathol 2011 48: 7 originally published online 27 July 2010

DOI: 10.1177/0300985810377187

The online version of this article can be found at:

<http://vet.sagepub.com/content/48/1/7>

Published by:



<http://www.sagepublications.com>

On behalf of:

[American College of Veterinary Pathologists](#), [European College of Veterinary Pathologists](#), & the [Japanese College of Veterinary Pathologists](#).

Additional services and information for *Veterinary Pathology Online* can be found at:

Email Alerts: <http://vet.sagepub.com/cgi/alerts>

Subscriptions: <http://vet.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Jan 25, 2011

[OnlineFirst Version of Record](#) - Jul 27, 2010

[What is This?](#)

Recommended Guidelines for the Conduct and Evaluation of Prognostic Studies in Veterinary Oncology

Veterinary Pathology
48(1) 7-18
© The American College of
Veterinary Pathologists 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0300985810377187
http://vet.sagepub.com



J. D. Webster¹, M. M. Dennis², N. Dervisis³, J. Heller⁴, N. J. Bacon⁵, P. J. Bergman⁶, D. Bienziele⁷, G. Cassali⁸, M. Castagnaro⁹, J. Cullen¹⁰, D. G. Esplin¹¹, L. Peña¹², M. H. Goldschmidt¹³, K. A. Hahn¹⁴, C. J. Henry¹⁵, E. Hellmén¹⁶, D. Kamstock¹⁷, J. Kirpensteijn¹⁸, B. E. Kitchell³, R. L. Amorim¹⁹, S. D. Lenz²⁰, T. P. Lipscomb²¹, M. McEntee²², L. D. McGill¹¹, C. A. McKnight²³, P. M. McManus²⁴, A. S. Moore²⁵, P. F. Moore²⁶, S. D. Moroff²⁷, H. Nakayama²⁸, N. C. Northrup²⁹, G. Sarli³⁰, T. Scase³¹, K. Sorenmo³², F. Y. Schulman²¹, A. M. Shoieb³³, R. C. Smedley³⁴, W. L. Spangler³⁵, E. Teske¹⁸, D. H. Thamm³⁶, V. E. Valli³⁷, W. Vernau²⁶, H. von Euler³⁸, S. J. Withrow³⁶, S. E. Weisbrode³⁹, J. Yager⁴⁰, and M. Kiupel^{34,41}

¹ Molecular Pathology Unit, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

² Faculty of Veterinary Science, University of Sydney, Camden, New South Wales, Australia

³ Department of Small Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan

⁴ School of Animal and Veterinary Sciences, Charles Sturt University, Wagga Wagga, New South Wales, Australia

⁵ Department of Small Animal Clinical Sciences, College of Veterinary Medicine, University of Florida, Gainesville, Florida

⁶ BrightHeart Veterinary Centers, Armonk, New York

⁷ Department of Pathobiology, University of Guelph, Guelph, Ontario, Canada

⁸ Department of General Pathology, Institute of Biological Science, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil

⁹ Department of Public Health, Comparative Pathology, and Veterinary Hygiene, University of Padua, Padua, Italy

¹⁰ Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, North Carolina

¹¹ ARUP Laboratories, Salt Lake City, Utah

¹² Department of Animal Medicine, Surgery and Pathology, Veterinary School, University Complutense of Madrid, Madrid, Spain

¹³ Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

¹⁴ Hill's Pet Nutrition, Inc, Topeka, Kansas

¹⁵ Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri, Columbia, Missouri

¹⁶ Department of Anatomy, Physiology, and Biochemistry, Swedish University of Agricultural Sciences, Uppsala, Sweden

¹⁷ Department of Microbiology, Immunology, and Pathology, College of Veterinary Medicine, Colorado State University, Fort Collins, Colorado

¹⁸ Department of Companion Animal Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, Netherlands

¹⁹ Department of Veterinary Clinics, Veterinary Medical School, UNESP, Sao Paulo State University, Botucatu, Sao Paulo, Brazil

²⁰ Department of Comparative Pathobiology, School of Veterinary Medicine, Purdue University, West Lafayette, Indiana

²¹ Department of Veterinary Pathology, Armed Forces Institute of Pathology, Washington, DC

²² Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York

²³ MPI Research, Mattawan, Michigan

²⁴ Idexx Reference Laboratories, Portland, Oregon

²⁵ Faculty of Veterinary Science, University of Sydney, Sydney, New South Wales, Australia

²⁶ Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, University of California, Davis, California

²⁷ Antech Diagnostics, Lake Success, New York

²⁸ Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo, Japan

²⁹ Department of Small Animal Medicine and Surgery, College of Veterinary Medicine, University of Georgia, Athens, Georgia

³⁰ Department of Veterinary Public Health and Animal Pathology, Faculty of Veterinary Medicine, University of Bologna, Bologna, Italy

³¹ Bridge Pathology Ltd, Bristol, United Kingdom

³² Ryan Veterinary Hospital, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

³³ Pfizer, Drug Safety and Research Development, Sandwich, United Kingdom

³⁴ Diagnostic Center for Population and Animal Health, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan

³⁵ NSG Pathology, Winters, California

³⁶ Animal Cancer Center, Colorado State University, Fort Collins, Colorado

³⁷ VDX Veterinary Diagnostics, Davis, California

³⁸ Department of Clinical Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

³⁹ Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio

⁴⁰ Yager-Best Histovet, Guelph, Ontario, Canada

⁴¹ Department of Pathobiology and Diagnostic Investigation, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan

Corresponding Author:

Dr M. Kiupel, Diagnostic Center for Population and Animal Health, Michigan State University, 4125 Beaumont Road 152A, Lansing, MI 48910

Email: kiupel@dcpah.msu.edu

Abstract

There is an increasing need for more accurate prognostic and predictive markers in veterinary oncology because of an increasing number of treatment options, the increased financial costs associated with treatment, and the emotional stress experienced by owners in association with the disease and its treatment. Numerous studies have evaluated potential prognostic and predictive markers for veterinary neoplastic diseases, but there are no established guidelines or standards for the conduct and reporting of prognostic studies in veterinary medicine. This lack of standardization has made the evaluation and comparison of studies difficult. Most important, translating these results to clinical applications is problematic. To address this issue, the American College of Veterinary Pathologists' Oncology Committee organized an initiative to establish guidelines for the conduct and reporting of prognostic studies in veterinary oncology. The goal of this initiative is to increase the quality and standardization of veterinary prognostic studies to facilitate independent evaluation, validation, comparison, and implementation of study results. This article represents a consensus statement on the conduct and reporting of prognostic studies in veterinary oncology from veterinary pathologists and oncologists from around the world. These guidelines should be considered a recommendation based on the current state of knowledge in the field, and they will need to be continually reevaluated and revised as the field of veterinary oncology continues to progress. As mentioned, these guidelines were developed through an initiative of the American College of Veterinary Pathologists' Oncology Committee, and they have been reviewed and endorsed by the World Small Animal Veterinary Association.

Keywords

cancer, guidelines, prognostication, standardization, veterinary oncology

Diagnostic markers are clinical, molecular, or pathologic characteristics of a patient or disease that are associated with a specific disease but not necessarily associated with a specific clinical outcome or response to treatment. In contrast, prognostic markers are clinical, molecular, or pathologic characteristics of a patient or disease that are associated with a clinical outcome, whereas predictive markers are characteristics associated with a treatment outcome.²⁹ Prognostic and predictive markers are used to identify the likely progression of a patient's disease and to determine treatment modalities that are most appropriate and efficacious for that patient. With an increasing number of treatment options available to veterinary oncology patients, the financial costs associated with these treatments, the potential for side effects, and the emotional stress associated with the disease and treatment experienced by owners, there is an increasing need for more accurate prognostic and predictive markers in veterinary oncology.

With an enormous amount of data being generated in the biomedical research community, there is marked variation in study design, assay performance, and result reporting. To address these variations and to allow meta-analyses, many interest groups have established minimum information requirements for publishing data based on specific assays (mRNA expression profiling,⁵ quantitative polymerase chain reaction (PCR),⁶ immunohistochemistry and in situ hybridization¹²) or technical disciplines (proteomics⁵⁰ and genomics^{15,49}). The human medical community has developed similar standards for prognostic biomarker reporting²⁹ and for the conduct of clinical trials.^{2,31} Many of these minimum information requirements and standards are living documents and will likely be continually amended to address new concerns and technologies.^{5,12,31,50} However, they suggest a baseline that should allow for increased transparency, more accurate and

repeatable research results, independent evaluation and validation of studies, and interstudy comparison.^{5,6,12,49}

Since January 2000, *Veterinary Pathology* published approximately 28 studies evaluating prognostic markers for neoplastic diseases in animals. The *Journal of Veterinary Internal Medicine* and *Veterinary and Comparative Oncology* published approximately 56 studies evaluating prognostic or predictive markers. These studies varied in their design, sample size, and results. The majority of studies in *Veterinary Pathology* focused on immunohistochemical markers associated with prognosis, whereas many studies published in the other 2 journals identified clinical markers associated with response and survival following administration of a tested chemotherapeutic or radiotherapy regimen. As veterinary medicine—specifically, veterinary oncology—continues to advance, there will be an increasing need to identify novel prognostic markers associated with metastasis, recurrence, disease-free interval, and overall survival. In addition, as new rationally targeted treatments, immunotherapeutics, and chemotherapeutics and radiotherapeutic protocols are developed and applied to veterinary patients, there will be a need to find predictive markers that will identify patients' likelihood of responding to a given therapeutic protocol.

To address these needs, the veterinary oncology research community needs to develop and perform rigorous controlled experiments to identify and characterize prognostic and predictive markers. Reporting of studies will need to be thoughtful and thorough so that the significance of the results can be independently evaluated, verified, reproduced, and objectively compared to those of other published studies. To accomplish these goals, the veterinary community should establish standards for study design and reporting. These standards should serve as benchmarks but not necessarily as inflexible

requirements, because there are many ways to design and conduct an experiment²⁹ and every study cannot meet all requirements. This article outlines the components of a prognostic study and highlights the points that should be evaluated to determine the quality and significance of a given study. The goal is not to suggest requirements for publication. Instead, the goal is to encourage thorough reporting and move toward standards for study design. These standards should facilitate critical review of current and future literature and extrapolation of conclusions to applied clinical settings.

Prognostic Study Objectives

An essential step in designing any study is to clearly define its aim. Investigators should clearly state the question being addressed and the study objectives. Study design and research methods stem from the research question; therefore, the study objectives should be preplanned and characterized in the introduction.

The goal of prognostic studies is to identify markers that are predictive of disease outcome, often by defining risk groups based on prognosis. Prognostic studies may focus on one or more markers to define risk groups, or they may generate a model that predicts outcomes, considering a range of markers in tandem. Screening numerous prognostic markers results in an increased probability of spurious associations. Hypotheses should be predefined to reduce the potential for such misleading observations. The prognostic markers being evaluated, the study's clinical endpoints, and intergroup comparisons should be included in a study's hypothesis. In addition, prognostic marker selection should be justified on the basis of biological plausibility, findings of previous studies, relevance to understanding disease pathogenesis or treatment, and/or clinical experience.⁴¹

Prognostic Study Design

Selection of the study population can be prospective or retrospective. A prospective approach is always preferred and is imperative when a marker cannot be measured in a stored sample. Ideally, studies evaluating prognostic markers should be prospective cohort studies. In these studies, patients with a defined neoplasm of interest are identified and followed until an outcome of interest is reached (eg, recurrence, metastasis, death).²⁰ Cohorts are defined within the study population by their exposure status, defined by the patient's status of the prognostic marker of interest¹³ (eg, present versus absent, degree or amount present). Differences in outcomes (eg, disease-free survival and overall survival times) are statistically compared between cohorts, usually by measuring relative risks, odds ratios, or survival curves.

Retrospective studies are necessary when diseases or outcomes are rare and when there is a long period from the onset of disease until its outcome. In retrospective cohort studies, historical cases of defined neoplasms are collected and placed into cohorts based on their prognostic marker status; then, the

differences in patient outcomes are statistically compared between cohorts. Disadvantages of retrospective studies include incomplete or inaccurate patient data, decreased standardization in treatment and data/specimen collection, and study populations biased toward patients with available tumor specimens. Therefore, results of retrospective studies are subject to a variety of unknown biases.

Regardless of study type, a comparison or reference group consisting of patients with the defined neoplasm but with a different prognostic marker status (present versus absent, different levels) needs to be included for biological and statistical comparisons. Case series describe the incidence of an outcome of interest in a collection of patients, without simultaneously evaluating the same outcome in a reference population. Few conclusions can be drawn from studies lacking a reference group.

Animals in all cohorts must be obtained from the same identifiable population and with the same neoplastic disease and treatment. Oftentimes, the animals within each cohort are not chosen randomly. Risk factors associated with prognosis, other than the prognostic marker being studied, may not be evenly distributed among each group. Therefore, controlling cohorts for equal representation of a known variable between groups to adjust for bias (ie, matching) might be necessary. In such a case, matching methods should be detailed.¹³

Sample Size

Inadequate sample size reduces the capability to identify potentially important associations. Some rationale or justification for a prognostic study's sample size ensures that its methods are appropriate for the study objectives. In survival analyses, the number of events (eg, recurrences, metastases, neoplasia-associated deaths), rather than the number of recruited patients, is used to define the required sample size. Therefore, requirements for sample size estimation include the length of the follow-up period, the presumed prognosis of the population, the suspected magnitude of contribution of each factor being evaluated, and the correlations with other factors.

The follow-up length is based on the time needed to collect an adequate number of events, and it can be estimated only when the likely outcome of the population is known. If a neoplasm of interest is consistently associated with short disease-free or survival intervals, a short follow-up time will be sufficient. However, if the tumor exhibits variable or prolonged disease-free or survival intervals, the follow-up time will likely need to be extended. If a tumor is associated with only rare events, then a larger sample size will be needed to ensure that enough events occur within the sample population.

In addition, because the sample size needed will vary, depending on the study's objectives and statistical methods, no single formula can be suggested for determining sample size. For example, if a study is to consider multiple prognostic markers or interactions between multiple markers, a larger sample size will be needed to confidently detect differences in outcomes. As a general rule for multivariable models, the

number of events should be at least 10 times the number of potential prognostic variables included in the model.^{23,45}

Prognostic studies should also consider and optimally describe the power needed to detect expected differences (ie, the probability that a statistically significant difference will be found if it exists). If a sample size is too small and its power is limited, a lack of statistically significant differences will not be meaningful. Many studies are limited by the number of available samples, which is an inherent problem for rare diseases. Multi-institutional databases may allow access to greater sample numbers; however, results from these samples may be biased by a variety of factors, including inconsistencies in treatment, sample collection, and record keeping.

Study Population

Study populations should comprise patients who represent the larger diseased-patient population and who are (preferably) all at the same point in disease progression. Dissimilarity among patients in terms of natural disease history may result in failure to detect true differences in prognostic markers. Additionally, if most animals in a study have end-stage disease at the time of diagnosis, it could be difficult to detect differences in outcomes among cohorts.

A lucid study provides inclusion and exclusion criteria to define how the study population was selected. Exclusion criteria often consider treatment or other factors that might reduce or confound outcome events, as well as the types of biological samples and medical data necessary to evaluate the prognostic marker status and patient outcome, respectively. Inclusion criteria often include a case definition for the study population and the period during which patients were enrolled. A case definition should be clearly stated, specifying the characteristics of the neoplasm of interest. All members of the study will share these characteristics. Case definitions are primarily defined by histologic criteria but can include a variety of other criteria, including tumor location, tumor subtype (eg, histomorphologic features, age, or expression of certain diagnostic markers), histologic tumor grade, tumor margins, clinical stage of disease, and treatment. Any study using histologic criteria as part of a case definition should state the tumor sampling strategy (ie, manner in which biopsy was collected) and the methods/criteria used to categorize tumor subtypes or grades. For example, if only dogs with a certain lymphoma classification subtype are enrolled in a study, then the specific criteria used to characterize the subtype of lymphoma should be clearly presented. Ideally, histologic criteria should be repeatable with minimal intraobserver and interobserver variability. If histologic criteria are highly subjective, then authors should acknowledge this as part of the study definition, with clearly stated measures of agreement among pathologists and with methods to overcome this subjectivity. For the purpose of data analysis, tumor subclassification based on histologic features should be included in the study only when there is reason to believe that subtypes have unique biological behaviors. However, it is appropriate to describe the histologic spectrum of the tumor and acknowledge

divergent morphologies as part of the case definition to document that different histologic subtypes have the same biological behavior. If a tumor has divergent histomorphologic features that confound classification, methods should be provided to address the criteria by which these tumors are classified (ie, by predominant features or features more likely to impact prognosis). Additionally, grading schemes often have poor repeatability^{34,35} and should be used only as part of a case definition if their biological significance has been validated. A study using immunohistochemical detection of a diagnostic marker as part of the case definition should include the methods employed for reproducible conduct of the assay and a clear description of the evaluation criteria (ie, location and level of immunolabeling and percentage of immunolabeled cells that constitute a positive result). Studies that include tumor margin assessment in the case definition should describe the methods used for margin evaluation.

If the results of a study are to be clinically applicable, then the study population should represent the larger population to which results will be extrapolated,²⁶ and it should be described in detail. Distributions of critical patient characteristics, such as species and age, should be summarized in the results. Other host factors that may affect prognosis, such as sex, breed, and sample source (primary care or referral hospital), should be included on the basis of previous literature results.

The study population should be as free from bias as possible. The percentage of the study population that constitutes referrals versus primary care patients may represent a serious bias. For example, patients at referral hospitals may be more likely to be in advanced stages of disease, to have already received undisclosed treatments, to have a variant of disease that is more difficult to manage, to be confined to a geographic location, and/or to have been seen by referring practitioners with varying levels of expertise. Other sources of bias may be found in retrospective studies, wherein specimen availability may be associated with factors such as tumor size, patient outcome, attending clinician, and preservation method.²⁹

Bias cannot be entirely avoided, and potential confounders should be accounted for within multivariable statistical analyses. Furthermore, because the population available to a study is not random, there will inevitably be biases. This limitation must be addressed with a thorough description of the study population.

Predictive Markers and Controlling for Treatment Effects

Ideally, studies evaluating predictive markers are cohort studies, which are similar to clinical trials in some regards. The elected treatment is often but not always mechanistically related to the function of the predictive marker of interest. In any study that involves adjunct therapies, treatment protocols should be described in detail in the Materials and Methods section. In exploratory studies, the treatment protocol is usually stringently fixed. In advanced studies, where more information is known regarding efficacy, flexible protocols may be devised. Treatments are allocated to subgroups within each study

cohort, and the association with the outcome of interest is measured and compared between groups. To prevent bias, investigators are blinded to the treatment being administered, which may require administration of a placebo to control groups. Comparisons can be made between new treatment groups relative to placebo groups (negative control), conventional treatment groups (positive control), or both. Most important, the outcomes in association with the new treatment should be compared with a predictive marker–negative cohort that is receiving the same treatment. Comparisons regarding predictive markers' associations with survival among patients that received different treatment protocols are inappropriate. In these cases, the presence of 2 independent variables (the predictive marker and the treatment protocol) will confound significant interpretation of the results. It is also inappropriate to combine animals given different treatment protocols into single cohorts to increase cohort sizes. Formal randomization is the ideal method for allocating patients into a treatment group. Clinical judgment or owner preference should not be used to assign patients to groups, given that it introduces severe bias.

Assessment of Clinical Outcomes

To identify associations between prognostic markers and clinical outcomes, it is paramount to have a clear definition of the endpoint, or event, under consideration. Measured events often include time to tumor progression or disease-free interval and overall survival time as continuous variables, whereas disease progression and mortality are dichotomous variables (occurred, did not occur). Regardless of the event selected, endpoints must be rigorously defined. Afterward, it is important to adhere to these endpoints and the defined start of the observation period. Histopathology—or, at minimum, cytology—is the gold standard for diagnosing local recurrence and metastatic lesions. Lymph nodes represent common sites of metastatic disease for many solid tumors. Additionally, lymph node metastases, as determined by cytologic or histologic examination, are an important prognostic indicator in many solid tumors. For example, the presence or absence of lymph node metastases is closely linked to the survival of dogs with lung cancer and osteosarcoma.^{24,39} Ideally, lymph node status should be determined by histologic or, minimally, cytologic evaluation. However, this is not always possible, as in the case of internal node involvement. In human medicine, sentinel lymph node examination is the ideal method to properly evaluate tributary lymph nodes; unfortunately, this methodology has not been routinely established in veterinary medicine.⁴⁸

When cytologic or histologic evaluations are not feasible, diagnostic imaging can be used to identify metastatic or recurrent disease. Currently, the size of a lymph node is the main imaging finding suggestive of metastatic lymph node involvement. Larger lymph nodes are more likely than smaller nodes to have metastasis.⁴⁴ Unfortunately, lymph node enlargement is a nonspecific finding because many other conditions besides metastasis, such as inflammation and hyperplasia, can cause lymphadenomegaly. Furthermore, it is well recognized that

micrometastases may be present in small nodes. Therefore, the accuracy of diagnostic imaging for detecting disease progression is questionable.⁵⁸ However, this is a limitation for any type of biomarker- or imaging-based analysis. Rather than try to define de novo guidelines to assess clinical outcomes and disease progression based on pathology or imaging techniques, we propose close adherence to the standards established by RECIST, PERSIST, and Choi.^{7,14,54}

Statistical Analysis

The use of statistical analysis is essential for validating prognostic markers. By definition, all prognostic studies rely on an estimate of survival, or time to the occurrence of a defined endpoint. Numerous considerations are required to ensure that analysis and reporting of survival data are undertaken appropriately. A nonexhaustive list of criteria needed for accurate reporting is detailed here. Ultimately, consultation with a statistician is strongly recommended before the initiation of any study.

Censoring

Censoring refers to the fact that a number of animals under evaluation will not have experienced the defined outcome event during the study period. Although initial cursory exploration of potential associations between putative prognostic markers and outcome may be undertaken without its consideration, censoring is essential for clinically useful and accurate analysis of survival data. Censoring is necessary if an animal has not experienced the relevant event (such as death or relapse) by the end of the study period, is lost to follow-up, or has experienced a different event that makes further follow-up impossible, such as death from an unrelated illness.⁹ Animals experiencing these situations are referred to as *right censored* and must be accounted for by using survival analysis methods. *Left-censored data* refers to observations where the event of interest occurs before the onset of observation. *Interval censoring* refers to observations made in cases where the event occurs at an unknown time between 2 observation periods. Most survival data reported may be appropriately analyzed with the consideration of right-censored data only. Censoring should be uninformative and unrelated to the event of interest to enable appropriate statistical inference from the methods applied. For example, it would be inappropriate to censor euthanized animals because of a paraneoplastic syndrome, given that euthanasia is related to the cancer of interest. All events contributing to censoring should be defined, and the number or percentage of animals censored should be indicated—particularly, those lost to follow-up. In some cases, it might not be possible to determine if death was due to malignancy or a concurrent condition, especially if necropsy data are not available. In these situations, it is especially important to define the criteria used to censor data and the criteria used to determine if death was due to the cancer of interest. Results should be interpreted with caution when a large proportion of

the study population has been lost to follow-up or when censoring is not considered to be uninformative.

Descriptive Indices

Valuable information can be obtained from descriptions of variable distributions, survival times, and calculations of summary statistics. Graphic representation of survival over time may be displayed with univariable Kaplan–Meier survival plots. This graphic display represents the proportion of patients surviving over time and allows for calculations of median survival time for the overall population or for subgroups. Graphs that include indications of censoring should be displayed, and estimates of median survival should be reported with confidence intervals and *P* values. Even if a magnitude of difference in median survival exists between 2 groups, this may not be statistically significant, because of intragroup variations, as evidenced by overlapping confidence intervals. It is also important to consider the potential influences of all variables when interpreting these graphs. If continuous and categorical variables are present in the multivariable model, separate plots may be needed to examine the effect of the continuous variable at multiple levels of the categorical variable. For example, if age and sex are variables in the model, it might be necessary to examine the effect of age separately in males and females.

Modeling Survival Data

Multivariate refers to a model where there is more than 1 outcome variable, whereas *multivariable* refers to a model where there is more than 1 explanatory variable.³⁸ Multivariable analysis allows the model to be adjusted for alternative putative prognostic variables as well as patient-related covariates. Because multiple variables can influence survival and confound associations between the prognostic marker of interest and survival, it is important to account for these variables by using multivariable analysis. All variables that may be associated with prognosis, including prognostic or predictive markers of interest, should be accounted for in multivariable statistical analyses. Each variable and how it was measured should be described in the Materials and Methods section. Studies should not neglect consideration of clinical or gross pathologic features, which can be easily measured and could clearly have an association with prognosis, such as extent of tumor necrosis, grossly evident invasion, metastasis, or tumor size. The description of each variable should include whether the data are continuous, dichotomous, or categorical. Putative alternative predictors and potential confounders should be relevant to the overall study aim and identified at the outset of the study to enable accurate sample size calculation. Inclusion of continuous or categorical covariates should be data driven. Consideration of relationships between the covariates and the outcomes should be made before their inclusion as continuous variables. Similarly, if a continuous variable is to be categorized, then the use of more than 2 categories should be considered to minimize information loss, and categorical

definitions should be based on data distribution and clinical relevance.¹⁰ Categorization based on statistical significance of varying cut points is inappropriate because of the introduction of bias.⁴⁵

Survival data are usually modeled with the hazard function rather than the survival function at a given time. The survival function describes the probability of surviving from the start of the study until at least time *t*, whereas the hazard function describes the risk of the event in a small interval after time *t*, assuming that the subject has survived until that time. The use of regression models (eg, Cox's proportional hazards model) or accelerated time function models allows for the inclusion of multiple prognostic variables.^{3,4} Model building should be based on clinical as well as statistical considerations, and whereas stepwise techniques are appropriate, fully automated implementation should be avoided. Model assumptions (eg, proportional hazards) should be carefully considered, and all models should be assessed for goodness of fit.

Model Reporting

There are many considerations for appropriate reporting of prognostic studies. Many of these issues have been outlined in previously published recommendations, such as CONSORT (Consolidated Standard for Reporting Trials),³¹ STROBE (Strengthening the Reporting of Observational Studies in Epidemiology),⁵² and REMARK (Reporting Recommendations for Tumour Marker Prognostic Studies).²⁹ In short, methods of statistical analysis should be thoroughly detailed, including information relating to variable selection and categorization, model-building strategies, assumptions and how they were verified, and the predetermined approaches to managing missing data. Results of univariable and multivariable analyses should be reported along with model-building criteria. It is critical to report confidence intervals for each variable's effect size in the final model. This allows for an estimation of reliability and a consideration of the effect of sample size, and it ultimately helps to determine whether results are truly statistically significant. McShane et al²⁹ also recommended reporting estimated effects of prognostic variables under consideration from a model that includes all standard prognostic variables, regardless of statistical significance.

It is important to remember that statistical significance does not equal biological significance or clinical value. Similarly, lack of statistical significance does not definitively indicate that the marker has no prognostic value.⁴⁵ Interpretation of results should be undertaken with clinical and statistical considerations in mind and with specific reference to the hypotheses addressed within the study.

Methods of Prognostic Marker Evaluation

The Materials and Methods section of any article should include enough information for the reader to independently reproduce the experiments. Therefore, reporting should be thorough and inclusive. In veterinary medicine, a variety of

assays and techniques have been used in attempts to identify prognostic markers for neoplastic diseases, including clinical parameters,^{1,18,24} histologic features,^{27,42,46} immunohistochemistry,^{33,56} quantitative PCR,¹⁶ evaluations for genomic mutations,⁵⁷ and specific serum protein levels or enzyme activity analyses.^{17,25} The variety of assays used to identify prognostic markers will likely increase as new technologies are developed. For some assays, such as quantitative PCR, proteomic, and mRNA expression–profiling experiments, suggested minimum information requirements for reporting have been published,^{5,6,12,15,50} and the veterinary oncology and pathology communities should strive to meet or exceed these requirements. Guidelines have recently been reported for the minimum information specification for *in situ* hybridization and immunohistochemistry experiments,¹² as well as suggested guidelines for the use of immunohistochemistry in veterinary diagnostics.⁴⁰ These resources should serve as baseline references and benchmarks for the use and reporting of immunohistochemistry, which is the most common technique used in recent veterinary prognostic studies. There is significant overlap among the minimum information requirements of many assays, whereas some requirements are unique to individual assays.⁴⁹ It is beyond the scope of this effort to review the individual reporting requirements of each assay that might have utility in prognostic biomarker discovery. However, some basic principles for reporting most, if not all, assays should be applied in prognostic studies.

The most important principle in the Materials and Methods section is that data should be reported in sufficient detail to allow the reader to validate the results and potentially apply the methods to a clinical setting. This begins with descriptions of the sample population; the population source and selection; the tissue subsection selection; and the sample collection, handling, preparation, and processing.⁵³ Details about how the marker was assessed should then be reported, including assay design, methodology, validation, and control. Assay design and methodology include all aspects of assay development—from sample handling and preparation to final data analysis and interpretation. Guidelines for many clinical, pathologic, and molecular measurements that could have prognostic utility have yet to be developed; however, detailed information regarding these key items can be included for almost every type of prognostic assay. These details allow editors, reviewers, and readers to critically evaluate the experiment, determine its application, and assess its validity. Additionally, sources of all materials used in the assay should be reported.

Assay validation should include demonstrations of the assay's sensitivity, specificity, and reliability, such as its positive and negative predictive values. For immunohistochemistry and other antibody-based assays, antibodies should be validated to confirm that they are detecting the protein of interest in the species examined. This is especially important in veterinary medicine because few antibodies are made specifically for companion animal species.⁴⁰ To demonstrate the antibody's specificity, Western blotting should be performed to document detection of a single protein at the predicted molecular weight

in the tissue of interest.⁴⁰ Additionally, immunohistochemical labeling should be shown to be restricted to the predicted tissue, cell type, and subcellular location (eg, nucleus, cytoplasm, cytoplasmic membrane) in positive-control tissues.⁴⁰ It is preferable that the control tissues be normal, given that aberrant protein expression can occur in neoplastic diseases. DNA- or RNA-based assays should describe not only the primer and/or probe sequence but also the genomic sequence from which probes were derived and the predicted splice variants and single nucleotide polymorphisms that they are expected to detect.⁶

In addition to demonstrating that the assay detects the predictive target (specificity), authors need to describe the quantitative ability of the assay (sensitivity). In many cases, such as quantitative PCR and immunoblots, quantification is based on an established relationship to reference transcripts or proteins. Not all reference transcripts and proteins are stably expressed in all systems; therefore, reference molecules should be validated and demonstrated to be stably expressed between samples.^{6,19,55}

Some assays and prognostic variables are intrinsically subjective, such as histologic grades, assessments of percentage and/or intensity of immunolabeling, mitotic counts, and proliferation and apoptotic indices. Some subjectivity can be reduced through the use of image analysis software and morphometrics,^{11,47} but these technologies are not readily available in many diagnostic settings. Therefore, it is important to include precise details regarding how assessments were made. These details should include criteria for histologic grades, areas included in the analyses (eg, areas with highest immunohistochemical labeling versus random fields, tumor margins versus centers, avoidance of necrosis), how counts were performed, assay cut points and how they were determined, and whether or not evaluators were blinded to clinical outcomes. Validation of such assays should include measurements of intraobserver and interobserver variations. To address interobserver variation, a subset of samples should be independently evaluated by at least 2 investigators, and the agreement of these evaluations should be reported as a weighted kappa statistic for categorical data or estimate of limits of agreement for continuous data.⁸ In addition, at least 1 investigator should reevaluate a subset of samples to provide a measure of intraobserver variation. These measures of intraobserver and interobserver variation are critical to validate any prognostic marker for translation to clinical use, which should be the goal of all prognostic studies.

Aside from validating an assay as a whole, it is necessary to validate each run of a given assay. This is usually done through the inclusion of appropriate positive and negative controls in each assay run and the description of these controls in the methods. The results should include description of any background signal apparent in the negative controls.

Drawing Conclusions and Applying Results

As discussed above, a study's sample population should be a well-defined group that represents a larger population, and the study's results can be applied only to patient populations that

1. Prognostic Study Objective <ul style="list-style-type: none"> Clearly defined, testable hypothesis 	Tumor classification based on histology, location, and/or other criteria Uniform therapeutic treatment of the study population Signalment Inclusion of criteria according to current literature Similar case definition but difference in prognostic factor status
2. Characterization of the Study Population <ul style="list-style-type: none"> Case definition Well-defined study population Reference population 	
3. Sample Size <ul style="list-style-type: none"> Justification of study's sample size Power analysis 	Number of events (eg, recurrence, metastases, tumor-associated mortalities) should be 10 times the number of potential prognostic variables Length of follow-up Number of prognostic factors to be analyzed Expected magnitude of each factor's contribution Expected prognosis of the population
4. Outcome <ul style="list-style-type: none"> Defined endpoints being measured Event definition 	<i>Continuous variables:</i> Time to tumor progression (disease-free interval) and/or survival time <i>Categorical variables:</i> Recurrence and/or metastasis and/or tumor-associated mortality Criteria to define an event consistent and clearly stated Pathology considered the gold standard for the diagnosis of disease progression <i>Event identification based on diagnostic imaging:</i> —Criteria defined a priori —Final diagnosis best confirmed with pathology (eg, necropsy)
5. Statistical Analysis <ul style="list-style-type: none"> Censoring Descriptive statistics Models 	Continuous survival data censored Censoring criteria/parameters defined a priori Median survival times reported with confidence intervals Kaplan–Meier survival curves used for graphic displays Based on the hazard function—the risk of an event in a small interval after time <i>Multivariable models preferred for multiple prognostic variables:</i> —Cox proportional hazard model or accelerated time function models —Confounding variables defined and considered for inclusion — <i>Continuous versus categorical data:</i> Based on the data type; categorization should be avoided —Results reported with confidence intervals
6. Methods of Prognostic Markers Evaluation <ul style="list-style-type: none"> Materials/methods (for reproduction) 	Study population, sample/reagent sources, collection, handling, processing <i>Assay validation:</i> —Probe sensitivity and specificity —Inclusion of positive, negative, and species controls —Overall assay sensitivity, specificity, and reliability <i>Subjective markers:</i> —Definition of observation criteria —Interobserver variability with kappa statistics —Intraobserver variability with kappa statistics —Attempts to address variability
7. Drawing Conclusions and Applying Results <ul style="list-style-type: none"> Results Conclusions Discussions 	Applied only to the represented population—specific tumor type, specific treatment Restricted to data supported in the results; conclusions of prognostic studies based on survival endpoints Information regarding the strengths and limitations of a marker in a clinical setting

Figure 1. Components to be included for the conduct of a prognostic study, as recommended by the American College of Veterinary Pathologists and the World Small Animal Veterinary Association.

truly represent the sample population. For example, the Patnaik histologic grading system for canine cutaneous mast cell tumors (MCTs)³⁷ can be applied only to canine cutaneous MCTs. This histologic grading system is not applicable to canine gastrointestinal MCTs or feline cutaneous MCTs, because neither of these tumor subtypes were evaluated by Patnaik et al and both are distinct diseases with unique biological behaviors.^{32,36} Similarly, results can be applied only to populations that received a similar treatment as the study population. For example, studies evaluating a defined prognostic marker in dogs treated by surgery alone (ie, no additional chemotherapy or radiation therapy) cannot be uniformly applied to dogs that receive additional therapy. The reason is that in some cases, prognostic markers that indicate a worse prognosis for untreated animals, such as increased cellular proliferation, might indicate a better response to a specific chemotherapeutic protocol, such as a drug that targets rapidly dividing cells. Just as the study population must be specifically defined, study conclusions and future applications should be restricted to the specific conditions represented by the study population.

Conclusions should also be applied only to the degree that there are supporting data. There should be a direct flow of information—from the hypothesis proposed in the introduction to the data presented in the results and, finally, to the conclusions—with each previous section providing support for statements in the next. Conclusions should be based on data presented in the Results section, and authors should avoid overstating conclusions that can be drawn from the results. For example, although mutated p53 is frequently associated with increased protein stability and accumulation, immunohistochemical detection of increased levels of p53 cannot be used as conclusive evidence for the presence of p53 mutations, given that p53 regulation is complex and multifaceted.^{21,22,43,59} Similarly, conclusions about protein expression or activity cannot be stated on the basis of detection of mRNA expression by real-time quantitative PCR, because mRNA concentrations, protein concentrations, and their functionality do not always correlate.^{28,51} Most important, the prognostic value of a marker can be defined only when evaluations are made on the basis of biological endpoints. Many studies evaluate markers in terms of their associations to histologic grades or histologic patterns, such as, respectively, the Patnaik canine MCT grading system or canine mammary tumors with varying histologic features. Although MCT grades and types of mammary carcinomas correlate with prognosis,^{37,60} some dogs with grade III and many with grade II MCTs will experience long-term survival following surgical excision.^{37,42} Similar remarks can be made for historically more aggressive subtypes of mammary carcinomas, such as solid or anaplastic carcinomas.^{30,60} Therefore, histologic classifications cannot serve as meaningful surrogate endpoints for prognostic studies, but they may serve as a basis for building hypotheses. Metastasis, recurrence, disease-free interval, and overall survival should remain as gold standard endpoints for prognostic studies, and conclusions about the significance of prognostic markers should be based only on evaluations of these endpoints.

Finally, the goal of prognostic studies is to identify markers that are significantly and independently associated with a clinical outcome and that can, ideally, be applied to a clinical or diagnostic setting.²⁹ In light of this goal, a discussion is warranted of the strengths and limitations of the marker's use in a clinical setting and the means of integrating a marker into a routine diagnostic setting. Because the goal of using prognostic markers is to indicate disease outcome, authors should consider not simply the biological and statistical validity of the assay but also the clinical validity. These discussions will stimulate readers to consider how they can use the data to plan their studies and, more important, how they can use the data to improve their ability to diagnose and treat patients.

Summary

The goal of this endeavor is not to establish publication rules for *Veterinary Pathology* or other journals; rather, this work is meant to provide guidelines for conducting prognostic studies and to establish standards by which studies can be critically evaluated. To provide the reader with an easier approach to the proposed guidelines, Figure 1 summarizes the critical points in bullet-point format. As we the veterinary community refine our investigative work, we will identify new areas upon which we can improve, and we will need to continually make adjustments to meet evolving standards. Therefore, this article will need to be periodically reevaluated as an continually evolving document. This reevaluation will be essential to maintain the one steadfast goal: to identify and characterize prognostic and predictive markers that are scientifically sound and clinically applicable.

Acknowledgements

These guidelines were developed through an initiative of the American College of Veterinary Pathologists' Oncology Committee and represent the consensus opinion of the committee and the listed authors. The guidelines have also been reviewed and endorsed by the World Small Animal Veterinary Association. We would like to thank both organizations for their support and guidance.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Financial Disclosure/Funding

The authors declared that they received no financial support for their research and/or authorship of this article.

References

1. Bailey D, Erb H, Williams L, Ruslander D, Hauck M: Carboplatin and doxorubicin combination chemotherapy for the treatment of appendicular osteosarcoma in the dog. *J Vet Intern Med* **17**:199–205, 2003.
2. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF: Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* **276**:637–639, 1996.

3. Bradburn MJ, Clark TG, Love SB, Altman DG: Survival analysis: part II. Multivariate data analysis: an introduction to concepts and methods. *Br J Cancer* **89**:431–436, 2003.
4. Bradburn MJ, Clark TG, Love SB, Altman DG: Survival analysis: part III. Multivariate data analysis: choosing a model and assessing its adequacy and fit. *Br J Cancer* **89**:605–611, 2003.
5. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M: Minimum information about a microarray experiment (MIAME): toward standards for microarray data. *Nat Genet* **29**:365–371, 2001.
6. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* **55**:611–622, 2009.
7. Choi H, Charnsangavej C, Faria SC, Macapinlac HA, Burgess MA, Patel SR, Chen LL, Podoloff DA, Benjamin RS: Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol* **25**:1753–1759, 2007.
8. Christley RM, Reid SW: No significant difference: use of statistical methods for testing equivalence in clinical veterinary literature. *J Am Vet Med Assoc* **222**:433–437, 2003.
9. Clark TG, Bradburn MJ, Love SB, Altman DG: Survival analysis: part I. Basic concepts and first analyses. *Br J Cancer* **89**:232–238, 2003.
10. Clark TG, Bradburn MJ, Love SB, Altman DG: Survival analysis: part IV. Further concepts and methods in survival analysis. *Br J Cancer* **89**:781–786, 2003.
11. Cregger M, Berger AJ, Rimm DL: Immunohistochemistry and quantitative analysis of protein expression. *Arch Pathol Lab Med* **130**:1026–1030, 2006.
12. Deutsch EW, Ball CA, Berman JJ, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, Christiansen JH, Daian F, Dauga D, Davidson DR, Gimenez G, Goo YA, Grimmond S, Henrich T, Herrmann BG, Johnson MH, Korb M, Mills JC, Oudes AJ, Parkinson HE, Pascal LE, Pollet N, Quackenbush J, Ramialison M, Ringwald M, Salgado D, Sansone SA, Sherlock G, Stoeckert CJ Jr, Swedlow J, Taylor RC, Walashek L, Warford A, Wilkinson DG, Zhou Y, Zon LI, Liu AY, True LD: Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). *Nat Biotechnol* **26**:305–312, 2008.
13. Dohoo I, Martin W, Stryhn H: Cohort studies. *In: Veterinary Epidemiologic Research*, ed. Dohoo I, Martin W, Stryhn H, pp. 151–162. AVC, Charlottetown, Prince Edward Island, Canada, 2003.
14. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J: New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**:228–247, 2009.
15. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, DePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glockner FO, Goldstein P, Guralnick R, Haft D, Hancock D, Hermjakob H, Hertz-Fowler C, Hugenholtz P, Joint I, Kagan L, Kane M, Kennedy J, Kowalchuk G, Kottmann R, Kolker E, Kravitz S, Kyrpides N, Leebens-Mack J, Lewis SE, Li K, Lister AL, Lord P, Maltsev N, Markowitz V, Martiny J, Methe B, Mizrahi I, Moxon R, Nelson K, Parkhill J, Proctor L, White O, Sansone SA, Spiers A, Stevens R, Swift P, Taylor C, Tateno Y, Tett A, Turner S, Ussery D, Vaughan B, Ward N, Whetzel T, San Gil I, Wilson G, Wipat A: The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* **26**:541–547, 2008.
16. Fieten H, Spee B, Ijzer J, Kik MJ, Penning LC, Kirpensteijn J: Expression of hepatocyte growth factor and the proto-oncogenic receptor c-Met in canine osteosarcoma. *Vet Pathol* **46**:869–877, 2009.
17. Garzotto CK, Berg J, Hoffmann WE, Rand WM: Prognostic significance of serum alkaline phosphatase activity in canine appendicular osteosarcoma. *J Vet Intern Med* **14**:587–592, 2000.
18. Gieger TL, Theon AP, Werner JA, McEntee MC, Rassnick KM, DeCock HE: Biologic behavior and prognostic factors for mast cell tumors of the canine muzzle: 24 cases (1990–2001). *J Vet Intern Med* **17**:687–692, 2003.
19. Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L: Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot* **60**:487–493, 2009.
20. Guyatt G: Determining prognosis and creating clinical decision rules. *In: Clinical Epidemiology: How to Do Clinical Practice Research*, ed. Haynes RB, Sackett DL, Guyatt GH, Tugwell P, 3rd ed., pp. 323–355. Lippincott Williams & Wilkins, Philadelphia, PA, 2006.
21. Hall PA, Lane DP: p53 in tumour pathology: can we trust immunohistochemistry? Revisited! *J Pathol* **172**:1–4, 1994.
22. Hall PA, McCluggage WG: Assessing p53 in clinical contexts: unlearned lessons and new perspectives. *J Pathol* **208**:1–6, 2006.
23. Harrell FE Jr, Lee KL, Matchar DB, Reichert TA: Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer Treat Rep* **69**:1071–1077, 1985.
24. Hillers KR, Dernell WS, Lafferty MH, Withrow SJ, Lana SE: Incidence and prognostic importance of lymph node metastases in dogs with appendicular osteosarcoma: 228 cases (1986–2003). *J Am Vet Med Assoc* **226**:1364–1367, 2005.
25. Ishiguro T, Kadosawa T, Takagi S, Kim G, Ohsaki T, Bosnakovski D, Okumura M, Fujinaga T: Relationship of disease progression and plasma histamine concentrations in 11 dogs with mast cell tumors. *J Vet Intern Med* **17**:194–198, 2003.
26. Justice AC, Covinsky KE, Berlin JA: Assessing the generalizability of prognostic information. *Ann Intern Med* **130**:515–524, 1999.
27. Luong RH, Baer KE, Craft DM, Ettinger SN, Scase TJ, Bergman PJ: Prognostic significance of intratumoral microvessel density in canine soft-tissue sarcomas. *Vet Pathol* **43**:622–631, 2006.

28. Maier T, Guell M, Serrano L: Correlation of mRNA and protein in complex biological samples. *FEBS Lett* **583**:3966–3973, 2009.
29. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM: REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer* **93**:387–391, 2005.
30. Misdrorp W: Tumors of the mammary gland. *In: Tumors in Domestic Animals*, ed. Meuten DJ, 4th ed., pp. 575–606. Iowa State Press, Ames, IA, 2002.
31. Moher D, Schulz KF, Altman DG: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* **357**:1191–1194, 2001.
32. Molander-McCrary H, Henry CJ, Potter K, Tyler JW, Buss MS: Cutaneous mast cell tumors in cats: 32 cases (1991–1994). *J Am Anim Hosp Assoc* **34**:281–284, 1998.
33. Mullins MN, Lana SE, Dernell WS, Ogilvie GK, Withrow SJ, Ehrhart EJ: Cyclooxygenase-2 expression in canine appendicular osteosarcomas. *J Vet Intern Med* **18**:859–865, 2004.
34. Northrup NC, Harmon BG, Gieger TL, Brown CA, Carmichael KP, Garcia A, Latimer KS, Munday JS, Rakich PM, Richey LJ, Stedman NL, Cheng AL, Howerth EW: Variation among pathologists in histologic grading of canine cutaneous mast cell tumors. *J Vet Diagn Invest* **17**:245–248, 2005.
35. Northrup NC, Howerth EW, Harmon BG, Brown CA, Carmichael KP, Garcia AP, Latimer KS, Munday JS, Rakich PM, Richey LJ, Stedman NL, Gieger TL: 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* **17**:561–564, 2005.
36. Ozaki K, Yamagami T, Nomura K, Narama I: Mast cell tumors of the gastrointestinal tract in 39 dogs. *Vet Pathol* **39**:557–564, 2002.
37. Patnaik AK, Ehler WJ, MacEwen EG: Canine cutaneous mast cell tumor: morphologic grading and survival time in 83 dogs. *Vet Pathol* **21**:469–474, 1984.
38. Peters TJ: Multifarious terminology: multivariable or multivariate? univariable or univariate? *Paediatr Perinat Epidemiol* **22**:506, 2008.
39. Polton GA, Brearley MJ, Powell SM, Burton CA: Impact of primary tumour stage on survival in dogs with solitary lung tumours. *J Small Anim Pract* **49**:66–71, 2008.
40. Ramos-Vara JA, Kiupel M, Baszler T, Bliven L, Brodersen B, Chelack B, Czub S, Del Piero F, Dial S, Ehrhart EJ, Graham T, Manning L, Paulsen D, Valli VE, West K: Suggested guidelines for immunohistochemical techniques in veterinary diagnostic laboratories. *J Vet Diagn Invest* **20**:393–413, 2008.
41. Riley RD, Abrams KR, Lambert PC, Sutton AJ, Altman DG: Where next for evidence synthesis of prognostic marker studies? Improving the quality and reporting of primary studies to facilitate clinically relevant evidence-based results. *In: Advances in Statistical Methods for the Health Sciences: Application to Cancer AIDS Studies, Genome Sequence Analysis, and Survival Analysis*, ed. Auget JL, Balakrishnan N, Mesbah M, Molenberghs G, pp. 39–58. Birkhauser, Boston, MA, 2007.
42. Romansik EM, Reilly CM, Kass PH, Moore PF, London CA: Mitotic index is predictive for survival for canine cutaneous mast cell tumors. *Vet Pathol* **44**:335–341, 2007.
43. Save V, Nylander K, Hall PA: Why is p53 protein stabilized in neoplasia? Some answers but many more questions? *J Pathol* **184**:348–350, 1998.
44. Schwartz LH, Bogaerts J, Ford R, Shankar L, Therasse P, Gwyther S, Eisenhauer EA: Evaluation of lymph nodes with RECIST 1.1. *Eur J Cancer* **45**:261–267, 2009.
45. Simon R, Altman DG: Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* **69**:979–985, 1994.
46. Spangler WL, Kass PH: The histologic and epidemiologic bases for prognostic considerations in canine melanocytic neoplasia. *Vet Pathol* **43**:136–149, 2006.
47. Sullivan CA, Chung GG: Biomarker validation: in situ analysis of protein expression using semiquantitative immunohistochemistry-based techniques. *Clin Colorectal Cancer* **7**:172–177, 2008.
48. Takei H, Kurosumi M, Yoshida T, Ninomiya J, Hagiwara Y, Kamimura M, Hayashi Y, Tozuka K, Suemasu K, Inoue K, Tabei T: Current trends of sentinel lymph node biopsy for breast cancer: a surgeon's perspective. *Breast Cancer* **14**:362–370, 2007.
49. Taylor CF, Field D, Sansone SA, Aerts J, Apweiler R, Ashburner M, Ball CA, Binz PA, Bogue M, Booth T, Brazma A, Brinkman RR, Michael Clark A, Deutsch EW, Fiehn O, Fostel J, Ghazal P, Gibson F, Gray T, Grimes G, Hancock JM, Hardy NW, Hermjakob H, Julian RK Jr, Kane M, Kettner C, Kinsinger C, Kolker E, Kuiper M, Le Novere N, Leebens-Mack J, Lewis SE, Lord P, Mallon AM, Marthandan N, Masuya H, McNally R, Mehrle A, Morrison N, Orchard S, Quackenbush J, Reecy JM, Robertson DG, Rocca-Serra P, Rodriguez H, Rosenfelder H, Santoyo-Lopez J, Scheuermann RH, Schober D, Smith B, Snape J, Stoeckert CJ Jr, Tipton K, Sterk P, Untergasser A, Vandesompele J, Wiemann S: Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat Biotechnol* **26**:889–896, 2008.
50. Taylor CF, Paton NW, Lilley KS, Binz PA, Julian RK Jr, Jones AR, Zhu W, Apweiler R, Aebersold R, Deutsch EW, Dunn MJ, Heck AJ, Leitner A, Macht M, Mann M, Martens L, Neubert TA, Patterson SD, Ping P, Seymour SL, Souda P, Tsugita A, Vandekerckhove J, Vondriska TM, Whitelegge JP, Wilkins MR, Xenarios I, Yates JR III, Hermjakob H: The minimum information about a proteomics experiment (MIAPE). *Nat Biotechnol* **25**:887–893, 2007.
51. Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, Yi EC, Dai H, Thorsson V, Eng J, Goodlett D, Berger JP, Gunter B, Linseley PS, Stoughton RB, Aebersold R, Collins SJ, Hanlon WA, Hood LE: Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol Cell Proteomics* **3**:960–969, 2004.
52. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* **61**:344–349, 2008.
53. Von Euler H, Khoshnoud R, He Q, Khoshnoud A, Fornander T, Rutqvist LE, Skog S: Time-dependent RNA degradation affecting cDNA array quality in spontaneous canine tumours sampled using standard surgical procedures. *Int J Mol Med* **16**:979–985, 2005.

54. Wahl RL, Jacene H, Kasamon Y, Lodge MA: From RECIST to PERCIST: evolving Considerations for PET response criteria in solid tumors. *J Nucl Med* **50**:122S–150S, 2009.
55. Warrington JA, Nair A, Mahadevappa M, Tsyganskaya M: Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiol Genomics* **2**:143–147, 2000.
56. Webster JD, Kiupel M, Kaneene JB, Miller R, Yuzbasiyan-Gurkan V: The use of KIT and tryptase expression patterns as prognostic tools for canine cutaneous mast cell tumors. *Vet Pathol* **41**:371–377, 2004.
57. Webster JD, Yuzbasiyan-Gurkan V, Kaneene JB, Miller R, Resau JH, Kiupel M: The role of c-KIT in tumorigenesis: evaluation in canine cutaneous mast cell tumors. *Neoplasia* **8**:104–111, 2006.
58. Williams LE, Packer RA: Association between lymph node size and metastasis in dogs with oral malignant melanoma: 100 cases (1987–2001). *J Am Vet Med Assoc* **222**:1234–1236, 2003.
59. Wynford-Thomas D: P53 in tumour pathology: can we trust immunocytochemistry? *J Pathol* **166**:329–330, 1992.
60. Yamagami T, Kobayashi T, Takahashi K, Sugiyama M: Prognosis for canine malignant mammary tumors based on TNM and histologic classification. *J Vet Med Sci* **58**:1079–1083, 1996.