International Guidelines for Veterinary Tumor Pathology: A Call to Action

Donald J Meuten,[1] Frances M Moore,[2] Taryn A Donovan,[3] Christof A Bertram,[4] Robert Klopfleisch,[4] Robert Foster,[5] Rebecca Smedley,[6] Michael J Dark,[7] Milan Milovancev,[8] Paul Stromberg,[9] Bruce H Williams,[10] Marc Aubreville,[11] Giancarlo Avallone,[12] Pompei Bolfa,[13] John Cullen,[1] Michelle Dennis,[14] Michael Goldschmidt,[15] Richard Luong,[16] Andrew D Miller,[17] Margaret A Miller,[18] John Munday,[19] Paola Roccabianca,[20] Elisa Salas,[16] Yvonne Schulman,[21] Renee L Amorim,[22] Midori Asakawa,[23] Linden Craig,[24] Nick Dervisis,[25] D Glenn Esplin,[26] Jeannie George,[†] Marlene Hauck,[27] Yumiko Kagawa,[28] Matti Kiupel,[6] Keith Linder,[1] Kristina Meichner,[29] Laura Marconato,[30] Michelle Oblack,[5] Renato de Lima Santos,[31] Marc Simpson,[32] Harold Tvedten,[33] Derick Whitley[34]

[1] College of Veterinary Medicine, North Carolina State University, Raleigh NC

[2] Heska Corporation, Loveland, CO

[3] Department of Anatomic Pathology, Animal Medical Center, New York, NY

[4] Institute of Veterinary Pathology, Freie Universität Berlin, Berlin, Germany

[5] Ontario Veterinary College, University of Guelph, Guelph, ON, Canada

[6] Diagnostic Center for Population and Animal Health, Michigan State University, East Lansing, MI

[7] University of Florida Veterinary Diagnostic Laboratories and Emerging Pathogens Institute, College of Veterinary Medicine, Gainesville, FL

[8] Department of Clinical Sciences, College of Veterinary Medicine, Oregon State University, Corvallis OR

[9] Department of Veterinary Biosciences, College of Veterinary Medicine, Ohio State University, Columbus OH

[10] Veterinary Pathology Service, Joint Pathology Center, Silver Spring, MD

[11] Department of Computer Sciences, Technische Hochschule Ingolstadt, Ingolstadt, Germany

[12] Department of Veterinary Medical Sciences (DIMEVET), University of Bologna, Ossano dell'Emilia, Italy

[13] Department of Biomedical Sciences, Ross University School of Veterinary Medicine, PO Box 334, Basseterre, Saint Kitts and Nevis

[14] Department of Microbiology, Immunology and Pathology and Veterinary Diagnostic Laboratory, Animal Population Health Institute, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO

[15] Department of Pathobiology, Laboratory of Pathology and Toxicology, University of Pennsylvania, School of Veterinary Medicine, Philadelphia, PA

[16] IDEXX Laboratories, Westbrook, ME

[17] Section of Anatomic Pathology, Department of Biomedical Sciences, Cornell University College of Veterinary Medicine, Ithaca, NY

[18] Indiana Animal Disease Diagnostic Laboratory and Department of Comparative Pathobiology, Purdue University, West Lafayette, IN

[19] Pathobiology, Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand

[20] Dipartimento di Medicina Veterinaria, Universita degli Studi di Milano, Milan, Italy

[21] Marshfield Labs, Veterinary Services, Marshfield, WI

[22] Department of Pathology, Botucatu Veterinary and Animal Science School, UNESP, Botucatu, São Paulo, Brazil

[23] Veterinary Specialists Emergency Center, Saitama, Japan

[24] University of Tennessee, Knoxville, TN

[25] Department of Small Animal Clinical Sciences, Animal Cancer Care and Research Center, VA-MD College of Veterinary Medicine, Blackburg, VA

[26] Animal Reference Pathology, Salt Lake City, UT

[27] College of Veterinary Medicine, Raleigh, NC

[28] North Lab, Sapporo, Hokkaido, Japan

[29] Department of Pathology, College of Veterinary Medicine, University of Georgia, Athens, GA

[30] Centro Oncologico Veterinario, Bologna, Italy

[31] Departamento de Clínica e Cirurgia Veterinárias, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, MG, BR

[32] Laboratory of Cancer Biology and Genetics, Center for Cancer Research, National Cancer Institute, Bethesda, MD

[33] Swedish University of Agricultural Sciences, Uppsala, Sweden

[34] Antech Diagnostics, Stillwater OK

† deceased

Corresponding author: Donald J Meuten, College of Veterinary Medicine, North Carolina State University, Raleigh NC, USA, (919) 605-1691, Email: Don_meuten@ncsu.edu

Introduction

Reproducibility is the foundation of experimental science but irreproducibility of published oncological studies is a crisis in human oncology and certainly also a problem in veterinary oncology. In many instances the lack of reproducibility stems from inadequate description of published methods.(Begley and Ellis 2012, Oransky Ivan 2017, Stark 2018, Wen, Wang et al. 2018) Efforts to address this crisis have been implemented in human medicine, including attempted reproduction of published studies and formulation of checklists for use by journal editors and reviewers to ensure inclusion and transparency of detailed methods and materials in publications.(Oransky Ivan 2017, Editorial 2018, Schott, Tatiersky et al. 2018, Stark 2018, Wen, Wang et al. 2018) Less than 10% of observational studies are able to be replicated and incredibly, less than 20% of preclinical trials can be replicated.(Begley and Ellis 2012) If that is the state of oncology studies in human medicine, how do we compare in veterinary oncology? How far have we advanced in the last 40 years? Attempts to validate existing studies and or grading schemes are almost nonexistent. Grading schemes and the methods employed require appropriate validation before they should be adopted and used to provide prognoses or direct clinical therapy. Yet our philosophy seems to be that once a system or method is created, it is put in use and remains in use regardless of whether the system has been validated or not. We do not know how that system or method will perform when different pathologists use it and when it is applied to new patients. Consensus statements that support use of studies are not validation. Authors should feel complimented when colleagues attempt to reproduce their methods and study designs. Results will not replicate exactly, but our methods must. Validation of new grading systems is impossible if the original methods cannot be duplicated by other investigators.

Reproducibility

One of the major reasons that published studies are not able to be reproduced is the lack of sufficient details of the methods used to assess basic histological parameters including mitotic figure (MF) recognition, mitotic count (MC), lymphovascular invasion, tumor necrosis and margin evaluation.(Meuten, Munday et al. 2018, Schott, Tatiersky et

al. 2018) Currently, the assessment of these parameters requires pathologists to make subjective evaluations without clearly defined methods. Due to the inconsistency of these qualitative evaluations, there is weak or poor concordance between pathologists. This may result in negativity of the parameters or rejection of the grading system. The materials and methods section of manuscripts should contain descriptions of each method in sufficient detail to allow others to reproduce the study and validate the results. Citing that the methods described in a prior published study were followed is acceptable protocol, providing that any modifications used are described in detail. Failure of studies to be replicated can be due to poorly described methods, not following methods, and other confounders. Interobserver variation between pathologists reported in manuscripts is often ascribed to a method that is inadequate or too subjective. However, investigators may not have considered that the original methods were flawed or inadequately described, such that pathologists could not consistently follow the method. Stating that MF were counted in 10 consecutive high-power fields (hpf) at 400X is insufficient detail for others to reproduce the method, compare results and validate the data. Since the area within one hpf can vary by greater than 200% because of different microscope configurations, then of course there will be interobserver variation in MC if the microscopes used by study pathologists were not configured in the same way. Not only MC, but all parameters that were enumerated histologically (e.g. pleomorphism) with a microscope or with whole slide images (WSIs) have the potential for error and variability of results if the area enumerated is not defined in mm^2. Until methods are detailed such that others can reproduce them, we will have inconsistent and contradictory data in the literature. Even with standardized methods, there will be variability that needs to be reported and practical considerations that cannot be standardized.

Standardizing methodology is one step toward consistent results, but it does not guarantee consistency and certainly not usefulness. The methods must be followed, done carefully, using the same materials (e.g. antibody), and methods applied to reference populations and diseased groups with known and well defined outcomes. Accredited clinical pathology labs use standardized methods, calibration of instruments

and quality control measures to quantitate substances and report the reproducibility, sensitivity, specificity, validation, and reliability, such as positive and negative predictive values of test results. Similar principles need to be applied to anatomic pathology. Quantitation of morphologic structures by semi- or fully automated means will need to be validated and applied with similar rigid standards.(Boyce, Dorph-Petersen et al. 2010) When the methods are reproducible, they can be applied to cases with known outcomes. It would be helpful if new methods to quantitate structures were compared to the existing more subjective means to enumerate structures in HE stained slides.(Puri, Hoover et al. 2019, Bertram, Aubreville et al. 2020) Then colleagues can compare results to determine if they wish to adopt the new method. Technology will continue to spur development of new methods that can be applied to diagnostic cases. Many owners will pay for new techniques at any cost, but other owners will decline based on practical considerations such as cost, age of pet, or emotional value of the pet to their family. How to balance best care with practicality of animal ownership is not simple. Researchers can help address this by comparing new methodologies with those that can be performed without additional costs, specialized equipment, or expertise. Development of interlaboratory proficiency programs to promote standardization of tumor grading system results and performance of ancillary testing, such as immunohistochemistry, is sorely needed. Although new methods may initially be restricted to the institutions in which they were developed, standardization and proficiency are critical as these techniques are validated in other laboratories and become routinely used for tumor diagnosis. Centers that develop novel tests (e.g. computational pathology or CPATH), artificial intelligence, molecular, genetic) should have a goal that the methods can be applied uniformly and are described in sufficient detail that other labs can perform and validate the tests. Newly developed, specialized assays should be compared to current methodology and to patient outcomes to assess their utility and ideally seek FDA approval.(Boyce, Dorph-Petersen et al. 2010, Puri, Hoover et al. 2019, Bertram, Aubreville et al. 2020)

Outcome assessment

In addition to a lack of standardized assessment of histological criteria, reproducibility in animal studies is also limited by a lack of standardized guidelines for outcome assessments of animal patients.(Meuten, Munday et al. 2018, Schott, Tatiersky et al. 2018) Euthanasia unrelated to tumor progression appears to be a significant confounder. Reported patient survival times are impacted by euthanasia which may be elected due to personal decisions, varying judgements regarding quality of life, owner income or other factors which do not reflect tumor behavior. Furthermore, the Start time (T=0) needs to be clearly and consistently defined in presented survival analyses.(Nguyen) How often has an assigned grade or reporting that a tumor is in a lymph node resulted in euthanasia and the patient may have lived significantly longer? Survival time statistics in veterinary oncology are influenced by many factors outside of tumor and host biology. Metastasis needs to be subdivided into confirmed or suspected. Evidence for metastases determined by imaging should be labelled as suspected or *metastases as determined by imaging* when reported in journals. Histopathology is required to confirm that metastases are present and are of the same tumor type. Multiple aggressive tumors can occur in the same patient and are well recognized in breeds such as Golden retriever, Rottweiler and Bernese mountain dogs.(Cullen and Breen 2016) Oncology studies no longer routinely include results of autopsy, the perceived value of which seems to have hit a nadir. Owner/client permission to perform autopsies should be pursued with sympathy and empathy but as vigorously as other tests. Autopsy findings greatly increase the objectivity of results such as metastases and recurrence and therefore confidence of study results. Veterinary oncology studies need institutions and labs to pool their resources so that large numbers of cases can be collected. If results of these studies are correlated with accurate patient outcomes, the archived materials are a precious resource. The materials from these studies (slides, blocks, images, statistical data) could be shared with others such that new methods can be applied to case series with known outcomes. This was done by Bergin et al.(Bergin, Smedley et al. 2011) in a study of canine oral and lip melanocytic neoplasms. In addition to a set of oral/lip melanocytic neoplasms from the authors' own diagnostic laboratory, this study used archived blocks from two previous studies(Spangler and Kass 2006, Esplin 2008) in order to validate the histologic parameters described in

those studies (nuclear atypia, mitotic count, and pigmentation) using the same methods and to compare them to a new parameter, Ki67 index. This can serve as an example of the value in validating prior reports, which adds confidence to conclusions and provides a new method. Archived images of tumors could also be used to test inter-pathologist variation on diagnoses, MC, necrosis, and other basic parameters from laboratories worldwide. Outcome assessment determines if a test predicts treatment or prognosis and may differ between tertiary and primary care patients. Standardization of outcome assessment data is as critical as the standardization of techniques involved in tumor assessment. These two components of oncology must be linked as it is useless to create a new tumor grading system without knowledge of patient outcomes

Appendices and Protocols

In 2011 Veterinary Pathology published a series of recommendations and reviews about tumors in animals and how they should be evaluated. The manuscripts in that issue are excellent. They exceed the goals of the present manuscript but like "all" veterinary pathology publications there is no provision to update information, which is now a decade old. Much of the information consisted of literature reviews, and descriptions of the multiple methods to perform a parameter, without prioritizing or choosing one. The present manuscript aspires to be a continuum of the information published in 2011 but with a focus on establishing standardized histopathology methods to evaluate tumors. These methods are guidelines that will help accrue similar data such that studies can be cross compared and validated. A website will be established to publish guidelines for standard methods of tumor evaluations with the purpose of advancing veterinary pathology and oncology. This will require modifying the contents when publications have substantial data driven results that warrant updating these guidelines. These changes will be dated, and references cited. The present system of waiting for publication of a book or a fascicle is outdated. Updates are also needed as *errors* are possible (authors are humans) in the present appendices and protocols, and it is possible that some important references were missed. The authors hope that readers will bring such errors to our attention by contacting one or more of the *communication authors*. Unlike an error or omission in a manuscript or book that

remains in print, and results in our names indelibly associated with the words *retraction* or *correction,* the website can be quickly and easily updated. Journals and books will remain vital to our professions as they provide the means to publish peer reviewed research and to describe in detail an entire topic. The website will attempt to remain focused, and current, more of a CliffNotes' version of a topic designed to aid pathologists, editors and researchers in the standard parameters used to evaluate tumors and checklists of information that should be gathered about specific tumor types.

Appendices are guidelines to be used for identification of MF, perform MC, assess lymphovascular invasion (LVI), margin evaluation, percent tumor necrosis, CPATH, lymph nodes and outcome assessments. These parameters have not been standardized for animal tumors. The methods are detailed for MF, MC, LVI and margin assessments while others are newly developed methods (CPATH) or need clarification. Tumor necrosis is used in grading systems for some tumors, yet the method to determine percent necrosis in tumors from pets has never been described or not in sufficient detail such that others can reproduce the method (see Appendix 4). At the end of each appendix is a section titled "Future Considerations", which provides a list of possible ways to improve that method. *Protocols* are designed to gather complete data sets for the evaluation of commonly graded canine neoplasms. Protocols are provided for Soft Tissue Tumors/Soft Tissue Sarcomas (STT/STS), and are in process for canine mast cell tumor, and canine melanoma (cutaneous and oral). Protocols for other tumors can be developed and are needed, including mammary, splenic, osteosarcoma, hematopoietic and lymphoid tumors as well as cytologic protocols. If we do not standardize the methods used to identify tumors, we will continue to have conflicting data in the literature. Protocols and Appendices can be used as guides for reviewers and editors of manuscripts to ensure all required data was included and standard methods were followed. Journals serve as a gatekeeper for scientifically sound data, and they should also not refrain from publishing negative results. Investigators can use protocols as a checklist to ensure complete data sets are included for study participants. The protocols are modeled after the College of American Pathologists with an emphasis on gathering uniform data on specific tumor types. What are the consequences of not following an appendix or protocol? Nothing, no accreditation or certification or plaques

of accomplishment will be awarded or rescinded. The methods described herein are intended to be “best practices” that will add consistency and reproducibility to our methods with an eye to our clients: clinicians, oncologists, patients, and the public. Appendices and Protocols extend beyond “best practices” as they provide brief literature reviews, areas of weaknesses and list suggested fields of investigation for future studies to improve a method. The guidelines described are based upon review of literature and authors’ expertise, and are intended to bring consistency and reproducibility to the evaluation of tumors in animals. These guidelines have not been certified, accredited or reviewed by any standards-creating body and represent the authors’ own interpretation and application of the data reviewed. Application of these guidelines may vary with different laboratories and personnel, and each pathologist should consider whether these guidelines are appropriate based on the equipment, tissues or other materials available. Whether a governing body will aid in further development in updating these guidelines will depend upon the success of the website and how widely it is used.

Future Collaboration

The website being constructed will address some of these needs, but additional personnel will be needed to maintain the site, develop different protocols, generate new data, and validate studies. The initiative of a website with living appendices and tumor protocols will be successful if others use this information in their diagnostic, research, and publication efforts and if the appendices and protocols are updated in a timely manner as new information becomes available. A key benefit of standardization of tumor evaluation is the ability to evaluate data accrued from studies of many investigators at various institutions world-wide. This will permit analysis of larger data sets and increase the statistical power of the observations. The eventual goal would be to develop veterinary pathology industry standards with international input and acceptance. The goal is to accrue data on the important parameters that should be evaluated for a specific tumor type so that, over time, large data sets with comparable information about specific tumor types can be evaluated to provide accurate prognostic information that improves patient care. This will take multi-institutional participation and

specialists from different disciplines. The driving force will likely come from younger generations. Future appendices might include molecular profiles, genetic tests, and checklists for surgical pathology reports. Protocols are unlimited, think of a tumor, write a protocol using these as templates. Edits and updates are encouraged: contact the communication authors of an appendix or protocol. Submission of additional tumor protocols is welcomed and can be accomplished by contacting the administrators of the website. Confirming the need for standardized parameters to evaluate animal tumors met with near unanimity. Agreement for the guidelines of each parameter is not always unanimous. To compare data between labs, and ultimately improve patient care, we need to apply the same methods to basic parameters used to evaluate tumors. Using unstandardized methods that can cause variation in results is not scientifically sound. Drawing conclusions for clinical cases based on methods that are not standardized is misleading.

Completed Appendices and Protocols are in the supplemental section of this manuscript and they will be posted on the website and updated as needed (www.vetcancerguidelinesandprotocols.cldavis.org). The following are excerpts and summaries of each appendix or protocol, not the completed documents. Readers interested in a parameter should read the details in completed documents provided in the supplemental section of this manuscript and on the website.

SUMMARY

The goal of this project is to help advance veterinary oncology and pathology by promoting standardization of tumor assessment and patient outcomes. Guidelines are proposed to increase the uniformity and consistency of methods used to evaluate tumors along with suggestions for future consideration to help improve their discrimination and utility. Scientific journals, editors and reviewers can ensure progress in the goals of tumor assessment standardization and study reproducibility by establishing certain requirements of manuscripts being reviewed. Oncology studies which include histopathologic and gross features of tumors should have a pathologist as a co-author and journals should require this. Data obtained from record review without knowledge of the diagnostic or grading criteria limits conclusions and confidence in the

study. Review of gross description and histologic slides or images by an authoring pathologist or multiple pathologists is needed to ensure accuracy and uniformity of the pathology data and that current methods and terminology are used. The appendices are designed to help accomplish this. Certain appendices are completed: MF recognition, MC, necrosis, LVI, margin assessment and synoptic reporting, while others are in progress. The key steps to performing each method are condensed into checklists within the appropriate appendix. These checklists should integrate well with synoptic reporting (see Appendix 8). There are also discussion and notes to clarify certain steps. The checklists for margin evaluation are subdivided by responsible persons, the list that a pathologist should report are short and practical. LVI can be evaluated in HE sections and methods to confirm and differentiate LVI from pseudo-vascular invasion (see Appendix 3). Future investigations need to determine the importance of identifying if the tumor thrombus is in a lymphatic or blood vessel, and if the distinction has practical importance it will need to be determined how capable pathologists are of distinguishing each type of vessel with HE stained sections. Some authors would like to see necrosis abandoned as a parameter but that will require additional investigations. Suggestions to improve how necrosis is determined are provided in Appendix 4. CPATH will aid new investigations and synoptic reporting will provide a means to summarize and readily retrieve information. Outcome assessments are central to improvement of prognostic parameters but are under the umbrella of oncologists. However, histopathology is needed to confirm it is the same tumor in a recurrence or metastasis.

Until there are data driven results that can be standardized and proven prognostically useful, tumor assessment will need to include a wide range of parameters. Some practices, such as reporting margins of benign tumors or mitotic counts in tumors in which significance is not established will be left to the discretion of the pathologist and clinician. Clinicians faced with decisions on patient therapy rely extensively on pathologists' assessments. The prognostic significance of various factors changes over time necessitating clarity in communication of pathological findings, giving clinicians the information needed. The website is a window for clinicians to see pathologists' perspective of tumor assessment. Fascial planes to the surgeon are not the same as to the pathologist, a high power field is not a standard unit of area, if surgical margins are

not inked by the clinician, there is no accuracy to HTFD and there are other examples to illustrate our different perspectives. We need to do our best for the clients, owner, and pet, but we also need to explain and defend our discipline. The latter will be easier if veterinarians entering our profession understand our roles, and the limitations of our techniques.

The appendices and protocols require updating and renewal to be useful documents. Pathologists, oncologists and other scientists are encouraged to submit suggestions and supporting data to enable thoughtful revision. Tumor types and behavior may differ in various geographic sites and we encourage communication from all points of the globe to enhance our overall understanding of tumor behavior. Protocols are needed for additional tumor types and appendices should be developed for other parameters such as cytological assessments to recognize and grade specific tumor types, cellular and nuclear pleomorphism and proliferative indices. Research needs to clarify which technique and modifications enhance diagnostic and prognostic accuracy and if they can be practically applied to diagnostic cases, and subsequently validated with robust data. As in most research endeavors, new technology should be directed to answer specific problems and not end up as a new method in search of a question to answer.

Prospective studies that follow rigorous guidelines are the standard we should strive for and which will help guide the way forward.(Webster, Dennis et al. 2011) We also propose a platform from which new data can be gathered and integrated into an ongoing approach to evaluate the practicality and utility of current, as well as newer methods of tumor evaluation. Publishers can aid this project by providing permission for authors to copy sections of manuscripts they authored without forcing them to rewrite their own sentences to avoid plagiarism. How long will it take to accomplish all of this is unknown, but we need to continue and expand upon what our colleagues started in 2011.

References:

Begley, C. G. and L. M. Ellis (2012). "Raise standards for preclinical cancer research." Nature **483**(7391): 531-533.
Bergin, I., R. Smedley, D. Esplin, W. Spangler and M. Kiupel (2011). "Prognostic evaluation of Ki67 threshold value in canine oral melanoma." Veterinary pathology **48**(1): 41-53.
Bertram, C. A., M. Aubreville, C. Gurtner, A. Bartel, S. M. Corner, M. Dettwiler, O. Kershaw, E. L. Noland, A. Schmidt, D. G. Sledge, R. C. Smedley, T. Thaiwong, M. Kiupel, A. Maier and R. Klopfleisch (2020). "Computerized Calculation of Mitotic Count Distribution in Canine Cutaneous Mast Cell Tumor Sections: Mitotic Count Is Area Dependent." Vet Pathol **57**(2): 214-226.
Boyce, R. W., K.-A. Dorph-Petersen, L. Lyck and H. J. G. Gundersen (2010). "Design-based stereology: introduction to basic concepts and practical approaches for estimation of cell number." Toxicologic pathology **38**(7): 1011-1025.
Cullen, J. M. and M. Breen (2016). "An overview of molecular cancer pathogenesis, prognosis, and diagnosis." Tumors in Domestic Animals: 1-26.
Editorial (2018). "Checklists work to improve science." Nature **556**(7701): 273-274.
Esplin, D. (2008). "Survival of dogs following surgical excision of histologically well-differentiated melanocytic neoplasms of the mucous membranes of the lips and oral cavity." Veterinary pathology **45**(6): 889-896.
Meuten, D., J. S. Munday and M. Hauck (2018). "Time to Standardize? Time to Validate?" Vet Pathol **55**(2): 195-199.
Oransky Ivan, M. A. (2017). "Project to replicate landmark cancer studies meets mixed results." STAT News: https://www.statnews.com/2017/2001/2018/replication-cancer-studies/.
Puri, M., S. B. Hoover, S. M. Hewitt, B.-R. Wei, H. A. Adissu, C. H. Halsey, J. Beck, C. Bradley, S. D. Cramer and A. C. Durham (2019). "Automated computational detection, quantitation, and mapping of mitosis in whole-slide images for clinically actionable surgical pathology decision support." Journal of pathology informatics **10**.
Schott, C. R., L. J. Tatiersky, R. A. Foster and G. A. Wood (2018). "Histologic grade does not predict outcome in dogs with appendicular osteosarcoma receiving the standard of care." Veterinary pathology **55**(2): 202-211.
Spangler, W. L. and P. H. Kass (2006). "The histologic and epidemiologic bases for prognostic considerations in canine melanocytic neoplasia." Vet Pathol **43**(2): 136-149.
Stark, P. B. (2018). "Before reproducibility must come preproducibility." Nature **557**(7706): 613-614.
Webster, J., M. M. Dennis, N. Dervisis, J. Heller, N. J. Bacon, P. J. Bergman, D. Bienzle, G. Cassali, M. Castagnaro and J. Cullen (2011). "Recommended guidelines for the conduct and evaluation of prognostic studies in veterinary oncology." Veterinary pathology **48**(1): 7-18.
Wen, H., H.-Y. Wang, X. He and C.-I. Wu (2018). "On the low reproducibility of cancer studies." National science review **5**(5): 619-624.

Mitotic count (MC) and Histologic Morphology of Mitotic Figures (MF) (See Appendix 1 and Appendix 2)

MC will remain an important parameter in the evaluation of tumors as it is easy to accomplish, incurs no additional costs, is predictive of tumor proliferation and is part of multiple grading schemes that help predict tumor behavior. However, certain components essential for performing reproducible MC must be defined including the region of the tumor where MC should be performed (ie “hot spot” or areas of highest mitotic density within a tumor)(van Diest, Baak et al. 1992, Baak, , Meyer, Cosatto et al. 2009, Al-Janabi, van Slooten et al. 2013, , Veta, Van Diest et al. 2015, Meuten, Moore et al. 2016, ,) and the amount of tumor area in which MF will be counted, expressed in standard units of measure (ie mm^2).(Meuten, Moore et al. 2016) Although enumerating MF has long been a mainstay of tumor assessment, until recently there has been no standardization of any element of this parameter in veterinary pathology. Perhaps, the assumptions were that we were all counting the correct structures with the same method, that these methods matched published manuscripts and therefore there was no need to standardize the technique. Unfortunately, some of those assumptions are false. Performing the MC is considered laborious but subjective with inter-pathologist variation.(, Tsuda, Akiyama et al. 2000, Meyer, Alvarez et al. 2005, Veta, Van Diest et al. 2015, Bertram, Gurtner et al. 2018,) Possible causes include counting differently sized areas, poorly defined methods, not following methods, counting too rapidly, counting ambiguous structures, experience level, tumor mitotic heterogeneity, inability to find hot spots, quality of sections (fixation, artifacts) and quality of images.

To achieve accurate and consistent counts the MC must be performed carefully following standardized procedures; when this is done, consistent counts can be achieved by pathologists.(van Diest, Baak et al. 1992) After we follow standardized methods, these criticisms should be re-evaluated for manual and automated MC. MC can be determined by partially automated means, using artificial intelligence (AI, more specifically, deep learning-based algorithms). MC performed with computer systems can correct for interobserver variations associated with manual counts. They can better

identify hot spots (Aubreville, Bertram Deep learning algorithms outperform...2020), and they can count thousands of fields but may introduce different hurdles. High quality training datasets that adequately reflect the variability of histopathology sections and scanned images, along with validation of AI methods are paramount for CPATH to produce accurate and verifiable counts. With high quality data sets that define MF, atypical mitotic figures (AMF) possibly along with hard negatives such as mitotic like figures (MLF such as inflammatory cells or cells undergoing necrosis or apoptosis), automated means to perform MC should eventually be able to address potential confounders. Regardless of which mode, manual or automated, we propose that each of these elements needs to be standardized: 1. definition of MF, AMF and MLF; 2. the size of the area in which MF and AMF are counted; 3. the area of the tumor to be evaluated and 4. how to handle confounders. Each of these is described in Appendix 1 and 2, CPATH is in Appendix 5. At the end of all appendices are considerations for future studies which should help improve the method and clarify issues associated with assessing the parameter.

Histologic Morphology of Mitotic Figures (MF) (See Appendix 2)

What morphological features define a MF to be included in a MC?

The morphologic characteristics of MF and AMF and features which distinguish these from MLF are detailed in a recent publication. Mitotic figures and AMF are most easily identified by the short “rods” of chromosomes protruding from the surface of aggregates of nuclear material (Figures 1-4). Identification of the different phases of mitosis or the type of AMF are not necessary, but an understanding of the mitotic continuum and that AMF may have prognostic significance should be appreciated. Counting AMF may correlate with poorer prognosis and outcome as seen in some human tumors.(Jin, Stewenius et al. 2007, Matsuda, Yoshimura et al. 2016) Definitive MF (figures 1-4) and AMF (figures 5-8) should be included in the MC; however, structures with ambiguous morphology create a dilemma in classification. This is not problematic if the MC is markedly high (e.g. >20 MF/2.37mm^2).However, if the MC is close to an established threshold which has clinical significance, then the identity of these candidate structures could be critical (see MC Appendix 1.0). New thresholds should be established following

the guidelines in Appendices 1 and 2 and those thresholds should be tiered (avoid thresholds based on a< or> single number). Clinicians that request recounts because the MC of a tumor is at a threshold should seek different parameters to help establish the prognosis or direct therapy. We all likely have “non-standardized methods” that we use while counting MF but would not necessarily like others to know we do them: enumerating doubtful structures under a column labelled “?”; looking at extra fields when no MF were seen; looking at extra fields because there were spaces created by blood vessels, ducts or cysts; what to do when the tissue sample is <2.37mm^2; and/or looking for MF when the diagnosis of inflammation vs neoplasia is not clear. Practical considerations while performing a MC are listed in Appendix 1. Pathologists and laboratories will develop their own procedures to address MC reporting in non-routine situations. When solutions are found, the appendix will be updated accordingly. Correct identification of histologic structures will improve MC consistency and accuracy obtained from manual (glass or WSI) or CPATH modes.

Does the FN of an ocular matter?

For light microscopy, absolutely. It is the limiting factor that determines the diameter and therefore the area in the field of view (FOV) when objectives of the same magnification are used. Engraved or printed on some ocular eyepieces is a field number (FN) ranging from 6-28 mm. Higher numbers have larger FOV diameters and small increases in the FN will produce large increases in the area of FOV (see Appendix 1). The diameter of the FOV can be measured with a stage micrometer or it can be calculated by dividing the FN (mm) by the objective magnification. The formula for the area of a circle is used to calculate the area in the FOV. Therefore, a microscope with an ocular FN 18mm, 40X objective has a diameter of 0.45mm in the FOV and an area of 0.16 mm^2 per “hpf”; FN 26.5mm, 40X objective has a diameter of 0.66 mm and an area of 0.34 mm^2 per “hpf” which is a 100% larger area, a two fold increase (see Figure 3; Table 3 in Appendix 1).(Meuten, Moore et al. 2016)

Some objectives will have FN and/or NA (numerical aperture) numbers engraved or printed on them. Both are defined in the Appendix 9 “definitions and abbreviations”. NA is critical for resolution and depth of field but it is not used to calculate FOV. The higher

the NA the greater the resolution, or sharpness of features. All objectives have an FN but it may not be engraved on the objective. The FN of an objective can influence the FOV: however, it is the ocular FN which limits the maximum size of FOV in a standard microscope, not the objective FN.

What is the area in 10 high power fields (hpf)?

The area in 10 hpf is not a standard size as it varies up to 200% or more with the objective and the FN of the ocular.(Meuten, Moore et al. 2016) We proposed replacement of the imprecise phrase, *10 hpf* with *2.37 mm*^{2} to reflect the area equating to 10 hpf using a 40X objective and a 10X ocular FN 22mm, the most common configuration of pathologists' microscopes today. Furthermore, 10 hpf is nebulous for whole slide imaging which is likely the number one means for diagnostic tumor evaluation worldwide. A standard size area in mm^2 is required so the characteristics of the monitor and the magnification at which the image is reviewed can be configured to a specific area (see Appendix 1). Temporarily, retaining the phrase "10 hpf " together with accurate terminology (2.37 mm^2) clarifies communication with clinicians and permits MC to be determined with microscopes or WSI.

Mitotic counts (MC) reported in terms of high-power fields (hpf) without specific units of measurement (mm^2) cannot be compared to other MC as the area within one or especially 10 hpf is too variable.(Meuten, Moore et al. 2016) Older microscopes were equipped with ocular FN 18 (smaller FOV) compared to current microscopes which commonly have oculars of FN 22 or greater. Most prior animal studies did not define the area (mm^2) in which the MC or other histological features were enumerated, or defined the area incorrectly limiting the utility of this data for formulating prognoses for current cases. These studies need to be repeated with standardized methods of determining the basic histological parameters used to evaluate tumors. New methods should be considered and all must be correlated with outcome assessments.

Does the standard area need to remain 2.37mm^2?

No, it can be changed with data driven results. The total area evaluated can be amended for different tumors or unique situations. e.g. total tissue submitted is

<2.37mm^2; cystic tumors etc. Perhaps tumors with low proliferative rates require larger areas to be enumerated (5-10mm^2) or perhaps it is the opposite. What might be more important than a MC in one spot is what proportion of an entire tumor (or section) has low vs high proliferative rates. Greater than 85% of canine cutaneous MCT are indolent (Kiupel, Webster et al. 2011); perhaps determining the percent of a MCT that is "cold" (few hot spots, or areas of high mitotic activity) will predict how aggressive the tumor is. For canine oral melanoma, it might be the proportion of the tumor that is "hot" which is predictive. We also do not know how many sections of a tumor should be enumerated for the MC to be most predictive? This is true for other histologic parameters as well. These changes require correlating the different methods with known outcomes in many cases to show which method is predictive. Once a method is validated for a tumor type, the same size area, same region of the tumor and means to identify MF and AMF need to be validated if we want to compare results between labs or use published cutoffs of histologic parameters.

When multiple sections or regions are enumerated, should an average MC be reported or the ranges?(Meyer, Cosatto et al. 2009) Various guidelines have been proposed for determining the optimum tumor area for performing the MC in human tumors. Different sized areas are recommended to perform MC for different tumors. Some authors recommend counting a series of 5 or more sets of MC and reporting the average. Others report the highest MC. There are a multitude of scenarios that need investigation to change how we determine MC, and CPATH will greatly aid these studies because MC can be performed faster, more consistently, and can be performed over differently sized areas in different regions of the tumors. CPATH can report the proportion of a tumor that is *hot* or *cold*. Manual counts for these types of studies will be laborious. Studies using CPATH should also include the standard means of determining the MC and compare the various methodologies to known outcomes. Hopefully, these studies will avoid creating MC cutoffs that are based on a single number (above or below) and develop scoring systems, confidence intervals, and ranges of predictability for MC for different tumors.

Until data driven results provide new methods, an area equivalent to 2.37mm^2 should be used for MC and should be reported as mm^2 rather than stating the FN of the ocular or how the scope is configured.

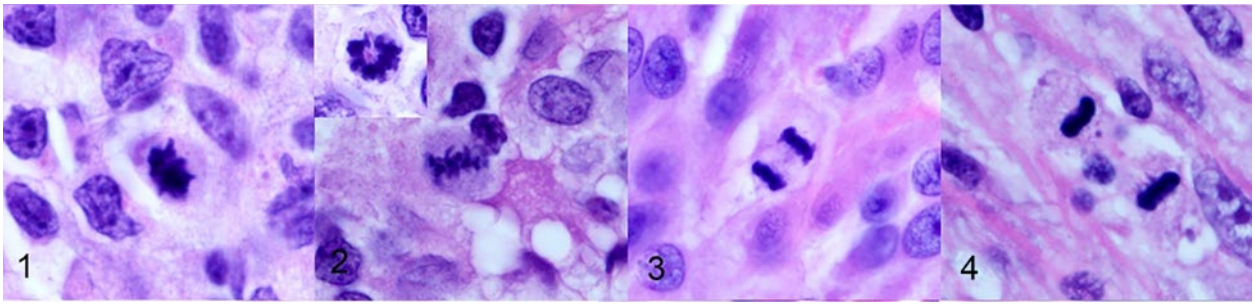
Where in the tumor should the MC be performed?

Presently MF and AMF should be counted in regions of *hot spots* or high mitotic activity in viable regions of tumor. It is logical to choose regions of high tumor cell proliferation because the cells in these areas may be more aggressive, they already may have the potential to metastasize or they have a greater opportunity to form a clone with metastatic potential. Until studies report that a different region is more predictive of outcomes, we should adhere to this method. There are no studies in animals that correlate MC determined in different regions with outcomes. Multiple studies in humans and one in dogs have demonstrated variability in the number of MF in different regions of tumors.(Bertram, Aubreville et al. 2020) We know there is heterogeneity of MF distribution in tumors, but we do not know if it matters, and we will not know until there are outcome assessments correlated to methods. Different regions and differently sized areas of different tumor types are used to perform MC in human tumors, and different cutoffs of MC are used to determine prognoses. Similar studies need to be done with animal tumors, and when these are performed, investigators should include newer technologies as well (molecular, CPATH etc., https://www.cap.org/protocols-and-guidelines/cancer-reporting-tools/cancer-protocol-templates)

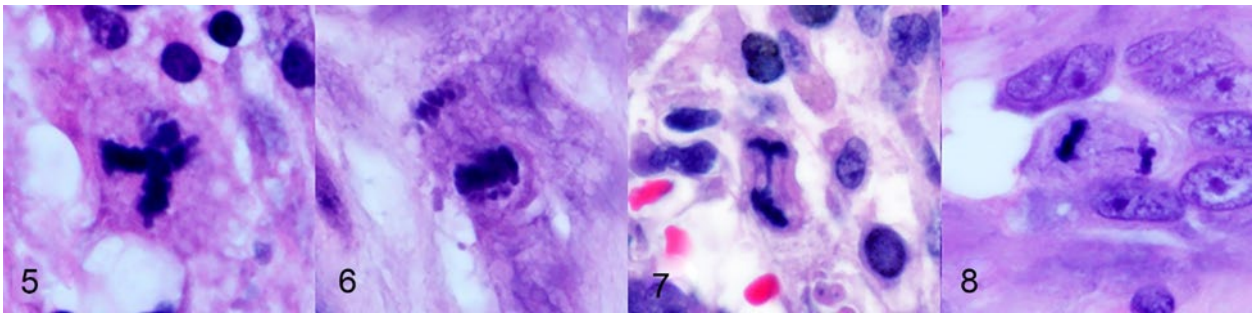
The periphery of some tumors is the preferred site because this is the invasive front, fixation is better, and there is a higher proliferative rate. A study of human breast carcinoma reported that the periphery contained more hot spots (using Ki67) than other regions and percentages of Ki67 positive nuclei obtained at the periphery changed the prognosis.(Gudlaugsson, Skaland et al. 2012) Other studies in humans reported that using Ki67 in hot spots, which were not just at the periphery of breast carcinoma, contributed the most prognostic information as compared to other methods.(Stålhammar, Robertson et al. 2018) Additionally, a study of canine cutaneous mast cell tumors did not find that the regions of highest mitotic activity were always at the periphery.(Bertram, Aubreville et al. 2020) Selecting the area of a tumor that is

predictive of outcome(s) or treatments needs to be found for each tumor type. Until those locations are identified, MC should be performed in regions of hot spots. However, determination of hot spots by routine light microscopy is subjective and a source of interobserver variation.(van Diest, Baak et al. 1992, Bertram, Aubreville et al. 2020) A study with canine MCT and one with canine melanoma showed that pathologists were not as capable of finding the hotspots as compared with computer-assisted localization of hot spots.)(Puri, Hoover et al. 2019, Aubreville, Bertram et al. 2020)

Summary: Appendix 1 and 2 detail the standard method of performing a MC including: definitions of MF, AMF and MLF, contiguous 2.37 mm^2 area, hot spot, practical considerations, and future considerations of how the MC can be improved. The present standard means to perform the MC will be modified when data-driven changes necessitate, and the appropriate appendices will subsequently be updated.



Figures 1-4: Mitotic Figures (MF) are characterized by dark aggregates of nuclear material with short rods and projections. Figure 1: Prometaphase/metaphase (dense nuclear cluster with short protruding rods). Figure 2: Metaphase with linear equatorial plate of darkly staining nuclear material and short protruding rods and spikes. Inset: Ring form of metaphase with end-on (non-perpendicular) view of the equatorial plate. Figure 3: Anaphase MF with two separate nuclear aggregates with irregular contours and short protruding spikes. Figure 4: Telophase MF with aggregates at opposite ends of the cell and formation of a cleavage furrow.



Figures 5-8: Atypical MF (AMF). Figure 5: Tripolar AMF (more than two spindle poles during any stage of mitosis). Figure 6: Asymmetric AMF (unequal sizes of the metaphase axes or anaphase poles). Figure 7: AMF with anaphase bridging (chromosomes stretching from one pole to the other). Figure 8: Lagging chromosomes left behind during anaphase (small dark purple streak in center of cell).

References (A full list of references is available in Appendices 1 and 2)

Aubreville, M., C. A. Bertram, C. Marzahl, C. Gurtner, M. Dettwiler, A. Schmidt, F. Bartenschlager, S. Merz, M. Fragoso, O. Kershaw, R. Klopfleisch and A. Maier (2020). "Deep learning algorithms out-perform veterinary pathologists in detecting the mitotically most active tumor region." Sci Rep **10**: 16447.

Bertram, C. A., M. Aubreville, C. Gurtner, A. Bartel, S. M. Corner, M. Dettwiler, O. Kershaw, E. L. Noland, A. Schmidt, D. G. Sledge, R. C. Smedley, T. Thaiwong, M. Kiupel, A. Maier and R. Klopfleisch (2020). "Computerized Calculation of Mitotic Count Distribution in Canine Cutaneous Mast Cell Tumor Sections: Mitotic Count Is Area Dependent." Vet Pathol **57**(2): 214-226.

Gudlaugsson, E., I. Skaland, E. A. Janssen, R. Smaaland, Z. Shao, A. Malpica, F. Voorhorst and J. P. Baak (2012). "Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer." Histopathology **61**(6): 1134-1144.

Jin, Y., Y. Stewenius, D. Lindgren, A. Frigyesi, O. Calcagnile, T. Jonson, A. Edqvist, N. Larsson, L. M. Lundberg, G. Chebil, F. Liedberg, S. Gudjonsson, W. Mansson, M. Hoglund and D. Gisselsson (2007). "Distinct mitotic segregation errors mediate chromosomal instability in aggressive urothelial cancers." Clin Cancer Res **13**(6): 1703-1712.

Kiupel, M., J. D. Webster, K. L. Bailey, S. Best, J. DeLay, C. J. Detrisac, S. D. Fitzgerald, D. Gamble, P. E. Ginn, M. H. Goldschmidt, M. J. Hendrick, E. W. Howerth, E. B. Janovitz, I. Langohr, S. D. Lenz, T. P. Lipscomb, M. A. Miller, W. Misdorp, S. Moroff, T. P. Mullaney, I. Neyens, D. O'Toole, J. Ramos-Vara, T. J. Scase, F. Y. Schulman, D. Sledge, R. C. Smedley, K. Smith, W. S. P, E. Southorn, N. L. Stedman, B. A. Steficek, P. C. Stromberg, V. E. Valli, S. E. Weisbrode, J. Yager, J. Heller and R.

642 Miller (2011). "Proposal of a 2-tier histologic grading system for canine cutaneous mast
643 cell tumors to more accurately predict biological behavior." Vet Pathol **48**(1): 147-155.
644 Matsuda, Y., H. Yoshimura, T. Ishiwata, H. Sumiyoshi, A. Matsushita, Y. Nakamura, J.
645 Aida, E. Uchida, K. Takubo and T. Arai (2016). "Mitotic index and multipolar mitosis in
646 routine histologic sections as prognostic markers of pancreatic cancers: A
647 clinicopathological study." Pancreatology **16**(1): 127-132.
648 Meuten, D. J., F. M. Moore and J. W. George (2016). "Mitotic Count and the Field of
649 View Area: Time to Standardize." Vet Pathol **53**(1): 7-9.
650 Meyer, J. S., E. Cosatto and H. P. Graf (2009). "Mitotic index of invasive breast
651 carcinoma. Achieving clinically meaningful precision and evaluating tertial cutoffs." Arch
652 Pathol Lab Med **133**(11): 1826-1833.
653 Puri, M., S. B. Hoover, S. M. Hewitt, B.-R. Wei, H. A. Adissu, C. H. Halsey, J. Beck, C.
654 Bradley, S. D. Cramer and A. C. Durham (2019). "Automated computational detection,
655 quantitation, and mapping of mitosis in whole-slide images for clinically actionable
656 surgical pathology decision support." Journal of pathology informatics **10**.
657 Stålhammar, G., S. Robertson, L. Wedlund, M. Lippert, M. Rantalainen, J. Bergh and J.
658 Hartman (2018). "Digital image analysis of Ki67 in hot spots is superior to both manual
659 Ki67 and mitotic counts in breast cancer." Histopathology **72**(6): 974-989.
660 van Diest, P. J., J. P. Baak, P. Matze-Cok, E. Wisse-Brekelmans, C. Van Galen, P.
661 Kurver, S. Bellot, J. Fijnheer, L. Van Gorp and W. Kwee (1992). "Reproducibility of
662 mitosis counting in 2,469 breast cancer specimens: results from the Multicenter
663 Morphometric Mammary Carcinoma Project." Human pathology **23**(6): 603-607.

664

Lymphovascular Invasion (See Appendix 3)

Neoplastic cell invasion of blood vessels or lymphatics is widely recognized as evidence of tumor aggressiveness and potential malignancy in both humans (Falvo, Catania et al. 2005, Mete and Asa 2011) and animals (Goldschmidt, Pena et al. 2011, Rasotto, Berlato et al. 2017) but, despite, the importance of this parameter, criteria to definitively identify lymphovascular invasion (LVI) and distinguish from pseudo-vascular invasion or retraction artifact are lacking in the veterinary literature. This lack of stringent, standardized criteria may have led to misdiagnosis of LVI in veterinary oncology studies. Assessing LVI with criteria of varying stringency has revealed key insights into the biological behavior of human cancers as has the distinction between blood vascular and lymphatic invasion.(Van den Eynden, Van der Auwera et al. 2006, Lin, Zhu et al. 2010, Mete and Asa 2011) In veterinary medicine, LVI is recognized as a marker of potential tumor malignancy but this parameter has only been extensively evaluated in canine and feline mammary tumors(Goldschmidt, Pena et al. 2011, Rasotto, Berlato et al. 2017) without establishment of strict criteria for LVI diagnosis or comparison of blood vascular and lymphatic invasion.

Mimickers of LVI, such as pseudo-vascular invasion and retraction artifacts are not adequately addressed in the veterinary literature; images of each can be found in Appendix 3 and on the website. Pseudo-vascular invasion is the presence of neoplastic cells within vascular spaces, but the cells are not present because of tumor invasion of vessels. Displacement of neoplastic cells into vessels secondary to manipulation of the neoplasm at the time of biopsy, surgical excision, grossing procedure or tissue sectioning (ie, “floaters”) can result in pseudo-vascular invasion.(Van den Eynden, Van der Auwera et al. 2006, Mete and Asa 2011) This is also reported for non-neoplastic lesions in the thyroid. (Mete and Asa, 2011) Neoplastic cells may protrude or impinge into adjacent vascular lumens without true invasion in which case endothelial cells cover the surface of the impinging tumor. However, endothelium may also line the surface of neoplastic cells which have invaded through the vascular endothelium but have undergone re-endothelialization, necessitating searching for other criteria of LVI to confirm which is the correct interpretation.

Distinguishing between these various manifestations of pseudo-vascular invasion and true LVI relies on identification of more robust LVI criteria. The two most definitive criteria used to define LVI in human tumors include: thrombus adherent to intravascular tumor and tumor cells invading through the vessel wall and endothelium. Additional criteria are listed in Appendix 3 along with a complete reference list.(Mete and Asa 2011. These criteria should be used to assess LVI in tumors from animals.

Retraction artifact, another mimicker of LVI, forms an artifactual space surrounding tumor foci and can be distinguished from intravascular neoplasia by the absence of an endothelial cell lining. Retraction artifact is seen in epithelial tumors in which tumor cells retract from surrounding stroma (Figure 5 in Appendix 3).

Studies of human breast, thyroid and prostate cancer show widespread metastases are more commonly associated with blood vascular invasion in contrast to lymphatic invasion.(Mete and Asa 2011) Animal tumors may show similar distinctions between blood and lymphatic vascular invasion, warranting detailed descriptions of the type of vessels invaded (ie, if a muscular wall is discerned in the involved vessels) or use of immunohistochemical markers to distinguish blood from lymphatic vessels. A variety of immunohistochemical markers have been used to identify endothelial cells in blood and lymphatic vascular channels in humans and animals (Von Beust, Suter et al. 1988, Sleeckx, Van Brantegem et al. 2013, Wennogle, Priestnall et al. 2019, Fitzgibbons, Connolly et al. 2020) Some markers, such as CD31 and Factor VIII related antigen, do not discriminate between lymphatic and blood vascular endothelium, whereas others, such as Lymphatic vessel endothelial receptor 1 (LYVE-1), D2-40 and prospero – related homeobox gene-1 (PROX-1) are specific for lymphatic endothelium.(Von Beust, Suter et al. 1988, Pusztaszeri, Seelentag et al. 2006, Sleeckx, Van Brantegem et al. 2013, Halsey, Worley et al. 2016, Wennogle, Priestnall et al. 2019, Fitzgibbons, Connolly et al. 2020) Use of IHC endothelial markers has been shown to facilitate identification of LVI in tumors in humans(O'Donnell, Feldman et al. 2008,) and in mammary and plasma cell tumors in dogs.(Sleeckx, Van Brantegem et al. 2013, Ehrensing and Craig 2018) Validation of IHC markers and antibodies used to differentiate lymphatic vs blood vessels for the different animal species is a necessity.

Although IHC confirms the identity of the vascular structure it does not confirm true LVI and, in fact, is not one of the more stringent criteria of LVI.

Studies of tumor lymphovascular density (LVD) in humans have been correlated with LVI in a number of human tumors.(complete list of references in Appendix 3) LVD is an enumeration of lymphatics within a defined area of a tumor and is used as an indicator of lymphangiogenesis and therefore probable lymph node metastasis. Both LVD and LVI are used as predictors of lymph node metastases in human breast cancer, and peritumoral lymphatic vessels may be the main route for dissemination of the tumor. Intratumoral microvascular density (IMD), the quantitation of blood vessels (number/mm^2) in tumors, is used as an indicator of angiogenesis or vasculogenesis and by extension LVI and the ability of a tumor to metastasize. New blood vessels in a tumor are required for tumors to grow beyond several millimeters; they are believed to facilitate metastasis and are associated with more aggressive neoplasms in humans and animals. Although IMD has been assessed in a number of animal tumors and has been associated with higher grade or more malignant histological features (ie canine: soft tissue sarcomas, mammary gland tumors, seminomas, cutaneous squamous cell carcinoma,) and cutaneous mast cell tumors),(full reference listing in Appendix 3) there have been no comprehensive studies of intratumoral versus peritumoral vascular density nor associations between IMD and blood vascular or lymphatic vascular invasion in domestic animals. Future veterinary studies comparing intratumoral versus peritumoral microvascular density and correlation with nodal and systemic metastases are warranted.

A thorough reassessment of LVI is needed in veterinary oncology with attention to the specific details described in the *appendix LVI* and under future considerations. These studies should use the criteria outlined to determine if LVI is present, especially focusing on the more definitive features: invasion through vessel wall and endothelium and thrombus adherent to the tumor. Studies should include detailed descriptions of criteria used to establish presence of LVI and clarify the importance of lymphatic versus blood vascular invasion. Quantitation of blood and lymphatic vessels (IMD, LVD) may benefit from the use of CPATH, and both subjective and quantitative analyses should be

755 correlated with nodal and systemic metastases and, most importantly, known patient
756 outcomes.
757
758 If individuals have images of true LVI and pseudo-vascular invasion please share them
759 with the communication author of appendix 3.
760

761 **References (**A full list of references is available within Appendix 3)

762 Ehrensing, G. and L. E. Craig (2018). "Intravascular neoplastic cells in canine
763 cutaneous plasmacytomas." J Vet Diagn Invest **30**(2): 329-332.
764 Falvo, L., A. Catania, V. D'Andrea, A. Marzullo, M. C. Giustiniani and E. De Antoni
765 (2005). "Prognostic importance of histologic vascular invasion in papillary thyroid
766 carcinoma." Annals of surgery **241**(4): 640.
767 Fitzgibbons, P. L., J. Connolly and S. Bose (2020). "Protocol for the Examination of
768 Resection Specimens from Patients with Invasive Carcinoma of the Breast." College of
769 american pathologists(www.cap.org/cancerprotocols.).
770 Goldschmidt, M., L. Pena, R. Rasotto and V. Zappulli (2011). "Classification and grading
771 of canine mammary tumors." Vet Pathol **48**(1): 117-131.
772 Im, K. S., N. H. Kim, H. Y. Lim, H. W. Kim, J. I. Shin and J. H. Sur (2014). "Analysis of a
773 new histological and molecular-based classification of canine mammary neoplasia." Vet
774 Pathol **51**(3): 549-559.
775 Lin, X., B. Zhu, Y. Liu and J. F. Silverman (2010). "Follicular thyroid carcinoma invades
776 venous rather than lymphatic vessels." Diagnostic pathology **5**(1): 1-5.
777 Mete, O. and S. L. Asa (2011). "Pathological definition and clinical significance of
778 vascular invasion in thyroid carcinomas of follicular epithelial derivation." Mod Pathol
779 **24**(12): 1545-1552.
780 Nishida, T., S.-i. Katayama and M. Tsujimoto (2002). "The clinicopathological
781 significance of histologic vascular invasion in differentiated thyroid carcinoma." The
782 American journal of surgery **183**(1): 80-86.
783 O'Donnell, R. K., M. Feldman, R. Mick and R. J. Muschel (2008). "Immunohistochemical
784 method identifies lymphovascular invasion in a majority of oral squamous cell
785 carcinomas and discriminates between blood and lymphatic vessel invasion." J
786 Histochem Cytochem **56**(9): 803-810.
787 Pifer, P., R. Bhargava, A. Patel, D. Ling, J. Vargo, B. Orr, P. Sukumvanich, M.
788 Courtney-Brooks, M. Boisen and J. Berger (2020). "Is the risk of substantial LVSI in
789 stage I endometrial cancer similar to PORTEC in the North American population?–A
790 single-institution study." Gynecologic Oncology.
791 Rasotto, R., D. Berlato, M. H. Goldschmidt and V. Zappulli (2017). "Prognostic
792 Significance of Canine Mammary Tumor Histologic Subtypes: An Observational Cohort
793 Study of 229 Cases." Vet Pathol **54**(4): 571-578.
794 Sleeckx, N., L. Van Brantegem, E. Fransen, G. Van den Eynden, C. Casteleyn, E.
795 Veldhuis Kroeze and C. Van Ginneken (2013). "Evaluation of immunohistochemical

markers of lymphatic and blood vessels in canine mammary tumours." J Comp Pathol **148**(4): 307-317.
Van den Eynden, G., I. Van der Auwera, S. Van Laere, C. Colpaert, P. van Dam, L. Dirix, P. Vermeulen and E. Van Marck (2006). "Distinguishing blood and lymph vessel invasion in breast cancer: a prospective immunohistochemical study." British journal of cancer **94**(11): 1643-1649.
Von Beust, B., M. Suter and B. Summers (1988). "Factor VIII-related antigen in canine endothelial neoplasms: an immunohistochemical study." Veterinary Pathology **25**(4): 251-255.
Wennogle, S. A., S. L. Priestnall, A. Suarez-Bonnet, S. Soontararak and C. B. Webb (2019). "Lymphatic endothelial cell immunohistochemical markers for evaluation of the intestinal lymphatic vasculature in dogs with chronic inflammatory enteropathy." J Vet Intern Med **33**(4): 1669-1676.

Necrosis (See Appendix 4)

The extent of tumor necrosis has been correlated with tumor biological behavior and is a parameter used in grading schemes in humans. Tumor necrosis has also been included as a grading scheme parameter in animals, primarily in dogs with STS/STT but is also used in other grading schemes (canine primary pulmonary carcinoma). Criteria for determining the percent of tumor necrosis in all species have not been adequately described (Kuntz, Dernell et al. 1997, Coindre 2006) Necrosis within a tumor is often subjectively and vaguely used to suggest a tumor is aggressive. In humans, the percent of tumor necrosis has been determined by estimating the amount seen grossly and histologically, whereas animal studies have not indicated if gross observations were used in combination with histological assessment, or if only histologic assessments were evaluated.(Trojani, Contesso et al. 1984, Kuntz, Dernell et al. 1997, Coindre 2006, McSporran 2009, Vayrynen, Vayrynen et al. 2016, Laurini, Blanke et al. 2017, Dobromylskyj, Richards et al. 2020) Many tumors are larger than a histological section, and measuring or estimating percentage of necrosis is more problematic.(Chiang and Oliva 2013)

In grading human soft tissue sarcomas, necrosis was found to be one of three parameters correlating with patient survival and tumor metastasis, along with tumor differentiation and mitotic count.(Trojani, Contesso et al. 1984) The thresholds for scoring necrosis histologically were no necrosis (score 0), less than 50% of tumor necrosis (score 1) and greater than 50% tumor necrosis (score 2) but how a pathologist was to estimate those percentages was not detailed. A grade of two could also be assessed for any neoplasm whose gross appearance was described as “mainly necrotic” by a surgeon or pathologist even if no necrosis was seen on the submitted sections.(Trojani, Contesso et al. 1984) We do not recommend this last criterion be adopted for animal tumors, and later authors and grading systems, as well as current College of American Pathologist protocol for assessment of soft tissue tumors in humans, require microscopic confirmation/validation of macroscopic evidence suggesting necrosis.(Coindre 2006, Laurini, Blanke et al. 2017) This brings us to the

problems associated with gross interpretation of necrosis (and to a lesser extent, its histologic interpretation). Even for an experienced pathologist, the gross diagnosis of necrosis may be problematic, and most pathologists in veterinary pathology will not see the gross specimen. Areas of edema or exudate may be interpreted as areas of necrosis grossly, and areas of hemorrhage, which are often associated with necrosis, may far exceed the boundaries of actual necrotic tissue. These problems are further compounded by certain histologic lesions such as myxomatous change, cystic space formation, edema, hemorrhage and exudate which can resemble or obscure necrotic areas. Gross/macroscopic assessment of necrosis requires histologic confirmation which, in large tumors, may not be practical for veterinary diagnosticians to submit an adequate number of sections (costs) but should be done in research studies. The number of sections examined at trimming and or submitted for histopathology for routine diagnostic cases is likely far fewer in veterinary than human pathology. If gross assessment is to be used as a parameter, numerous confounders must be clarified in future studies. This requires documentation of systematic sampling of both necrotic and viable tissue during the gross examination and confirmation of necrosis by histological evaluation. Alternatively, we can abandon the use of gross assessment and only use light microscopy. This would be straightforward, but if gross assessment of tumor necrosis improves the prognostic utility of grading systems then it would be lost as a parameter.

Although it seems obvious that the means to assess various histologic parameters need to be defined prior to implementation, this has not always happened, e.g. the area in which MF were counted was never standardized and the same seems true for percent necrosis. The percent of tumor necrosis in soft tissue mesenchymal tumors/soft tissue sarcomas (STT/STS) is included in grading schemes, yet the means to assess necrosis has not been clearly defined or standardized. Was the percent necrosis determined by examination of the tumor during gross sectioning, and were areas appearing necrotic confirmed microscopically? Was the percent necrosis used in the grading system based upon visual estimate of necrosis in random histologic tumor sections? Was a consistent portion of the tumor submitted for microscopic examination? A recent publication suggested preparation of 1 tissue block for each 2 cm diameter of soft tissue

tumors(Roccabianca, Schulman et al. 2020). Since no formulae for number of blocks/slides per tumor have been described in published grading systems for dogs this seems like a good starting point, but no studies using this guidance were referenced.

The necrosis appendix (Appendix 4) provides guidelines for recording and scoring extent of tumor necrosis on gross and histologic tumor evaluation which should enable evaluation of the utility of this parameter to assess tumor prognosis and patient outcomes. The scoring system proposed is based on prior reports and is indicated above but includes an unusual percentage of <10% for future studies and explains the logic for this. Additionally for necrosis to be objectively assessed as a parameter for future grading schemes, new studies must determine if gross assessment of necrosis can be documented in a standardized fashion and if this parameter correlates with outcome assessment independently or as part of a grading system. For this to be accomplished, grossing personnel must include sectioning of tumor sites which appear necrotic, hemorrhagic, or edematous, regions typically avoided in most grossing procedures. Most veterinary pathologists will only have microscopic sections to estimate necrosis and these sections are likely to be a small percentage of the entire tumor. Furthermore, in many cases, the gross description will be inadequate unless grossing personnel are instructed to search and report the percent of the entire tumor that appears necrotic. The usual practice of only sampling viable tissue for histological examination might bias the utility of tumor necrosis as an independent or a component parameter in grading systems. Importantly, the size of the tumor, method of sectioning, number of cut surfaces examined grossly and histologically must be documented and at some point, standardized. Based on size of tumor, a recommendation is needed for how many sections should be examined grossly and microscopically. It seems obvious that if pathologist A examines 5 histologic sections and pathologist B only 1 section of a tumor with $5cm^3$ dimensions that the data gathered will not be comparable.

This brings us to the dilemma of how to currently approach reporting tumor necrosis. Given the lack of established guidance, the pathologist can estimate necrosis either visually with glass slides, WSI or measure necrosis with annotation software in WSI. If WSI has drawing software, simply outline the entire tumor circumference (X) as well as

the areas of necrosis (Y), followed by calculation of X/Y = % necrosis in one section (Fig 1 Appendix 4). In the absence of software or if using a microscope then visually estimate with varying magnifications (to confirm areas are indeed necrotic) if the percent necrosis is <50%>. The range of <50%> seems like a wide target and perhaps that is sufficient for estimates. We “assume” prior studies that estimated necrosis in canine tumors only used histology. But how representative the slide(s) are of overall tumor necrosis is unknown and inconsistent sampling of the tumor, purposely avoiding areas of necrosis in tissue selection can skew any determination of percent necrosis in histologic sections. Given the wide target of greater than or less than 50% necrosis, it may be possible to assess this level of necrosis histologically, even with inconsistent sampling. However, determining a 10% threshold of necrosis may prove problematic, as reported in one study indicating that dogs with tumors with > 10% necrosis were 2.7 times more likely to die of tumor related causes.(Kuntz, Dernell et al. 1997)

Future studies can clarify how to determine the percent of tumor necrosis, particularly in larger tumors, and establish a standardized means of gross tissue selection for histologic examination. Various means of assessing for necrosis in histologic sections can be compared and statistically evaluated. Results of standardized assessments for tumor necrosis can be compared to outcomes in univariate and multivariate analysis in concert with other histologic parameters and prognostic utility determined.

References

Chiang, S. and E. Oliva (2013). "Recent developments in uterine mesenchymal neoplasms." Histopathology **62**(1): 124-137.

Coindre, J. M. (2006). "Grading of soft tissue sarcomas: review and update." Arch Pathol Lab Med **130**(10): 1448-1453.

Dobromylskyj, M. J., V. Richards and K. C. Smith (2020). "Prognostic factors and proposed grading system for cutaneous and subcutaneous soft tissue sarcomas in cats, based on a retrospective study." J Feline Med Surg: 1098612X20942393.

Kuntz, C. A., W. S. Dernell, B. E. Powers, C. Devitt, R. C. Straw and S. J. Withrow (1997). "Prognostic factors for surgical treatment of soft-tissue sarcomas in dogs: 75 cases (1986-1996)." J Am Vet Med Assoc **211**(9): 1147-1151.

Laurini, J. A., C. D. Blanke, K. Cooper, G. D. Demetri, R. P. Dematteo, C. D. Fletcher, J. R. Goldblum, T. Krausz, J. Lasota and A. Lazar (2017). "Protocol for the Examination of Specimens From Patients With Gastrointestinal Stromal Tumor (GIST)."

McSporran, K. D. (2009). "Histologic grade predicts recurrence for marginally excised canine subcutaneous soft tissue sarcomas." Vet Pathol **46**(5): 928-933.
Roccabianca, P., F. Y. Schulman, A. G., R. A. Foster, J. L. Scruggs, K. Dittmer and M. Kiupel (2020). "Tumors of Soft Tissue." **Volume 3** 1-306.
Trojani, M., G. Contesso, J. M. Coindre, J. Rouesse, N. B. Bui, A. de Mascarel, J. F. Goussot, M. David, F. Bonichon and C. Lagarde (1984). "Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system." Int J Cancer **33**(1): 37-42.
Vayrynen, S. A., J. P. Vayrynen, K. Klintrup, J. Makela, T. J. Karttunen, A. Tuomisto and M. J. Makinen (2016). "Clinical impact and network of determinants of tumour necrosis in colorectal cancer." Br J Cancer **114**(12): 1334-1342.

Computational pathology (CPATH) (See Appendix 5)

Computational pathology (CPATH) is an umbrella term used to broadly encompass computerized/automated gathering of information on disease in patients.(Abels, Pantanowitz et al. 2019) Although CPATH may use a large variety of information sources (raw medical data: histology images, radiology images, gene sequences, clinical records), Appendix 5 focuses only on *automated image analysis* (AIA) of microscopic tumor images, particularly whole slide images (WSI). When used appropriately, CPATH is an exciting tool which uses microscopic images (input data) and automatically produces output information (counts or scores of patterns, classification of images etc.). It allows the evaluation of large amounts of tumor data on an unprecedented scale, which is likely to reveal novel trends of prognostic importance. As AIA is a relatively new modality of analysis in veterinary pathology with a vast number of relevant methods, this field can be overwhelming with respect to terminology, technical aspects, requirements for developing algorithms, performance validation, and implementation strategies. Therefore, the associated appendix aims to give an overview of relevant terms, general considerations of CPATH methods and specific recommendations for individual prognostic parameters. Generally, two broad categories of AIA approaches are applicable for microscopic tumor prognostication: 1) thresholding-based and 2) advanced data-driven approaches. Thresholding-based algorithms use a set of simple, often programmer-designed image processing steps based on the color information of individual pixels, which are especially useful for scoring immunohistochemical labeling intensity. Data-driven approaches learn to retrieve meaningful patterns from images in order to derive the desired information using artificial intelligence (AI). AI can be used with traditional machine learning methods that require “hand-crafted” (by developer) information about relevant features of the pattern, or more sophisticated deep learning methods that autonomously extract relevant features (decision criteria are unknown to developers, “black box”). Deep learning is generally more powerful than traditional machine learning methods, but necessitates larger amounts of data. For histological images, supervised learning (as opposed to unsupervised learning) is a very useful method that learns by “feedback”

from ground truth labels assigned to the input images. Creating those labels is a very time-consuming task and is prone to several biases (see Appendix 5).

Possible applications of AIA for tumor prognostication are seemingly limitless and various benefits of these approaches have been determined in previous studies.(Stålhammar, Robertson et al. 2018, Steiner, MacDonald et al. 2018, Aubreville, Bertram et al. 2020) Compared to manual assessment by pathologists, algorithms have higher reproducibility, may have higher accuracy, may increase efficiency of repetitive tasks (such as counting of mitotic figures (MF)), and can carefully assess vast amounts of data per case (every image section of multiple WSIs at high magnification) without fatigue. AIA of immunohistochemical labeling intensity was reported to have higher reproducibility and improved prognostic value compared to the manual approach by pathologists for Ki-67 index in human breast cancer,(Stålhammar, Robertson et al. 2018) and membrane-binding biomarkers in human esophageal adenocarcinomas.(Feuchtinger, Stiehler et al. 2015) An automated topometric segmentation mapping algorithm of immunolabeled MF (anti-phospho-histone H3) was used to identify mitotic ‘hot spots’ in canine melanomas and subsequently used image registration in order to assign the same region to H&E stained tumor sections(Puri, Hoover et al. 2019) Deep learning approaches for MF identification in H&E stained tumor sections have been developed for human(Veta, van Diest et al. 2016, Aubreville, Bertram et al. 2020) and canine(Aubreville, Bertram et al. 2020) breast cancer as well as canine mast cell tumors.(Bertram, Aubreville et al. 2019) Deep learning-based algorithms are comparable with pathologists for counting MF (in the same tumor regions)(Veta, van Diest et al. 2016) and outperform pathologists in identifying the ‘hot spot’ regions in WSI.(Aubreville, Bertram et al. 2020) However, correlation of algorithmic MC to patient outcome has not yet been investigated in human and animal tumors. For automated metastasis identification in H&E sections, deep learning-based algorithms can be used for prescreening of images, and a computer-assisted approach has been shown to have higher sensitivity and diagnostic speed compared to the unassisted pathologist.(Steiner, MacDonald et al. 2018) Recent studies on tumors from humans reported that the systems used could even predict if a tumor was benign, carcinoma in

situ, or invasive carcinoma(Aresta, Araujo et al. 2019) as well as predict genetic alterations and gene expression from H&E tumor sections.(Kather, Heij et al. 2020)

Algorithms are not flawless, have multiple sources of error (depending on the algorithmic approach and available dataset) and therefore require very careful validation (see Appendix 5). While thresholding-based approaches have high explainability of algorithmic predictions, data-driven approaches are often considered a “black box” as decision criteria of the algorithms are typically unavailable. Although algorithms are 100% reproducible (same result for the same image using the same model), they may not necessarily cope with variability introduced via biological and pre-analytic factors (tumor type, tissue types present, section preparation and image acquisition). For example, a deep learning-based algorithm for MF may perform poorly on images obtained from a WSI scanner that was not used for the training images.(Aubreville, Bertram et al. 2020) If not part of the training data, algorithms can be compromised by images with very poor tissue or image quality (artifacts, poor fixation etc.). In contrast to thresholding-based approaches, data-driven algorithms are, however, capable of learning a certain degree of image variability and training datasets should include realistic variability that reflects the intended use. Performance evaluation should be done with great care, and data-driven approaches can be assessed by mathematical evaluation (see Appendix 5),(Abels, Pantanowitz et al. 2019) whereas thresholding-based approaches are often only assessed visually by a pathologist.(Aeffner, Wilson et al. 2016) As opposed to pathologists, current algorithms are not capable of modifying their decision based on surrounding tissue (spatial awareness), which can lead to false detections. For example, pathologists are more careful when classifying a MF in an area of necrotic tissue as it may be a MLF but algorithms will not use surrounding tissue and will use the decision criteria programmed to evaluate the candidate structure.

Besides the numerous hurdles in development of AIA algorithms, there are practical issues to consider for bringing AIA into diagnostic workflows. Basic requirements include consistent tissue preparation steps, a digital image acquisition workflow, appropriate IT infrastructure, and sufficient computational power. Increasing implementation of digital microscopy in veterinary laboratories(Bertram and Klopfleisch

2017) will augment access to WSI and facilitate AIA. Nevertheless, acceptance of AIA may be hampered by unfamiliarity, limited research results and poor explainability of machine learning-based algorithms ("black box"). However, there are approaches that can convert the "black box" into a more transparent "glass box" that are likely to have higher acceptance. For example, some algorithms can be implemented as computer-assisted prognosis systems (as opposed to fully computerized decisions) that always require review by a pathologist. These approaches will improve the reliability of the computer assisted prognosis system and allow the reviewing pathologist to retain responsibility in making final decisions with regards to these prognostic parameters. AIA could greatly improve tumor prognostication by providing vast amounts of reproducible and possibly accurate information on the tumor section, but interpretation of the result remains the responsibility of the pathologist.

References:

Abels, E., L. Pantanowitz, F. Aeffner, M. D. Zarella, J. van der Laak, M. M. Bui, V. N. Vemuri, A. V. Parwani, J. Gibbs, E. Agosto-Arroyo, A. H. Beck and C. Kozlowski (2019). "Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association." J Pathol **249**(3): 286-294.

Aeffner, F., K. Wilson, B. Bolon, S. Kanaly, C. R. Mahrt, D. Rudmann, E. Charles and G. D. Young (2016). "Commentary: Roles for Pathologists in a High-throughput Image Analysis Team." Toxicol Pathol **44**(6): 825-834.

Aresta, G., T. Araujo, S. Kwok, S. S. Chennamsetty, M. Safwan, V. Alex, B. Marami, M. Prastawa, M. Chan, M. Donovan, G. Fernandez, J. Zeineh, M. Kohl, C. Walz, F. Ludwig, S. Braunewell, M. Baust, Q. D. Vu, M. N. N. To, E. Kim, J. T. Kwak, S. Galal, V. Sanchez-Freire, N. Brancati, M. Frucci, D. Riccio, Y. Wang, L. Sun, K. Ma, J. Fang, I. Kone, L. Boulmane, A. Campilho, C. Eloy, A. Polonia and P. Aguiar (2019). "BACH: Grand challenge on breast cancer histology images." Med Image Anal **56**: 122-139.

Aubreville, M., C. A. Bertram, T. A. Donovan, C. Marzahl, A. Maier and R. Klopfleisch (2020). "Dogs as Model for Human Breast Cancer: A Completely Annotated Whole Slide Image Dataset." arXiv preprint arXiv:2008.10244.

Aubreville, M., C. A. Bertram, C. Marzahl, C. Gurtner, M. Dettwiler, A. Schmidt, F. Bartenschlager, S. Merz, M. Fragoso, O. Kershaw, R. Klopfleisch and A. Maier (2020). "Deep learning algorithms out-perform veterinary pathologists in detecting the mitotically most active tumor region." Sci Rep **10**: 16447.

Bertram, C. A., M. Aubreville, C. Marzahl, A. Maier and R. Klopfleisch (2019). "A large-scale dataset for mitotic figure assessment on whole slide images of canine cutaneous mast cell tumor." Sci Data **6**(1): 274.

Bertram, C. A. and R. Klopfleisch (2017). "The Pathologist 2.0: An Update on Digital Pathology in Veterinary Medicine." Vet Pathol **54**(5): 756-766.
Feuchtinger, A., T. Stiehler, U. Jutting, G. Marjanovic, B. Luber, R. Langer and A. Walch (2015). "Image analysis of immunohistochemistry is superior to visual scoring as shown for patient outcome of esophageal adenocarcinoma." Histochem Cell Biol **143**(1): 1-9.
Kather, J. N., L. R. Heij, H. I. Grabsch, C. Loeffler, A. Echle, H. S. Muti, J. Krause, J. M. Niehues, K. A. Sommer and P. Bankhead (2020). "Pan-cancer image-based detection of clinically actionable genetic alterations." Nature Cancer **1**(8): 789-799.
Puri, M., S. B. Hoover, S. M. Hewitt, B.-R. Wei, H. A. Adissu, C. H. Halsey, J. Beck, C. Bradley, S. D. Cramer and A. C. Durham (2019). "Automated computational detection, quantitation, and mapping of mitosis in whole-slide images for clinically actionable surgical pathology decision support." Journal of pathology informatics **10**.
Stålhammar, G., S. Robertson, L. Wedlund, M. Lippert, M. Rantalainen, J. Bergh and J. Hartman (2018). "Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer." Histopathology **72**(6): 974-989.
Steiner, D. F., R. MacDonald, Y. Liu, P. Truszkowski, J. D. Hipp, C. Gammage, F. Thng, L. Peng and M. C. Stumpe (2018). "Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer." Am J Surg Pathol **42**(12): 1636-1646.
Veta, M., P. J. van Diest, M. Jiwa, S. Al-Janabi and J. P. Pluim (2016). "Mitosis Counting in Breast Cancer: Object-Level Interobserver Agreement and Comparison to an Automatic Method." PLoS One **11**(8): e0161286.

Margins (see Appendix 6)

Margin assessment is one of the most important histological parameters evaluated in oncology.(Kamstock, Ehrhart et al. 2011, Stromberg and Meuten 2017, Liptak 2020) Patient management decisions often hinge on the results of margin assessment, and clinicians may value margin assessment as highly or more than a diagnosis. Appendix 6 provides the types of data that are required to standardize the reporting of margins for both clinical management and future studies.

Histologic margin evaluation only needs to be reported on tumors where the aim of surgery is to completely remove the neoplasm (achieve local control). Samples where there was no intent to totally excise the tumor, including intralesional or incisional biopsies are for diagnosis only. Certain tumors or the anatomic location of a tumor dictate that excision for local control will be attempted but adjacent structures limit how much margin can safely be taken, and therefore margin assessment by the pathologist should not be requested, e.g. thyroid, anal sac tumors, adrenal glands. Consensus was not reached whether margins should be reported for benign tumors. Appendix 6 provides contrasting philosophies (*Why not? vs Why bother?)* and the recommendation that considerations such as this should be left at the discretion of the pathologist and or their lab as there was no data to support either approach.

For the overall evaluation of surgical margins, the members of the cancer treatment team are the clinician, surgeon, laboratory technologist and pathologist. The responsibilities of each are detailed in Appendix 6. Although terms such as complete, clean, clean but close, narrow, and dirty are ingrained in the clinical and pathology lexicon, practitioners, surgeons, and oncologists should discourage their use and not expect these to be used in pathology reports. Ultimately it is the clinician and/or surgeon that judges if the margin is deemed adequate after consideration of all factors. Observations by the pathologist include 1) relationship of neoplastic cells to the surrounding tissue including presence of a capsule, tissue compression, peripheral invasion and lymphovascular invasion 2) the distance from neoplastic cells to the narrowest or closest inked margin (histologic tumor-free distance (HTFD, Figure 1) and

3) the relationship of neoplastic cells to the boundaries of the *compartment* in which the tumor is located. In many cases, measuring the HTFD alone is not enough to determine the adequacy of surgical margins, yet it is the parameter that is often used to determine 'completeness' of excision by clinician and pathologist. Inking the margin by the clinician/surgeon immediately after tumor excision is required if a HTFD is expected. Although surgical margin identification/inking is routinely performed by most surgeons, this practice is not commonplace in general practice. Details of how to apply ink have been reported (Kamstock, Ehrhart et al. 2011, Appendix 6) and this information should be included in veterinary school curricula. If ink is not present when the sample arrives at the lab this should be noted. Only a small portion of the circumferential surgical margin is evaluated histologically (approximately 0.1- 0.01% of the total margin)(Rapini 1990, Becker 2007, Selmic and Ruple 2020). HTFD should be further studied by comparing different methods of margin analysis (radial, tangential, parallel slicing) with outcome assessments for different tumor types (Milovancev, Townsend et al. 2017, Dores, Milovancev et al. 2018). Until those studies provide comparative data, radial sections are recommended. Regardless of the method used, any margin measured histologically may not accurately represent the tumor and its relationship to the normal surrounding tissue in the patient. It is important to note that HTFD is made on a histopathology specimen that has undergone shrinkage, (ranges reported from 13-50%) and can underestimate the surgically obtained margins by up to 40%.(Miller and Dark 2014, Upchurch, Klocke et al. 2018) Most of the shrinkage occurs immediately after removal and prior to fixation.(Clarke, Banks et al. 2014, Miller and Dark 2014, Upchurch, Klocke et al. 2018) The important margin is between neoplastic cells and "normal tissues" (non-neoplastic) in the patient and this can only be estimated from histopathology. It is recommended to use whole numbers and ranges when reporting HTFD as reporting distances with decimals implies a level of precision and confidence that could be misleading. Furthermore, data is accumulating that the biological behavior of the tumor may be a more important predictor of recurrence than identification of neoplastic cells at a margin. Certainly, this seems to be the case with low-grade canine MCT and STS/STT. Most low grade MCT do not recur even with tumor cells at the margin and approximately one-third of high-grade MCTs will recur when the histologic

margins are free of tumor cells.(Donnelly, Mullin et al. 2015) Similarly, for canine soft tissue tumors/sarcomas, greater than 95% of canine STT do not recur if margins greater than 1mm are free of neoplastic cells and one study reported that when margins are less than 1 mm, three of 41 grade 1 tumors (7%), 14 of 41 grade 2 tumors (34%), and 3 out of 4 grade 3 tumors recurred.(McSporran 2009) The biology of the tumor and the host (immune system, genes) are important factors that influence tumor recurrence and metastases.

Of importance to surgeons is the concept of compartmental boundaries, which are used to plan and perform surgical removal of tumors.(Enneking, Spanier et al. 1980, Kawaguchi, Ahmed et al. 2004) The surgical margins for tumors can be planned differently if the tumor is in a well-delineated anatomic compartment (such as bone, joint, muscle) or is infiltrating poorly demarcated interfascial planes and spaces.(Enneking, Spanier et al. 1980, Kawaguchi, Ahmed et al. 2004) For a well-delineated compartment, it should be reported whether the tumor penetrated the anatomic structure forming the boundary (e.g. periosteum, epimysium or cortical bone). The deep fascia has been described as a barrier of the subcutaneous tissue compartment(Enneking, Spanier et al. 1980) but this structure is not always included in sections of cutaneous and subcutaneous tumor resections. Compartment boundaries may be natural barriers to tumor extension.(Enneking, Spanier et al. 1980, Kawaguchi, Ahmed et al. 2004) It is unclear if these structures function as a true barrier to tumor growth; if they do then it is likely multifactorial and depends upon the aggressiveness of the neoplasm as well as the components of the barrier (eg cortical bone vs adipose tissue; cytokines). Furthermore, what a surgeon vs a pathologist sees as a *fascial plane* may not be the same. If pathologists report the facts of what structures were seen between the tumor and the inked margin, surgeons and oncologists can decide if they are appropriate barriers, and if so, the clinical significance of their presence. Future studies need to clarify if anatomic structures can prevent tumor infiltration, if so how and what the pathologist should identify for skin and subcutaneous “tissue barriers” and fascial planes.(Fulcher, Ludwig et al. 2006) Appendix 6 lists references that describe using CT and MRI for visualizing tissue compartments and assessing the relationship of tumor to adjacent structures, even differentiating aggressive from benign soft tissue

tumors in humans. It is reported that the tunica serosa fascia in peritoneal cavities is a barrier to migration of tumor cells using an in vitro system.(Gao, Ye et al. 2013)

When a delineated anatomic compartment is not obvious, the HTFD is of critical importance. HTFD for lateral and deep margins in samples from skin and subcutis tumors should be reported separately. In a review of surgical biopsy reports of canine cutaneous mast cell tumors, details about the margins and consistency of how histologic margins were reported were generally lacking.(Reagan, Selmic et al. 2018) For example, while some margins were reported in 92% of cases, lateral and deep margins were described separately in 77% of cases, margin direction was only given in 16% of cases and descriptions of the deep margin were only available in 11% of cases.(Reagan, Selmic et al. 2018) The deep margin is difficult for surgeons to visualize intraoperatively. At the end of appendix 6 are considerations for future studies (M1-M4 or R0-RX)(Stromberg and Meuten 2017, Liptak 2020)

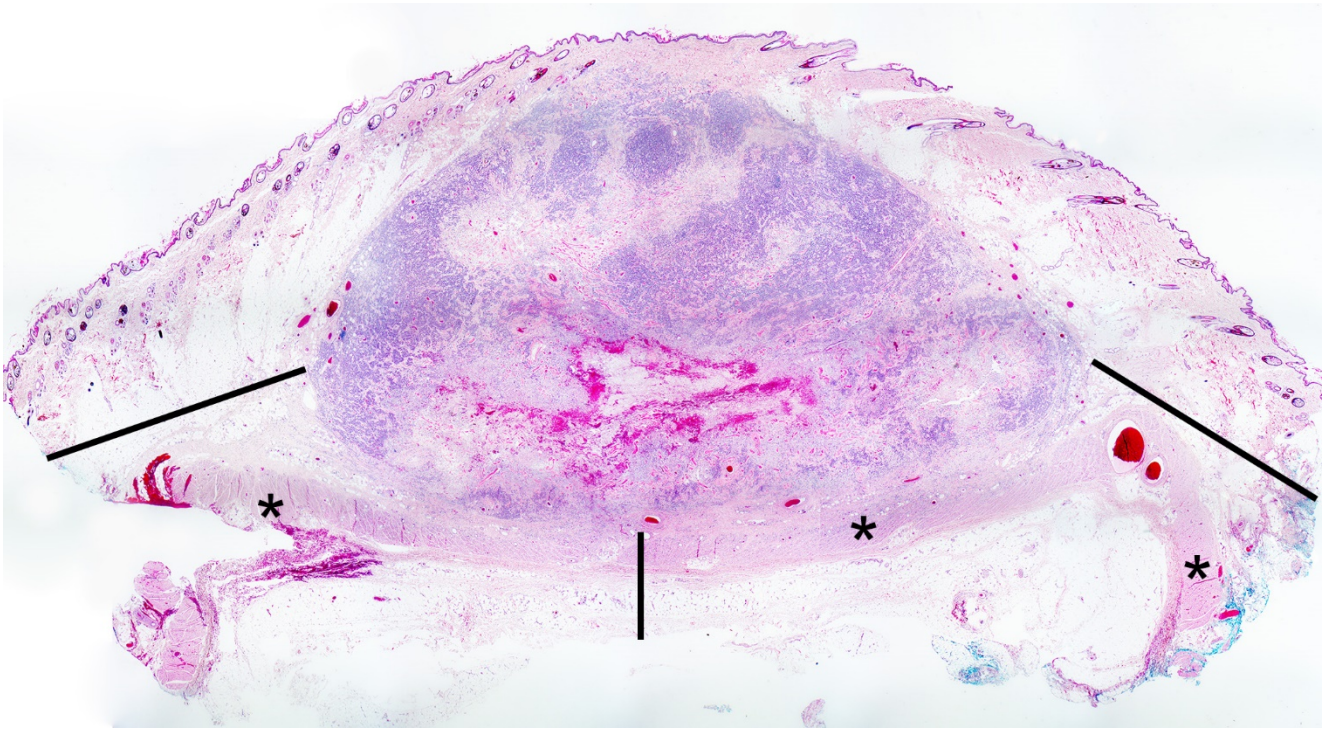


Figure 9: Canine cutaneous mast cell tumor involving the dermis and subcutaneous tissues. The histologic tumor free distance (HTFD) is depicted with horizontal and vertical black lines and can be measured with manual or digital means. Note that ink

can be observed at the lateral (or peripheral) margins, but is not visible at the deep margin. Therefore, the deep margin measurement represents an approximation given the lack of ink. Additional sections into the formalin fixed, paraffin embedded block may resolve this issue. A potential tissue barrier within the subcutaneous tissue is the striated muscle (also called panniculus carnosus or cutaneous trunci in the truncal region, denoted by the asterisks). This muscle is not always visible in histologic sections of cutaneous and subcutaneous tumors; it has variable distribution and continuity in different body regions.(Ahmed, Kulikowska et al. 2019) The subcutaneous fat and loose connective tissue are considered a weak barrier as compared to epimysium, epineurium, or periosteum. The effectiveness of tissue barriers is likely multifactorial and depends upon the aggressiveness of the neoplasm as well as the components of the barrier.

References:

Ahmed, W., M. Kulikowska, T. Ahlmann, L. C. Berg, A. P. Harrison and V. S. Elbrond (2019). "A comparative multi-site and whole-body assessment of fascia in the horse and dog: a detailed histological investigation." J Anat **235**(6): 1065-1077.

Becker, G. D. (2007). "The many faces of surgical margins." Am J Clin Oncol **30**(5): 556-559.

Clarke, B. S., T. A. Banks and L. Findji (2014). "Quantification of tissue shrinkage in canine small intestinal specimens after resection and fixation." Canadian Journal of Veterinary Research **78**(1): 46-49.

Donnelly, L., C. Mullin, J. Balko, M. Goldschmidt, E. Krick, C. Hume, D. C. Brown and K. Sorenmo (2015). "Evaluation of histological grade and histologically tumour-free margins as predictors of local recurrence in completely excised canine mast cell tumours." Vet Comp Oncol **13**(1): 70-76.

Dores, C., M. Milovancev and D. Russell (2018). "Comparison of histologic margin status in low-grade cutaneous and subcutaneous canine mast cell tumours examined by radial and tangential sections." Veterinary and comparative oncology **16**(1): 125-130.

Enneking, W. F., S. S. Spanier and M. A. Goodman (1980). "A system for the surgical staging of musculoskeletal sarcoma." Clinical Orthopaedics and Related Research® **153**: 106-120.

Fulcher, R. P., L. L. Ludwig, P. J. Bergman, S. J. Newman, A. M. Simpson and A. K. Patnaik (2006). "Evaluation of a two-centimeter lateral surgical margin for excision of grade I and grade II cutaneous mast cell tumors in dogs." Journal of the American Veterinary Medical Association **228**(2): 210-215.

Gao, Z., Y. Ye, W. Zhang, D. Shen, Y. Zhong, K. Jiang, X. Yang, M. Yin, B. Liang and L. Tian (2013). "An anatomical, histopathological, and molecular biological function study of the fascias posterior to the interperitoneal colon and its associated mesocolon: their relevance to colonic surgery." Journal of anatomy **223**(2): 123-132.

Kamstock, D., E. Ehrhart, D. Getzy, N. Bacon, K. Rassnick, S. Moroff, S. Liu, R. Straw, C. McKnight and R. L. Amorim (2011). "Recommended guidelines for submission, trimming, margin evaluation, and reporting of tumor biopsy specimens in veterinary surgical pathology." Veterinary Pathology **48**(1): 19-31.

Kawaguchi, N., A. R. Ahmed, S. Matsumoto, J. Manabe and Y. Matsushita (2004). "The concept of curative margin in surgery for bone and soft tissue sarcoma." Clin Orthop Relat Res(419): 165-172.

Liptak, J. M. (2020). "Histologic margins and the residual tumour classification scheme: Is it time to use a validated scheme in human oncology to standardise margin assessment in veterinary oncology?" Veterinary and Comparative Oncology **18**(1): 25-35.

McSporran, K. D. (2009). "Histologic grade predicts recurrence for marginally excised canine subcutaneous soft tissue sarcomas." Vet Pathol **46**(5): 928-933.

Miller, J. L. and M. J. Dark (2014). "Evaluation of the effect of formalin fixation on skin specimens in dogs and cats." PeerJ **2**: e307.

Milovancev, M., K. L. Townsend, E. Gorman, S. Bracha, K. Curran and D. S. Russell (2017). "Shaved margin histopathology and imprint cytology for assessment of excision in canine mast cell tumors and soft tissue sarcomas." Veterinary Surgery **46**(6): 879-885.

Rapini, R. P. (1990). "Comparison of methods for checking surgical margins." J Am Acad Dermatol **23**(2 Pt 1): 288-294.

Reagan, J. K., L. E. Selmic, C. Fallon, E. A. Driskell and L. D. Garrett (2018). "Evaluation of information presented within mast cell tumour histopathology reports in the United States: 2012-2015." Vet Med Sci **4**(3): 252-262.

Selmic, L. E. and A. Ruple (2020). "A systematic review of surgical margins utilized for removal of cutaneous mast cell tumors in dogs." BMC veterinary research **16**(1): 1-6.

Stromberg, P. C. and D. J. Meuten (2017). "Trimming tumors for diagnosis and prognosis." Tumors in domestic animals **5**: 35-51.

Upchurch, D. A., E. E. Klocke and J. N. Henningson (2018). "Amount of skin shrinkage affecting tumor versus grossly normal marginal skin of dogs for cutaneous mast cell tumors excised with curative intent." Am J Vet Res **79**(7): 779-786.

Outcome Assessment (See Appendix 7)

Standardized methods of histologic and outcome assessment parameters for animal tumors are essential if we wish to compare studies and apply the data to clinical cases. The protocols and the appendices herein are an attempt to address this goal and complement existing literature. Outcome assessment of clinical patients is required for determining the predictability of histologically determined prognostic measures (e.g., tumor classification, grade, etc.) Outcome assessment data need to be collected as carefully and accurately as the techniques used to assess tumors.(Webster, Dennis et al. 2011) Some criteria are subjective, clinical, and out of the realm of pathology. Clinicians must carefully select and standardize clinical outcome measures to avoid potential confounders. For example, reporting either disease- or progression-free interval is preferable to median survival time, in order to avoid the confounding effect of timing of euthanasia, which reflects individual biases present within owners and clinicians. Pathologists play a critical role in accurately determining both progression- and disease-free intervals by allowing definitive determination of whether the same tumor recurred and/or metastasized given the appropriate tissue. Obtaining samples for histopathology presents more of a challenge than non-invasive imaging modalities. Although many advances have been made in this realm, microscopic examination of tissues remains the gold-standard. Histopathologic assessment has advantages over cytologic evaluation as more definitive information regarding tumor type can be gained from histopathology. Spindle cell tumors pose a particular problem for cytologic evaluation as it is difficult (impossible) to distinguish reactive fibroplasia (granulation tissue) from sarcomas and difficult to specifically identify tumor type. If we want to know if there is reactive fibroplasia or recurrent perivascular wall tumor at the original excision site, histologic assessment is ideal. However, even with histopathology it is difficult to differentiate these two processes and can be difficult to find tumor cells in re-excision specimens. There is no standard means to evaluate these cases (clinically and histologically) and in at least one study of STTs, presence or absence of tumor in re-

excision specimens did not accurately predict recurrence.(Bacon, Dernell et al. 2007) Future studies could include imaging modalities, and correlate outcome with the presence of normal tissue at the margins of resected samples (eg. no spindle cells of any type). The type of tumor being evaluated will influence the feasibility of visualizing residual tumor cells in margin excisions as well as the concern for local recurrence.

Genetic studies have shown human and animal breed susceptibilities to develop specific tumor types and multiple tumors in the same host. We know that multiple aggressive tumors can be present in the same dog, (Golden Retriever, Rottweiler, Bernese Mountain dogs and others)(Cullen and Breen 2016) . Given this tendency, it is essential to make a definitive diagnosis of tumors in metastatic sites. Combining methodologies is ideal but practical considerations of costs and emotional factors impact study results. Imaging can provide an alternative means to assess for suspected metastatic lesions and can provide useful clinical information for patient management but leaves a gap in outcome assessment studies provided no other confirmatory data is available. Imaging cannot determine whether the tumor suspected of being in the lungs is the same tumor as was excised previously. These methods to identify suspected neoplasia are the methods of choice for clinical settings but not research models. It is important to differentiate the information from a test being used to help treat one patient or predict how a population of animals with the same tumor will react to that tumor type. The latter will be applied to the former when we gather and analyze data carefully. Histopathology remains the gold standard to develop ground truths if the tumor type is the same. We can substitute other methods for histopathology, but the data should be labelled suspected neoplasia/metastases (e.g., as determined by imaging or physical exam) but not confirmed unless histopathology is used. In the future, molecular testing of suspected tumor tissue may be superior to histopathology.

Other appendices have detailed how to assess parameters used to evaluate a tumor, recurrence, margins and metastases. In order to use morphologic diagnoses, margins, LVI, MC, lymph node status, or CPATH to predict tumor behavior and/or to select treatment options, oncologists must acquire sufficient outcome assessment information to allow interpretation of tumor parameters. Knowing actual survival times of geriatric

pets or including pets in which no treatments were performed provides control groups to which treatments and outcomes can be compared. Determining the least invasive means to characterize tumor behavior is ideal but cannot be accomplished without adequate outcome assessment studies. Appendix 7 utilizes and expands upon published guidelines for conduct and evaluation of prognostic studies(Webster, Dennis et al. 2011) and for response assessment in canine solid tumors,(Nguyen, Thamm et al. 2015) citing specific information gained from studies of canine soft tissue sarcoma and canine mast cell tumor.

Standardized criteria, such as RECIST and RECIST 1.1(Therasse, Arbuck et al. 2000, Schwartz, Seymour et al. 2016) should be used to document the patient's response to treatment and progressive disease. The RECIST 1.1 criteria have been robustly evaluated for use in human clinical trials and can be easily adapted to the evaluation of veterinary patients. Pathologists, oncologists, surgeons, clinicians and students should be familiar with the terms explained in these manuscripts which indicate response to treatment and include Complete remission (CR), Partial response (PR), Progressive disease (PD), Stable disease (SD) and Not evaluable (NE).(Nguyen, Thamm et al. 2015) Documented progression is needed in the cases of questionable lesions, or a minimum size is required to determine whether neoplastic disease is present within a lymph node. Additionally, there may be specific anatomical locations evaluated depending on the tumor type. For example, prostate cancer may favor bone metastases, pulmonary carcinoma in cats requires assessment of all digits, and hemangiosarcoma is the most common metastatic tumor to the brain of dogs. Ideally, imaging will be used in concert with biopsy or autopsy in order to confirm recurrence and metastasis with the utmost accuracy.

Metastasis should be subdivided into confirmed and suspected. Metastases determined by imaging only should be labelled suspected. Histopathology is required to confirm metastases are present and are of the same tumor type. The preferred methodology of evaluation in humans, the CT scan, should be used if possible as it avoids some of the technical problems associated with the use of radiographs, whereas ultrasound is not an acceptable method of assessing disease state(Nguyen, Thamm et al. 2015) The use

of functional imaging (PET scans) is increasingly common to better determine sites of disease; however, it cannot be used for measuring purposes. Following these standardized criteria will ensure that studies can be reproduced and compared between institutions, resulting in more useful correlates of clinical data to prognostic information, and ensuring progress in veterinary oncologic pathology.

Euthanasia is a reality of veterinary medicine, and oncology studies that use pets must carefully evaluate how decisions to euthanize influenced survival times. Reported patient survival times are impacted by euthanasia which may be elected due to perceived pet value, owner income, primary vs referral centers or other factors which do not reflect tumor behavior. When patients are euthanized, clinicians should determine and/or record the cause of death with as much accuracy as possible. If euthanasia is due to an unrelated disease process, this must be noted. If euthanasia is caused by the neoplasm being studied, and cachexia is present, then histologic confirmation of the extent of the neoplastic disease helps verify clinical observations and reliability of study conclusions. Oncology studies no longer include results of autopsy, the perceived value of which seems to have hit a nadir. Permission to perform autopsies should be pursued as autopsy greatly increases the confidence in results from the case. Studies should set a goal of autopsies on at least 20% of the cases.

References

Bacon, N. J., W. S. Dernell, N. Ehrhart, B. E. Powers and S. J. Withrow (2007). "Evaluation of primary re-excision after recent inadequate resection of soft tissue sarcomas in dogs: 41 cases (1999-2004)." J Am Vet Med Assoc **230**(4): 548-554.

Cullen, J. M. and M. Breen (2016). "An overview of molecular cancer pathogenesis, prognosis, and diagnosis." Tumors in Domestic Animals: 1-26.

Nguyen, S. M., D. H. Thamm, D. M. Vail and C. A. London (2015). "Response evaluation criteria for solid tumours in dogs (v1.0): a Veterinary Cooperative Oncology Group (VCOG) consensus document." Vet Comp Oncol **13**(3): 176-183.

Schwartz, L. H., L. Seymour, S. Litiere, R. Ford, S. Gwyther, S. Mandrekar, L. Shankar, J. Bogaerts, A. Chen, J. Dancey, W. Hayes, F. S. Hodi, O. S. Hoekstra, E. P. Huang, N. Lin, Y. Liu, P. Therasse, J. D. Wolchok and E. de Vries (2016). "RECIST 1.1 - Standardisation and disease-specific adaptations: Perspectives from the RECIST Working Group." Eur J Cancer **62**: 138-145.

Therasse, P., S. G. Arbuck, E. A. Eisenhauer, J. Wanders, R. S. Kaplan, L. Rubinstein, J. Verweij, M. Van Glabbeke, A. T. van Oosterom, M. C. Christian and S. G. Gwyther (2000). "New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada." J Natl Cancer Inst **92**(3): 205-216.

Webster, J., M. M. Dennis, N. Dervisis, J. Heller, N. J. Bacon, P. J. Bergman, D. Bienzle, G. Cassali, M. Castagnaro and J. Cullen (2011). "Recommended guidelines for the conduct and evaluation of prognostic studies in veterinary oncology." Veterinary pathology **48**(1): 7-18.

Synoptic Reporting in Veterinary Medicine (See Appendix 8)

Synoptic reporting (as opposed to the traditional narrative reporting) is a method for reporting specific pieces of prognostically-relevant data in a discrete format in pathology reports (Renshaw, Mena-Allauca et al. 2018). In human medicine, these have progressed from individual efforts (Markel and Hirsch 1991) to being mandated by the College of American Pathologists (CAP) for accreditation (College of American Pathologists 2020). In general, a synoptic pathology report consists of data elements and responses (see Figures S1, S2 Appendix 8, supplemental), which may be either required or optional. For CAP purposes, the report must have all required core components reported, all conditional core components reported when applicable, must be listed with the element next to its associated response, and all elements and responses must be on separate lines and in one place in a report. Synoptic reporting has been shown to make pathology reports more readable to clinicians and patients (Renshaw, Mena-Allauca et al. 2018), as well as making reports more likely to include all data elements needed (Karim, van den Berg et al. 2008, Kang, Devine et al. 2009, Srigley, McGowan et al. 2009, Messenger, McLeod et al. 2011) To develop an effective synoptic report typically requires the efforts of pathologists and clinicians, who develop the checklist of required and recommended items after reviewing the relevant literature (Chamberlain, Wenckebach et al. 2000). Currently, there are two main groups producing templates in human medicine, CAP and the International Collaboration on Cancer Reporting (ICCR). Both require a committee of pathologists, oncologists, and other interested representatives (e.g., World Health Organization working groups, etc.) to develop a new protocol.

A number of studies have found that synoptic reporting produces reports that are more likely to contain all significant pieces of information than narrative reports. For pancreatic tumors, 100% of synoptic reports had information about small vessel and perineural invasion, compared to 66% and 84% of narrative reports, respectively (Gill, Johns et al. 2009). In addition, the stage could be determined in 100% of synoptic

reports compared to 56% of narrative reports. In a comparison of melanoma reports, mitotic count, histologic subtype, predominant cell type, vascular and lymphatic invasion, neurotropism, desmoplasia, and distance to the nearest margin were all reported significantly more frequently in synoptic reports than narrative reports, both at the teaching institution responsible for the study and the outside reports sent in to the teaching institution for a second opinion(Karim, van den Berg et al. 2008).

While full implementation of standardized reporting would allow for easy automated data collection(Ellis and Srigley 2016), even simple implementations of synoptic reporting can allow for significant automated information extraction. For example, if all deep margins are listed as “DEEP MARGIN: <xx>mm” on a line by itself, it is comparatively easy to extract all margins from reports using standard text search and manipulation tools (e.g., grep, cut, etc.). Not only can this improve retrospective studies, but can also provide valuable clinical information, as extracted information can be compared between services, clinicians, and other variables to determine if these influence patient outcomes.

From the beginning of synoptic reporting, clinicians have reported increased satisfaction with synoptic vs. narrative reports (Markel and Hirsch 1991). A study of treating physicians and pathologists in Canada found that both groups found synoptic reports easier to find information in, facilitate a consistent approach to interpretation of diagnostic and prognostic factors, and provide higher overall satisfaction (Lankshear, Srigley et al. 2013). While pathologists felt that reports took approximately 25-50% longer to complete, treating physicians did not notice a difference in the length of time it took pathology reports to be completed.

The major problem in veterinary medicine is a lack of knowledge about factors involved in prognosis. As discussed in the other appendices in this document, there is little standardization of methods used in determining prognostic factors. There are also no standards for terminology, such as immunohistochemical findings (e.g., “positive” vs. “immunoreactive” vs. “present”), which hinders design of standardized reports. Another issue for many pathologists, particularly in academia, is the effect switching to synoptic reports would have on resident training. Given the necessity of writing descriptions for

boards and the lack of universal adoption of synoptic reporting, residents still require significant experience in writing narrative reports. This can be mitigated by requiring narrative reports in other resident educational settings (such as rounds) to provide practice in writing narrative reports for neoplasms.

Many pathologists are concerned about increased time to finish reports with synoptic reporting, including physicians (Lankshear, Srigley et al. 2013); however,when synoptic reports have been implemented many of these concerns have been deemed technology related rather than issues with the reporting format. As with many new processes, we assume that once the pathologists become familiarized with the new format, there will be a decrease in time to write these types of reports. A standardized formatted template will be created and added to the website we propose. In veterinary medicine, no current laboratory information management system (LIMS) can use synoptic reporting, which may seem like an obstacle to implementation of synoptic reporting. However, any word processor can be used to implement synoptic reporting without specialized software(Ellis and Srigley 2016); all that is required is to type the data element, a separator (such as TAB), and the response. Templates can be saved containing required and optional data elements, making it easier for pathologists to fill out reports quickly. These can then be copied and pasted into any LIMS or word processor for subsequent reporting.

Finally, another major obstacle to implementation of synoptic reporting is a lack of awareness of synoptic reporting and its benefits in veterinary medicine. Establishing working groups with pathologists and oncologists to develop guidelines for specific neoplasms would help promote awareness and develop reporting checklists that would benefit both pathologists and treating clinicians.

The next step beyond synoptic reporting is standardized reporting, that is, having a standardized, specific set of responses for each required question(Srigley, McGowan et al. 2009). Ultimately, this can lead to automated staging and grading, as well as improving data harvesting for future research and clinical applications. The addition of free text fields associated with standardized options would allow for customization of reports while retaining standardization for further applications.

References:

Chamberlain, D. W., G. F. Wenckebach, F. Alexander, R. S. Fraser, A. Kolin and T. Newman (2000). "Pathological examination and the reporting of lung cancer specimens." Clin Lung Cancer 1(4): 261-268.

Ellis, D. W. and J. Srigley (2016). "Does standardised structured reporting contribute to quality in diagnostic pathology? The importance of evidence-based datasets." Virchows Arch 468(1): 51-59.

Gill, A. J., A. L. Johns, R. Eckstein, J. S. Samra, A. Kaufman, D. K. Chang, N. D. Merrett, P. H. Cosman, R. C. Smith, A. V. Biankin, J. G. Kench and N. New South Wales Pancreatic Cancer (2009). "Synoptic reporting improves histopathological assessment of pancreatic resection specimens." Pathology 41(2): 161-167.

Karim, R. Z., K. S. van den Berg, M. H. Colman, S. W. McCarthy, J. F. Thompson and R. A. Scolyer (2008). "The advantage of using a synoptic pathology report format for cutaneous melanoma." Histopathology 52(2): 130-138.

Lankshear, S., J. Srigley, T. McGowan, M. Yurcan and C. Sawka (2013). "Standardized synoptic cancer pathology reports - so what and who cares? A population-based satisfaction survey of 970 pathologists, surgeons, and oncologists." Arch Pathol Lab Med 137(11): 1599-1602.

Markel, S. F. and S. D. Hirsch (1991). "Synoptic surgical pathology reporting." Hum Pathol 22(8): 807-810.

Messenger, D. E., R. S. McLeod and R. Kirsch (2011). "What impact has the introduction of a synoptic report for rectal cancer had on reporting outcomes for specialist gastrointestinal and nongastrointestinal pathologists?" Arch Pathol Lab Med 135(11): 1471-1475.

Renshaw, A. A., M. Mena-Allauca, E. W. Gould and S. J. Sirintrapun (2018). "Synoptic Reporting: Evidence-Based Review and Future Directions." JCO Clin Cancer Inform 2: 1-9.

Srigley, J. R., T. McGowan, A. Maclean, M. Raby, J. Ross, S. Kramer and C. Sawka (2009). "Standardized synoptic cancer pathology reporting: a population-based approach." J Surg Oncol 99(8): 517-524.

Synoptic Report

MASS SIZE:	3 cm x 2 cm x 2 cm
	Gross measurement
BIOPSY TYPE:	Excisional
LOCATION:	Forelimb proximal to elbow
ASSESSMENT METHOD:	Manual light microscopy with glass slides
HISTOLOGIC TYPE:	Nerve sheath tumor
DEEPEST LAYER INFILTRATED:	Dermis
	Via histology
DIFFERENTIATION SCORE:	2
MITOTIC COUNT:	12 per 2.37mm^2
NECROSIS:	11-50%
TOTAL SCORE:	5
HISTOLOGIC GRADE:	2
	Kuntz system
LYMPHOVASCULAR INVASION:	None
METASTASIS:	None
MARGINS INKED:	By laboratory
MARGIN TYPE:	Radial
DEEP MARGIN:	Complete
DEEP MARGIN HTFD:	3 mm
LATERAL MARGIN:	Complete
LATERAL MARGIN HTFD:	6 mm

Narrative Report

In one transverse and two longitudinal sections (from a 3 x 2 x 2cm mass from the left forelimb, per submitter), the dermis is disrupted by a highly cellular, infiltrative, unencapsulated mass. The mass is composed of cells forming bundles and whorls surrounding empty capillaries. The cells have indistinct borders and eosinophilic cytoplasm. The nuclei are medium to large and fusiform, with finely stippled chromatin. Mitoses average 12 per ten 400x fields (2.37mm^2). The central 30% of the mass is necrotic. The mass is separated from the deep and lateral sample margins by 3mm and 6mm, respectively.

DIAGNOSIS:

Nerve sheath tumor, grade II, left forelimb

Figure 10: Comparison of synoptic and narrative reports. The same information in each report is in the same color.

Skin and Subcutaneous Soft Tissue Tumors (STT/STS) (See Protocol 1)

This protocol is intended for use with soft tissue tumors arising in the skin and subcutaneous tissues which are predominantly of mesenchymal tissue origin and which are commonly referred to as soft tissue sarcomas (STS).(Bostock and Dye 1980, Kuntz, Dernell et al. 1997, McSporran 2009, Roccabianca, Schulman et al. 2020) Modifying a name generally meets with resistance and lack of unanimity. The term sarcoma suggests the group of neoplasms are aggressive (malignant), however present outcome assessment data does not indicate that is the case.(Bostock and Dye 1980, Kuntz, Dernell et al. 1997, McSporran 2009, Roccabianca, Schulman et al. 2020) Thus, it is proposed to remove sarcoma from the acronym. These neoplasms are predominantly mesenchymal, however, a subset (namely nerve sheath tumors) are not solely derived from the mesoderm, therefore, soft tissue mesenchymal tumor is not entirely accurate. These neoplasms can be accurately encompassed by the term soft tissue tumors (STT) (which is admittedly vague), however, ensures that more users of this term will be satisfied. The purpose of this protocol is to provide standards for accruing data so that, over time, large data sets with comparable information can be evaluated to enable meaningful conclusions and accurate prognostic information.

The term STT/STS encompasses a wide range of benign and malignant tumor types in humans (Byerly S, Chopra S, Nassif NA et al, 2016) The different types are much more limited in animals and, although the veterinary terminology and various grading schemes have, in many instances, been borrowed from the human literature, the types of neoplasms which commonly comprise soft tissue tumors in humans are very different from the tumor types typically encountered in animals. This is exemplified by liposarcomas, which are common in humans and rare in dogs, and perivascular wall tumors (PWT), very common in dogs, are rare in humans. Furthermore, STS in humans have extensive molecular profiles to help subtype them, which is not established for canine tumors. The common denominators between species appears to be an origin in non-epithelial, extraskeletal soft tissues exclusive of hematopoietic system. (Bostock and Dye 1980, Trojani, Contesso et al. 1984, Kuntz, Dernell et al. 1997, Coindre 2006, McSporran 2009, Dennis, McSporran et al. 2011, Roccabianca, Schulman et al. 2020)

This protocol is intended for use with the following types of tumors: Perivascular wall tumors (PWT), nerve sheath tumors (NST), fibrosarcoma, myxosarcoma, leiomyosarcoma, liposarcoma, rhabdomyosarcoma or unclassified spindle cell tumor/sarcoma arising in the dermis or subcutis. PWT and NST are the most common types of STT/STS and their biological behavior is primarily indolent.(Roccabianca, Schulman et al. 2020)The effect of grouping of disparate tumors within the same grading scheme needs to be compared to grading tumors segmented into specific histological diagnoses so that important predictive parameters may be determined.

The current scheme used for grading dog STT/STS is patterned after Trojani's grading of human STS.(Trojani, Contesso et al. 1984) Unlike the human grading scheme, however, the studies of dog STT/STS only evaluated three histological features. Some criteria, such as determination of the percentage of necrosis via gross and/or histological criteria, are poorly defined in the human literature and were not clarified in the veterinary manuscripts.(Bostock and Dye 1980, Kuntz, Dernell et al. 1997, McSporran 2009) Percent necrosis for human tumors was determined by estimating the amount seen grossly and histologically (see Appendix 4).(Trojani, Contesso et al. 1984, Coindre 2006, Rubin, Cooper et al. 2010, Nguyen, Thamm et al. 2015) There are a number of distinctions between the grading systems used for human tumors and how the they are applied to dogs, which have not been addressed in the canine papers;(Bostock and Dye 1980, Trojani, Contesso et al. 1984, Kuntz, Dernell et al. 1997, Coindre 2006, McSporran 2009) in particular, the need to determine histological tumor type and confirmation of the diagnosis of sarcoma *prior* to applying the human grading systems. Four additional histological features evaluated by Trojani but not found useful for human tumors were not assessed in the dog STT/STS grading studies. Our existing scheme needs to be broadened to determine if parameters originally rejected for human STS may, in fact, be predictive in dogs. The methods described to assign scores for necrosis, MC and differentiation for canine tumors are not detailed enough that others can replicate them, and the number of dogs reported with high-grade STT/STS that have outcome assessments is small. These studies need to be repeated with additional parameters evaluated, more detailed description of methods and greater case numbers paired with standardized outcome assessments. The

protocol in this appendix provides details of the histological findings that should be noted in STT/STS which will enable more thorough assessment of these tumors and should provide a database for performance of studies and validation of grading schemes.

For any proposed veterinary tumor grading system, the tumor type should be designated as precisely as possible and the criteria used to designate that diagnosis be provided (H&E, IHC etc). Each graded element must be clearly defined. For instance, the means to assess percent necrosis (gross, histology, both; Appendix 4) must be clarified if this is an element of a grading system and others are expected to duplicate the method.(Kuntz, Dernell et al. 1997) Histologic classification of some types of STT/STS is difficult. A particular conundrum is differentiating PWT from NST. Histological features characteristic of PWT and NST have been described, but there is overlap of histological patterns found in these two tumor types(Avallone, Helmbold et al. 2007, Suzuki, Uchida et al. 2014, Loures, Conceição et al. 2019, Vučićević, Marinković et al. 2019, Avallone, Stefanello et al. 2020, Roccabianca, Schulman et al. 2020) which can complicate definitive diagnosis in routinely stained sections. How specific can, or should our diagnoses be from HE slides and how does this influence differentiation scores used to grade these tumors? Examples: Should PWT be subtyped, and similarly as NST is not just one tumor, should neurofibroma, Schwannoma and malignant NST be identified? Classification of some tumors, including some cases of PWT, may require IHC or other ancillary tests. In veterinary medicine, the costs for these tests are incurred by owners and, if the tests are declined, it is unreasonable to expect a precise classification of some of these tumors with H&E. These practical factors influence our diagnoses and grading systems.

Present canine studies have not determined if identifying tumor type is predictive of tumor behavior. Until we use a grading system for specific tumor types as well as for the entire group of STT/STS, we will not know which approach is more predictive. A grading scheme that can be applied to any tumor within the STT/STS group is easier to apply then requiring identification of the specific tumor type before grading, particularly in instances in which a definitive diagnosis cannot be made with evaluation of routinely

stained tissue sections. However, future studies should validate if this is “best practice”. On the one hand, identifying the precise tumor type may have prognostic information unrelated to a specific grade and, on the other hand, perhaps all tumors within either the entire STT/STS group or within specified subsets of the group (for instance PWT/NST) may behave according to assigned grades. For instance, group PWT and NST together, based on H&E histologic morphology and determine outcome and determine if there is a difference in outcome assessment if these two tumor types are evaluated separately. Identification of these two tumor types may involve IHC or electron microsocopy. If the biological behavior of these two tumors was such that distinguishing them at the H&E level was not needed that would have practical use for a diagnostic pathologist and oncologist. The only means to determine the prognostic utility of grouping or separating tumor types within the STT/STS category is to perform studies which evaluate outcomes related to the STT/STS group as a whole and ALSO evaluate outcomes in relation to specific histologic type of tumors. Studies must have sufficient numbers of animals within each tumor grade to generate statistically significant findings. This latter issue will be a problem for uncommon tumors, such as liposarcoma, for which it may be problematic to find enough high-grade tumors with accurate outcome assessments, but using criteria in which two tumor types (ie, PWT and NST) comprise more than 80% of the cases to predict how uncommon tumors behave needs to be validated.

Future considerations should address existing and new grading systems for STT/STS (see protocol 1). The present grading system should be followed with methods described in sufficient detail to permit other investigators to duplicate the methods and the scoring systems. Consideration should be given to assessment of weighted scores for parameters, such as differentiation or mitotic count, in determining grade and correlation with outcome assessment. Additional histological features should be evaluated for their prognostic utility, for instance, tumor cellularity, presence of atypical nuclei or multinucleated giant cells and presence of lymphovascular invasion (see Appendix 3). The benefit of applying a new, better-detailed scoring system for histological differentiation should be assessed as this is the most subjective parameter

in human tumors and likely canine tumors. The use of a defined area in mm^2 should be applied to all parameters enumerated in a grading system. New grading systems should be compared to older systems, and there must be sufficient numbers of animals in each tumor grade to enable interpretation of results. Studies should be initiated to assess the criteria for diagnosis of NST and PWT and the reproducibility of the criteria. Finally, the use of computational pathology and molecular profiling should be explored in determining grades and outcomes of STT/STS.

References (A complete list of references is available in protocol 1)

Avallone, G., P. Helmbold, M. Caniatti, D. Stefanello, R. C. Nayak and P. Roccabianca (2007). "The spectrum of canine cutaneous perivascular wall tumors: morphologic, phenotypic and clinical characterization." Vet Pathol **44**(5): 607-620.

Avallone, G., D. Stefanello, R. Ferrari and P. Roccabianca (2020). "The controversial histologic classification of canine subcutaneous whorling tumours: The path to perivascular wall tumours." Vet Comp Oncol **18**(1): 3-8.

Bostock, D. E. and M. T. Dye (1980). "Prognosis after surgical excision of canine fibrous connective tissue sarcomas." Vet Pathol **17**(5): 581-588.

Byerly, S., S. Chopra, N. A. Nassif, P. Chen, S. F. Sener, B. L. Eisenberg and W. W. Tseng (2016). "The role of margins in extremity soft tissue sarcoma." J Surg Oncol **113**(3): 333-338.

Coindre, J. M. (2006). "Grading of soft tissue sarcomas: review and update." Arch Pathol Lab Med **130**(10): 1448-1453.

Dennis, M., K. McSporran, N. Bacon, F. Schulman, R. Foster and B. Powers (2011). "Prognostic factors for cutaneous and subcutaneous soft tissue sarcomas in dogs." Veterinary Pathology **48**(1): 73-84.

Kuntz, C. A., W. S. Dernell, B. E. Powers, C. Devitt, R. C. Straw and S. J. Withrow (1997). "Prognostic factors for surgical treatment of soft-tissue sarcomas in dogs: 75 cases (1986-1996)." J Am Vet Med Assoc **211**(9): 1147-1151.

Loures, F., L. Conceição, R. Lauffer-Amorim, J. Nóbrega, E. Costa, R. Torres Neto, J. Clemente, M. Vilória and J. Silva (2019). "Histopathology and immunohistochemistry of peripheral neural sheath tumor and perivascular wall tumor in dog." Arquivo Brasileiro de Medicina Veterinária e Zootecnia **71**(4): 1100-1106.

McSporran, K. D. (2009). "Histologic grade predicts recurrence for marginally excised canine subcutaneous soft tissue sarcomas." Vet Pathol **46**(5): 928-933.

Nguyen, S. M., D. H. Thamm, D. M. Vail and C. A. London (2015). "Response evaluation criteria for solid tumours in dogs (v1.0): a Veterinary Cooperative Oncology Group (VCOG) consensus document." Vet Comp Oncol **13**(3): 176-183.

Roccabianca, P., F. Y. Schulman, G. Avallone, R.A. Foster, J.L. Scruggs, K. Dittmer, and M. Kiupel (2020). "Tumors of Soft Tissue." **Volume 3**(Davis Thompson DVM Foundation): 1-306.

Rubin, B. P., K. Cooper, C. D. Fletcher, A. L. Folpe, F. H. Gannon, J. L. Hunt, A. J. Lazar, A. G. Montag, T. D. Peabody and R. E. Pollock (2010). "Protocol for the examination of specimens from patients with tumors of soft tissue." Archives of pathology & laboratory medicine **134**(4): e31-e39.

Suzuki, S., K. Uchida and H. Nakayama (2014). "The effects of tumor location on diagnostic criteria for canine malignant peripheral nerve sheath tumors (MPNSTs) and the markers for distinction between canine MPNSTs and canine perivascular wall tumors." Vet Pathol **51**(4): 722-736.

Trojani, M., G. Contesso, J. M. Coindre, J. Rouesse, N. B. Bui, A. de Mascarel, J. F. Goussot, M. David, F. Bonichon and C. Lagarde (1984). "Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system." Int J Cancer **33**(1): 37-42.

Vučićević, I., D. Marinković, V. Kukolj, S. Nešić, M. Aničić, B. Đurđević and S. Aleksić-Kovačević (2019). "Immunohistochemical distinguishing between canine peripheral nerve sheath tumors and perivascular wall tumors." Acta Veterinaria **69**(3): 290-299.