UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI MEDICINA VETERINARIA

PH.D. COURSE IN VETERINARY AND ANIMAL SCIENCES

(CLASS XXXIII)

Ph.D. THESIS

PATHOLOGY OF ROUND CELL TUMORS AND SARCOMAS: A DIAGNOSTIC AND CLINICALLY-ORIENTED APPROACH, WITH A FOCUS ON BASIC INVESTIGATION

(VET/03)

Ph.D. Candidate

Dr. Matteo Gambini
(Matricola: R12036)

Tutor: **Prof. Chiara Giudice**

Ph.D. Course Coordinator: **Prof. Valeria Grieco**

ACADEMIC YEAR 2019/2020

Alle Dottoresse Sala, Di Cesare, Diotti e Buzzi,

senza le quali questo percorso non sarebbe giunto alla sua conclusione.

INDEX

ABSTRACT

Neoplasms represent a constantly increasing threat for companion animals, requiring fast and reliable diagnostic techniques. In this context, cytology is a diagnostic tool widely applied to preliminarly investigate the nature of lesions observed in daily veterinary practice. Nevertheless, considered the limitations of cytology as well as the key role that a cytological diagnosis might cover for the clinical approach to each patient, estimation of the reliability of this technique represents a fundamental step. With this in mind, diagnostic accuracy studies can provide a proof of reliability of cytological results, when the latter are compared to the histopathological gold standard. With these premises, the main aim of the current Thesis was to evaluate the diagnostic accuracy of cytology applied to different round cell tumors and sarcomas currently representing a serious threat for the canine and feline species. **Chapter 1** describes a study investigating the diagnostic accuracy of cytology in the evaluation of canine splenic neoplasm. To the best of our knowledge, our work was the first study conjunctively reporting overall accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of cytology in the diagnosis of these lesions. Diagnostic accuracy indexes identified limitations of negative cytological results in excluding a dog to be truly free from neoplasia; however, high specificity and positive predictive value still highlighted cytology as a valuable tool in the diagnostic approach to splenic neoplasms. In **Chapter 2** is described a study investigating interobserver agreement and diagnostic accuracy of cytology in the immunophenotype prediction of feline nodal lymphoma. Our results revealed a low inter-observer agreement and a low diagnostic accuracy in immunophenotype prediction, thus highlighting potential marked differences among laboratories and even among different cytologists within the same laboratory, and consequently the mandatory need of histopathology and immunohistochemistry for a correct interpretation of feline nodal lymphoma immunophenotype. **Chapter 3** describes two studies focusing on the application of cytology in the evaluation of nodal metastasis of canine mast cell tumor (MCT). The preliminary investigations described in **Section 1** focused on the quantification of mast cells in cytological nodal samples obtained

from both non oncological dogs and MCT-bearing dogs, revealing that mast cell (MC) quantification in lymph node (LN) cytological samples obtained from the latter might be useful to determine the nodal metastatic status. Our findings further suggested that neither the sampling technique applied to collect cytological samples nor the quantification method applied to estimate the number of MCs, influence the number of nodal MCs observed. **Section 2** describes a study investigating interobserver agreement and diagnostic accuracy of cytology in the evaluation of the metastatic status of lymph nodes obtained from MCT-bearing dogs. Specifically, the study investigated the diagnostic performance of the cytological interpretative system currently available in literature as well as that of 2 amendments of the latter (AM1 and AM1.2 system). Our results revealed that the AM1.2 system, which include MC quantification besides the cytological criteria reported in the previously published cytological interpretative system, could represent a valid alternative to the latter, being characterized by an almost overlapping interrater agreement, and sensibly higher accuracy and sensitivity without substantial changes of the other diagnostic accuracy indexes. Besides studies focusing on the diagnostic accuracy of cytology, the **Addendum** briefly describes the results of 2 research projects investigating viral oncolysis in a cell culture model of canine histiocytic sarcoma, performed by the Ph.D. candidate during his externship at the University of Veterinary Medicine of Hannover (TiHo Hannover, Germany).

PREFACE AND AIMS

Neoplasms represent a constantly increasing threat for companion animals, similarly to what happens in human beings.[1,21,186,212]

Neoplasm can be divided according to histogenesis (i.e. the embrional layer the neoplastic cells originated from) into ectodermal, mesodermal, and endodermal, with the second including all mesenchimal tumors.[146] According to biological behavior, i.e. the tendency to infiltrate surrounding tissue and/or to metastatize, neoplasms can be further subdivided into benign and malignant, with the terms “sarcoma” generally referring to malignant mesenchymal tumors.[146] Finally, neoplasms can be divided according to cell morphology and tendency to aggregate into epithelial, mesenchymal, and round cell tumors, which approximately correspond to neoplasms of epithelial, mesenchymal, and hematopoietic origin, although exceptions are reported.[164] In practice, however, the aformentioned classifications are intimately intermingled. As few examples, it would be sufficient to think about lymphomas or histiocytic sarcomas, which are simultaneously classified as “sarcomas” on the basis of the histogenesis, and “round cell tumors” as per cell morphology.[146]

The present Ph.D. Thesis is specifically concerned with the study of round cells tumors and sarcomas. More specifically, while a detailed review of classification, morphological features, and biological behavior of these neoplasms falls beyond the aims of the current Thesis, the work of the Ph.D. candidate was focused on the evaluation of diagnostic accuracy of cytology applied to different round cell tumors and sarcomas currently representing a serious threat for the canine and feline species.

Cytology is a diagnostic tool widely diffuse in daily veterinary practice whose main aim in diagnostic clinical settings is to preliminarily identify the nature of a lesion (neoplastic versus reactive), and possibly to distinguish benign from malignant tumors.[3,23,43,128,189,190,227,229] Cytology is indeed a diagnostic technique able to provide useful information to guide clinical decisions in a minimally invasive and cost effective way, in association with the tremendous advantage of a short turnaround time (i.e.

the time comprised between sampling and diagnosis).[23,32,189,190,229] Furthermore, cytological sampling can be virtually applied *ex vivo* to almost all organs and tissues, being generally well-tolerated by patients and requiring none or minimal pharmacological restraint.[23,122,190] Additionally, cytological sampling is generally safe and relatively easy to perform with a low rate of associated complications,[3,9,148,202,232] and might be guided by different imaging techniques (e.g. ultrasounds) which can improve precision in specimen collection and interpretation of morphological findings.[7,9,11,40,44,124,125,134,155,170,173] Finally, when compared to histology, cytology allows a better identification of etiological agents and an improved definition of cellular and nuclear details.[190]

The most important limitation of citology is perhaps represented by the impossibility to evaluate the microanatomical architecture of the investigated tissue,[23,32,189] although cellular arrangements suggestive of certain structures such as glandular tubules and capillaries can be still observed.[123,164] Lack of recognizable tissue architecture, along with intrinsical features of the cells (e.g. poorly differentiated tumors) might lead to non definitive cytological diagnoses.[23,32,128,189,190,229] Inconclusive sampling, i.e. specimens that are acellular or poorly-cellular or characterized by marked blood contamination,[23,41,81,189,193] represents an additional limitation of cytology that can depend upon the nature of the sampled organ (e.g. the high content of blood in spleen) or the lesions affecting it (e.g hemhorrages or necrosis). [81,128,189,190,193] Finally, non-definitive diagnoses or misdiagnoses can also be due to the inexperience of the cytologist who evaluates the sample,[42,128,174,190] or by the inexperience of the operator who collect and prepare the smear or the touch imprint.[3,81,189] The issue is further complicated by the fact that a variable degree of interobserver agreement,[52,208] as well as a considerable lack of consensus among cytologists in the use of modifiers (such as “possible” or “probable”),[42] can affect cytological diagnosis.

According to the balance between the aforementioned advantages and limitations, cytology is still considered an irreplaceable diagnostic tool for a preliminary evaluation of lesions observed in daily clinical practice.[43,189,190] Nevertheless, the estimation of the reliability of this technique represents a fundamental step, given the

key role that a cytological diagnosis might cover for the clinical approach and for the selection of the most appropriate treatment for each patient.[42,52,189,190] In this regards, diagnostic accuracy studies can provide a proof of reliability of cytological results, when the latter are compared to the histopathological gold standard,[23,189–191] which should be more precisely defined as the "reference test" or "reference standard".[27] According to 2007 FDA guidelines,[215] "the *diagnostic accuracy* of a new test refers to the extent of agreement between the outcome of the new test and the reference standard". In this context, histopathology represents the best alternative to cytology, being characterized by the tremendous advantage of allowing the evaluation of tissue and of lesion architecture.[23,32,128,189,191,229] Although histology can be still affected by limitations too, such as the number of sections evaluated [92,120] or a variable degree of interrater agreement in specific conditions,[52,147] as well as by its intrinsic increased invasiveness, this technique is still widely recognized as the reference standard.[23,189–191]

Unfortunately, despite the aforementioned premises, studies investigating the diagnostic accuracy of cytology in veterinary medicine are still too few. Some examples of studies specifically focusing on the evaluation of cytological diagnostic accuracy in the canine and feline species are listed below, according to the organ system investigated:

- Lymph nodes;[71,106,108]
- Bone;[19,176]
- Mammary gland;[55,61]
- Liver;[7,11,173,202]
- Oral cavity;[24]
- Nasal cavity;[35,36]
- Kidney;[124,125]
- Multiple organs;[30,48,58,155]
- Spleen;[9,40,44,148]
- Prostate;[170]
- Skin;[76]
- Perianal glands. [177]

In other cases, cytological diagnostic accuracy studies focused on specific pathological entities:

- Soft tissue sarcomas;[134]
- Feline intestinal lymphoma;[175]
- Feline nasal hamartoma;[28]
- Melanoma;[83,163]
- Pheochromocytoma.[18]

As common sense would suggest, a detailed comparison of results among the aforementioned studies would be useless and time-wasting, due to the marked differences between the two species (i.e. dog and cat) and among the organs and pathological conditions investigated. Disregarding the common belief that for cytology sensitivity and negative predictive value are generally lower than specificity and positive predictive value, respectively,[189] a high variability in terms of these indexes has reported among studies (as examples, the readers are invited to consult the following references: [7,36,71,124,125]). This means that the accuracy of cytology in ruling in or ruling out a certain diagnosis is strictly dependent on the tissue and on the lesion investigated.[189] This observation correlates with the observation that, based on cell morphology, a different tendency to exfoliate has been reported for different neoplasms, with round cell tumors showing the highest exfoliation rate and spindle cell tumors the lowest.[36,41,48,204,229] Another advantage of diagnostic accuracy investigations is the possibility to apply this kind of study design to refine cytological criteria and quantification cut-off for certain cellular elements in view to improve consistency with the histopathological reference standard.[28,202] This approach allows the establishment of classification schemes characterized by a reduction of the interobserver agreement among different Institutions as well as by the possibility of establishing potential correlations of the same schemes with the clinical outcome.[52,227] Additionally, diagnostic accuracy studies can be further strengthen by the investigation of interobserver agreement,[157,174] according to the observation that interrater variation might influence reliability of diagnostic accuracy studies.[71,83,106] Nonetheless, only a few of the aforementiond studies investigating diagnostic

accuracy of cytology in veterinary medicine included the evaluation of interobserver agreement among different cytologists.[18,28]

Besides the unavoidable differences in study designs implied by the investigation of such different organs and pathological entities, some of the aforementiond studies investigating diagnostic accuracy of cytology in veterinary medicine included results which cannot be compared with successive similar works due to incomplete reporting of data.[189,190] Some specific examples are represented by the reporting of cumulative results of diagnostic accuracy of cytology simultaneously referring to different species,[9,48,58] or by lack of reporting of clear indications concerning the number of animals affected by a specific clinical condition and included in the study.[106] In this context, it should be remembered that incomplete reporting has been described as one of the major sources of avoidable waste in biomedical research.[52,78] Further factors which might contribute to make difficult critical appraisal and replication of studies is the lack of clear indications concerning selection criteria for investigated cases, according to the observation that diagnostic accuracy indexes are not fixed properties of the test under evaluation and are influenced by the prevalence of the investigated condition.[27,189,233] To reduce the amount of avoidable waste in biomedical research in the field of diagnostic accuracy studies, in 2003 a group of different professional figures introduced specific guidelines, called “Standards for Reporting Diagnostic Accuracy Studies” (abbreviated as STARD).[26] Nowadays, the latest and most recent version of these guidelines is the one updated in 2015,[27] which has been accompanied by a detailed explanation and elaboration document,[46] and followed by the publication of guidelines for drafting abstracts in diagnostic accuracy studies.[47] The aims of STARD guidelines are to avoid incomplete reporting in diagnostic accuracy studies and to improve the general quality of the latter, reducing problems related to study identification, critical appraisal, and replication.[27] Although these guidelines have been available since 2003,[26] only a very limited number of studies investigating the diagnostic accuracy of cytology in veterinary medicine applied them.[7,18,19,36,71,124,125] Therefore, further studies are warranted to investigate in further details the reliability of cytology in terms of diagnostic accuracy in veterinary medicine.

In this context, the main aim of this Thesis was the investigation of diagnostic accuracy of cytology in several fields of round cell neoplasms and sarcomas in veterinary medicine, including:

- Diagnostic accuracy of cytology in the evaluation of canine splenic neoplasm (**Chapter 1**);
- Diagnostic accuracy of cytology and interobserver agreement in the immunophenotype prediction of feline nodal lymphomas (**Chapter 2**);
- Diagnostic accuracy of cytology and interobserver agreement in the evaluation of the metastatic status of lymph nodes obtained from mast cell tumor-bearing dogs, preceded by preliminary investigations on the quantification of mast cells in cytological nodal samples obtained from the same patients (**Chapter 3**).

Each Chapter opens with a review of the literature in each of the aforementioned topics.

Besides investigations in the field of the diagnostic accuracy of cytology, during his externship at the University of Veterinary Medicine of Hannover (TiHo Hannover, Germany) the PhD candidate was involved in 2 research projects concerning viral oncolysis in a cell culture model of canine histiocytic sarcoma. The 2 projects are the main focus of investigation of the Ph.D. Thesis of Dr. Federico Armando from the Department of Veterinary Medicine Sciences of the University of Parma (Parma, Italy); therefore, only a brief summary of the main findings is reported in the current Thesis (**Addendum**).

CHAPTER 1 - DIAGNOSTIC ACCURACY OF CYTOLOGY IN THE EVALUATION OF CANINE SPLENIC NEOPLASMS

SCIENTIFIC PRODUCTION OF THE PH.D. CANDIDATE IN THE FIELD

The study described in the current chapter has been published in 2019 in the peer-reviewed Journal "PLOS ONE", with the Ph.D. candidate as one of the 2 co-first Authors:

- Tecilla M*, Gambini M*, Forlani A, Caniatti M, Ghisleni G, Roccabianca P (2019) Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. PLoS ONE 14(11): e0224945.

The manuscript, as well as Supplementary Table S1 and Supplementary Tabel S2, can be found in Open Access at the following link: https://doi.org/10.1371/journal.pone.0224945.

INTRODUCTION AND AIMS

The complex microanatomical architecture of the spleen, composed by an intricate bed of vascular sinuses intermingled with lymphoid follicles and supported by thick stromal septa, is somehow mirrored by the wide range of tumors that can develop within this organ.[220,227] Literature is widely concordant in describing hemangiosarcoma (HES) as the most common primary splenic malignant tumor of dogs. [44,49,64,97,198] However, there is no consensus regarding the incidence of benign and malignant splenic lesions, with some studies reporting a higher prevalence of benign lesions,[17,44,57,198,199] including hematoma and hyperplasia, and other works describing malignancies as the most frequent pathologic conditions.[4,49,64,86,97] In this context, due to the poor prognosis associated with HES, splenectomy is still the routine approach to most canine splenic masses for both diagnostic and therapeutic purposes.[44,49,69,232]

In human beings, splenectomy has been associated with the occurrence of the so-called “overwhelming post-splenectomy sepsis”, which is a lethal syndrome mostly affecting pediatric and young patients.[51,149] Despite not so common, this syndrome represents a serious concern in human medicine, thus splenic preservation, especially in children, is now the standard care in hemodynamically stable patients when underlying splenic neoplasia is not suspected.[8] The latter is an underestimated problem in veterinary medicine, where splenectomy is still the routine therapeutic approach to splenic lesions [199] despite the recent indications that the surgical procedure should be avoided as possible.[40] Post-splenectomy survival rate has been described lower in dogs with HES compared to dogs with benign lesions.[97,199] Although this can be interpreted as fair survival data, it has been reported that a proportion of dogs (7.6%) might develop complications secondary to splenectomy because of thrombotic or coagulopathic syndromes.[236] Additional adverse effects following splenectomy in dogs reported in the literature included peri- and post-operative ventricular arrhythmias,[121,151,236] reduced blood filtration and renewal,[132,133] impairment of humoral immune response,[98] reduced immune-surveillance against bacteria and parasites,[100,101,110,168,211] and higher incidence of gastric dilatation-volvulus.[2,69,115,181]

In this context, any preoperative diagnostic approach to splenic lesions might be beneficial in preventing unnecessary splenectomy. Ultrasonographic examination of nodular splenic lesions in dogs is not reliable to differentiate with certainty benign and malignant processes, requiring the use of additional, ideally minimally invasive, diagnostic tests.[109] In veterinary medicine, cytologic sampling represents one of the most useful approaches to the pre-surgical diagnosis of splenic disease.[9,41] Indeed, the spleen is readily accessible due to its superficial anatomical location.[97] Notwithstanding the common belief that splenic aspiration can be dangerous especially when investigating cavitated masses,[40,44,69] complications from splenic aspiration procedures are rarely elicited even in thrombocytopenic animals.[9,148,199,232] Similarly, in human medicine splenic fine needle aspiration (FNA) cytology is seldom associated with complications,[79,82] resulting in 5.2% secondary complications with fewer than 1% considered severe and consisting mostly of controllable hemorrhage.

Given the aforementioned considerations, attempts to minimize unnecessary splenectomy should prompt an increased use of additional diagnostic techniques as preoperative screening tests to characterize splenic disease. Fine needle aspiration cytology can provide diagnostic information useful to distinguish inflammatory, benign and malignant nodular lesions and to assess generalized splenomegaly.[9,41,44]

Despite the relatively high frequency of splenic diseases in dogs, data regarding usefulness and validity of diagnostic cytology are fragmentary. In veterinary medicine, no studies have comprehensively assessed overall accuracy, sensitivity, specificity, positive and negative predictive values of canine splenic cytology against histopathology utilizing the Standards for the Reporting of Diagnostic Accuracy Studies (STARD) guidelines.[9,30,40,48,58,148,232] STARD guidelines have been created to avoid incomplete reporting in diagnostic accuracy studies and to improve the general quality of the latter, reducing problems related to study identification, critical appraisal, and replication.[27] Overall agreement between cytology and histology of canine splenic lesions is the most frequent index reported, ranging from 38 to 100%.[9,30,40,44,48,58,148,232] It is noteworthy that this index has been evaluated on a limited number of splenic cytological specimens (range: 5-40).[9,30,40,44,48,58,148,232] In most reports, sensitivity, specificity, positive and negative predictive values of splenic cytology were not calculated and cannot be properly estimated because caseloads simultaneously included multiple species,[9,30,48,58,148] multiple tissues and organs,[30,48,58] or “equivocal” or “provisional” cytological and histological diagnoses.[9,30,40] Application of STARD guidelines in the current study allowed cross-tabulation of cytological results (i.e. the index test) against those of histopathology (i.e. the reference standard) to generate sensitivity and specificity values.[27] These data have not been included in previous studies and will be useful for future researchers comparing diagnostic methods for canine splenic neoplasms.

To avoid unnecessary splenectomy, a diagnostic test with a high sensitivity and negative predictive value would be desirable because these indexes measure the percentage of diseased dogs correctly diagnosed with splenic neoplasia and the probability that dogs with a negative cytology truly do not have a neoplasm, respectively. In this context, the aim of this study was to determine the diagnostic

accuracy of cytology in the diagnosis of canine splenic neoplasms utilizing the corresponding histopathology as the diagnostic reference standard,[9,30,36,48,71] following STARD guidelines.[27] Additionally, sensitivity of cytology in the diagnosis of specific tumor types and in the identification of nodular versus diffuse neoplasms was evaluated.

MATERIAL AND METHODS

Criteria of selection of cases

In this retrospective study, the electronic cytological database of the Diagnostic Pathology Service of the Department of Veterinary Medicine (DIMEVET) of the University of Milan was searched for splenic samples collected between January the 1st, 1998 to January the 31st, 2018. The database was searched for specific key words in the following combinations: 1) “dog” and “spleen”, 2) “dog” and “splenic”. A consecutive series of canine splenic cytologies was obtained. Cytological samples came from external referring private practices or from the Veterinary Teaching Hospital (VTH) of the DIMEVET, or were prepared from fresh surgical biopsies and necropsies. Samples were submitted or collected to evaluate splenomegaly or nodular lesions.

The histopathology database was then searched for the histopathological specimen(s) corresponding to the same lesion examined by cytology. Histopathological samples were obtained from splenic biopsies (nodular lesions) or whole spleens (from splenectomies or necropsies) submitted to the Diagnostic Pathology Service of the DIMEVET, or were represented by slides submitted by external pathologists as second opinion cases. A time interval >45 days between cytological and histological sampling was considered as an exclusion criterion. Histopathological samples collected via needle core biopsies (NCBs) were excluded from the study as they may bear reduced diagnostic reliability compared to incisional and excisional histological samples.[19,232] Cases were included in the study only when at least 1 cytological and 1 histological slide of the same lesion were available for review.

Additional information collected from the archives for cases investigated in the study included sex, age, breed, cytological sampling technique (e.g. fine needle aspiration – FNA, touch imprinting, scraping smearing), and gross appearance of the lesion (i.e. diffuse versus nodular lesion).

To improve data completeness, transparency, and reproducibility, the study was conducted following the STARD guidelines [27] to the best of Authors' ability. All canine splenic samples included in the current retrospective work, regardless of the sampling technique applied, were submitted to the Diagnostic Pathology Service of the DIMEVET for diagnostic purposes of spontaneous developing diseases. No animals were sampled or euthanized for research use. The use of animal tissue in the current study was approved by the Ethics Committee in charge for animal welfare of the University of Milan (Organismo Preposto al Benessere degli Animali, OPBA) with protocol number OPBA_86_2019. Sensitive information regarding owners and animals were stored, managed and preserved according to European and Italian laws.

Sample processing

Cytological samples were air dried and manually stained with May-Grünwald-Giemsa (Merck KGaA, Frankfurt, Germany). Tissue samples for histopathology were fixed in 10% neutral buffered formalin, processed routinely, and embedded in paraffin wax. Sections of 5 µm were stained with Hematoxylin and Eosin. Second opinion cases were provided as Hematoxylin and Eosin stained slides by the referring pathologists.

Case review

All cytological and histopathological samples were independently reviewed in a blinded fashion by three cytologists (an ECVP diplomate, an ECVCP diplomate, and an ECVP trainee) and by three anatomical pathologists (2 ECVP diplomates and an ECVP trainee), respectively. Both cytologists and anatomical pathologists were blinded to signalment information related to each case.

For each cytological case, one slide for each sampling technique was reviewed. First, slides were examined at low-power magnification (i.e. 100x) to assess the adequacy of the specimen. Those samples characterized by marked hemodilution in the absence of both stromal elements and a mixed population of leukocytes, were defiend as “poorly cellular”.[41] Poorly cellular and poorly smeared samples (i.e. samples too thick or where most cells were ruptured), and samples where stain quality impaired adequate definition of the cell type (e.g. formalin-contaminated smears), were considered inconclusive.[41] As previously reported,[9,19,30,36,58,71,232] inconclusive cases were excluded from the statistical analysis.

Cytological diagnoses were expressed according to those reported in the literature (Figure 1).[17,41,128,165] To facilitate comparison of the agreement between cytological and histological results, each cytological sample was further classified as neoplastic or non-neoplastic according to the main pathologic process. Non-neoplastic samples were those characterized by degenerative, reactive (including extramedullary hematopoiesis), and inflammatory changes, as well as normal specimens consisting of stromal elements with mixed leukocyte population.[40,41,165] Reviewing cytologists were not allowed to use diagnostic modifiers such as “probably”, “most likely”, “suggestive of”, as previously reported,[19] nor to provide equivocal diagnoses (i.e. reporting more than one differential diagnosis). When a univocal diagnosis was not reached, cytologists reviewed the case collaboratively to find an agreement. Only the definitive diagnosis was included in the consecutive statistical analysis.

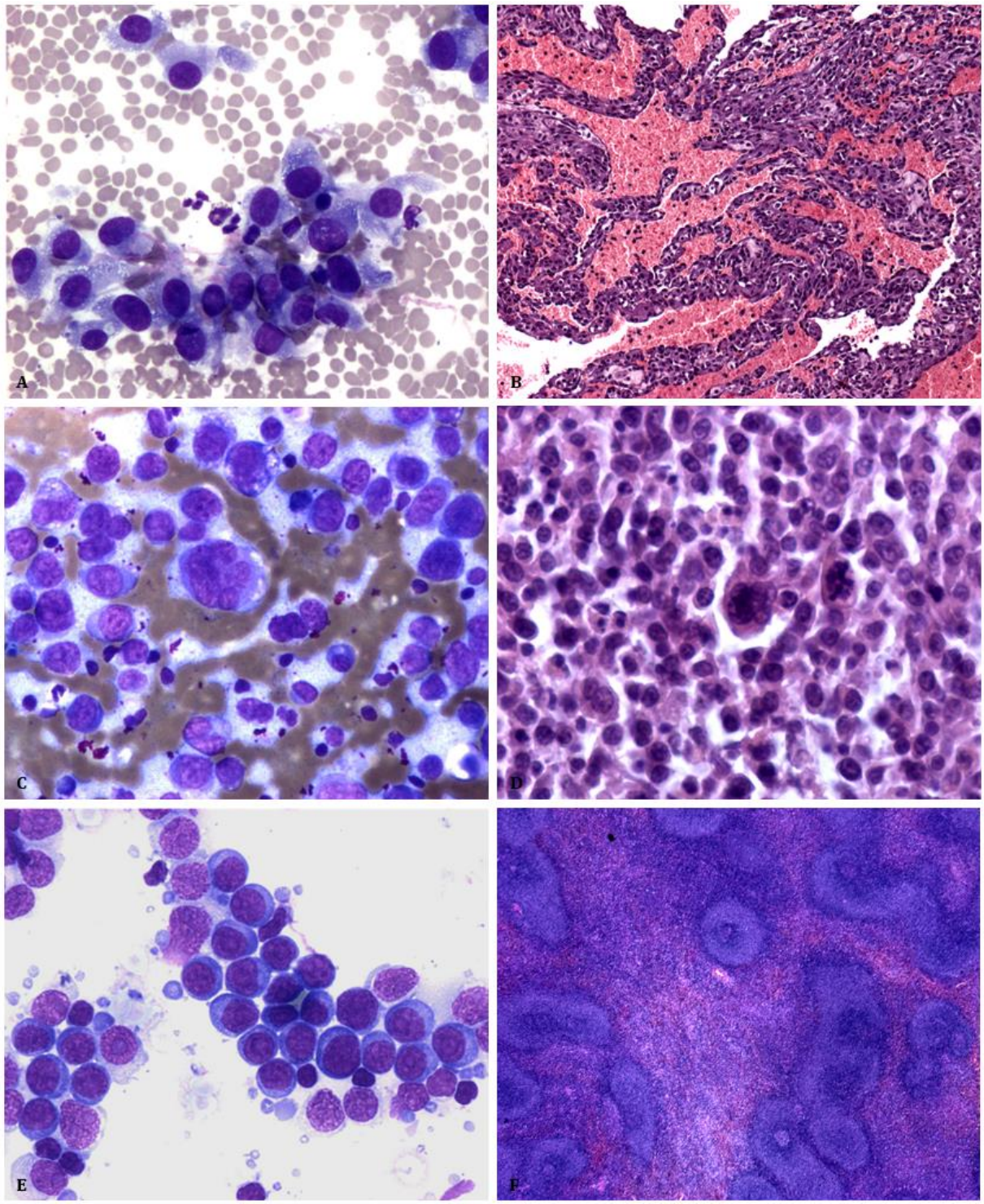
Neoplastic cytological samples were further subdivided by tumor type into the following subcategories: benign soft tissue tumor including angioma (BSTT), angiosarcoma (HES), soft tissue sarcoma other than angiosarcoma (STS), lymphoma (LYM), mast cell tumor (MCT), histiocytic tumors including hemophagocytic sarcoma (HS), other round cell tumors including plasma cell tumor, myeloid leukemia, and undifferentiated round cell tumor (ORCT), carcinoma (CARC), and malignant neoplasm not otherwise specified (MNNOS).

Cases were included only when the three anatomical pathologists were in agreement because histopathology served as the diagnostic reference standard to evaluate cytological diagnostic accuracy. Neoplasms were classified applying the World

Health Organization's histologic classification of tumors in domestic animals (Figure 1).[89,102,139,220–227] To further standardize histopathological diagnoses, anatomic pathologists were invited to classify some specific pathological entities (i.e. lymphomas, histiocytic proliferative disorders, nodular lesions previously classified as "fibrohistiocytic nodules") according to criteria reported in recent literature.[137,140,218,219] For nodular lesions, histological samples were considered conclusive and therefore included in the statistical analysis only if at least one slide containing at least one margin between the nodule and the adjacent splenic parenchyma was available for review.[57,198,199]

Histological samples were then classified as neoplastic and non-neoplastic. Neoplastic cases were further subclassified utilizing the same subcategories applied to cytological samples.

Figure 1. Cytological smears and corresponding histopathology of canine splenic lesions.



A) Hemangiosarcoma, fine needle aspiration cytology (FNA). Moderately cellular sample composed of small aggregates of spindle cells with distinct cell margins, blue cytoplasm, and numerous cytoplasmic vacuoles. The nucleus is oval, with homogeneous chromatin. May-Grünwald-Giemsa (MGG). **B)** Hemangiosarcoma, histopathology. Corresponding histopathology of case illustrated in Figure 1. Hematoxylin-Eosin (HE). **C)** Lymphoma, FNA. Highly cellular sample composed of pleomorphic neoplastic round cells, of 14-60 µm, frequently

characterized by multinucleation and nuclear clefts. MGG. **D)** Lymphoma, histopathology. Histopathological evaluation of the sample represented in Figure 3. The splenic parenchyma is replaced by a solid sheet of round cells with the same cytological features detected in the cytological sample. HE. **E)** Marginal zone hyperplasia, FNA. Highly cellular sample, composed by a monomorphic population of medium sized lymphocytes, with one prominent round central nucleolus and abundant bluish cytoplasm, mimicking a marginal cell lymphoma. MGG. **F)** Marginal zone hyperplasia, histopathology. Histological evaluation of the sample represented in Figure 5. Expansion of the marginal zones without formation of a splenic nodular lesion and with maintenance of the normal splenic architecture. HE.

Data analysis

For all cases, the cytological diagnosis was compared with its paired histopathological diagnosis. Since histological samples maintain tissue architecture and are not biased by cellularity,[19,58,229,232] histopathology was set as the reference standard as previously reported.[9,30,36,48,71]

To determine diagnostic accuracy indexes, cytological specimens were classified according to four correlation categories (true positive, true negative, false positive, false negative). Specifically, the True Positive (TP) category included all cytological samples diagnosed as neoplastic with a corresponding neoplastic histopathology. The True Negative (TN) category comprised all cytological samples diagnosed as non-neoplastic with a corresponding non-neoplastic histopathology. The False Positive (FP) category included all cytological samples diagnosed as neoplastic with a corresponding non-neoplastic histopathology. The False Negative (FN) category comprised all non-neoplastic cytological diagnoses with a corresponding neoplastic histopathology.

To evaluate the sensitivity of cytology in the diagnosis of specific tumor types, only those cases histologically confirmed as neoplastic were taken into account. The subcategories assigned to each cytological and corresponding histological sample were then compared. When cytological and histological diagnoses matched for both neoplastic categorization and tumor type subcategorization, the case was defined as “true positive with complete agreement”. When cytological and histological diagnoses matched for the neoplastic categorization but did not match for the tumor type subcategorization, the case was defined as “true positive with partial agreement”.

When a histopathological diagnosis categorized as neoplastic corresponded to a cytological diagnosis categorized as non-neoplastic, the case was considered in disagreement and defined as “false negative case”.

Statistical methods

Cytological-histological correlation categories (TP, FP, TN, FN) were included in a 2x2 contingency table and used to calculate point estimates of overall accuracy, sensitivity, specificity, positive and negative predictive values.[24,71] Overall accuracy was defined as the ability of cytology to correctly identify neoplastic and non-neoplastic lesions, and was calculated as the sum of cases in which cytology and histology agreed in diagnosing a lesion as neoplastic (i.e. TP) or non-neoplastic (i.e. TN), divided by the total number of cases included in the study.[58,106] Given that pre-determined acceptability criteria for diagnostic performance of splenic cytology to distinguish between neoplastic and non-neoplastic lesions have not been previously established, overall accuracy, sensitivity, specificity, positive and negative predictive values were considered low if <70%, moderate if ≥70% and <80%, high if ≥80% and <90%, and very high when ≥90%. (Bonfanti et al., 2015) To increase data comparability with other studies, positive and negative likelihood ratios were also calculated.[71]All diagnostic accuracy indexes and the corresponding ninety-five percent (95%) confidence interval were calculated for each of the above mentioned indices of diagnostic test accuracy using a web-based application (MEDCALC - https://www.medcalc.org/calc/diagnostic_test.php).

The level of agreement between cytology and histopathology in the diagnosis of splenic neoplastic conditions was further investigated calculating the Cohen's kappa coefficient (κ), which was then corrected by the standard error. The value of k can be indicative of no agreement (if k <0), slight agreement (k = 0 - 0.20), fair agreement (k = 0.21 - 0.40), moderate agreement (k = 0.41 - 0.60), substantial agreement (k = 0.61 - 0.80), almost perfect agreement (k = 0.81 - 0.99), or perfect agreement (k = 1).[24,71] Cohen's kappa and standard error were calculated utilizing GraphPad QuickCalcs Web site (GraphPad Inc. - https://www.graphpad.com/quickcalcs/kappa2/).

Sensitivity of cytology in differentiating splenic tumor types was defined as the ability of cytology to correctly identify as neoplastic a sample belonging to a specific neoplastic subcategory. Therefore, sensitivity for each tumor type was calculated as the sum of cases in complete and partial agreement (i.e. true positive cases) divided by the total number of cases with that specific neoplasm.[48]

Similarly, the sensitivity of cytology in the diagnosis of neoplastic lesions according to their distribution pattern (i.e. diffuse or nodular) was evaluated. For each distribution pattern, sensitivity was calculated as the sum of cases in complete and partial agreement (i.e. true positive) divided by the total number of cases with a specific distribution pattern. Sensitivity of cytology according to distribution pattern was calculated for neoplastic lesions in general (i.e. the general sensitivity value obtained in our study), as well as for those specific neoplastic subcategories including cases with either diffuse or nodular distribution pattern.

Chi-square analysis applied to pairwise comparison was performed to evaluate whether statistically significant differences existed in the sensitivity of cytology for different tumor types, as well as in the diagnosis of neoplastic lesions with diffuse or nodular distribution pattern.[19,71] Specifically, the sensitivity of cytology for each neoplastic subcategory was compared with the sensitivity for splenic neoplasms in general, the sensitivity for all other neoplastic subcategories grouped together, and the sensitivity for any other neoplastic subcategory. Similarly, the difference between sensitivity for nodular or diffuse lesions among neoplasm in general and for each neoplastic subcategory was statistically investigated. Chi-square analysis was performed only on sensitivity values different from 0% and 100%, using MEDCALC (https://www.medcalc.org/calc/comparison_of_proportions.php). A p-value <.05 was considered statistically significant.

RESULTS

Animals and samples

From a total of 950 splenic cytological samples retrieved between 1998 and 2018, 92 cytological samples from 91 dogs were included in the study, with one dog sampled

for two distinct splenic lesions. A total of 858 splenic cytological cases were excluded for the following reasons: lack of a corresponding histopathological sample (832 cases), unavailable cytological and/or histological samples (16 cases), and histological diagnosis expressed on NCBs (10 cases). Among the selected cytological cases, 14 were considered inconclusive, and therefore excluded from the consecutive statistical analysis. Detailed evaluation of diagnostic accuracy was performed on 78/92 reviewed cytological samples (retrieval rate: 84.78%) obtained from 77 dogs.[19]

Sex was available for 76/77 dogs: 19 were neutered females, 6 neutered males, 21 intact females, and 30 intact males. Mean age was 9.05 years (age range 2 months-16 years; age was not available for 2 cases). Twenty-six (26) breeds other than mongrels were represented; in one case breed was not provided.

The time interval between cytological sampling and the corresponding histopathology collection ranged from 0 to 44 days for all cases.

Of the 78 cases included in the study, 81 cytological slides were evaluated (3 cases were sampled with two different techniques, i.e. touch imprinting and scraping smearing). Cytological samples consisted of 43/81 touch imprints (53.09%) collected from both surgical biopsies and necropsies, and 28 FNAs (34.57%). Among the latter, 21 were ultrasound guided, 1 was CT-scan guided, 1 was obtained during surgery, while for 5 FNA biopsies no additional sampling information was available. Finally, in 6 cases scrapings were obtained from surgical and necropsy specimens (7.41%). In 4 cases (4.94%) the sampling technique was not specified.

Complete agreement among anatomical pathologists was reached for all the 78 corresponding histopathological samples. Histopathological specimens were distributed as follows: 51 surgical samples from partial or complete splenectomies (51/78 cases, 65.38%), 24 spleens from necropsies (24/78, 30.77%), and 3 cases submitted as a second opinion (3/78, 3.85%).

Detailed signalment information and cytological sampling techniques applied for dogs investigated in the current study are reported in Supplementary Tables S1 and S2 (available at the following link: https://doi.org/10.1371/journal.pone.0224945).

Cytological and histological diagnoses

All cytological and corresponding histopathological diagnoses (78 cases) are listed in Supplementary Table S1 (available at the following link: https://doi.org/10.1371/journal.pone.0224945). The diagnoses for the cytological-histological pairs excluded due to inconclusive cytology are listed in S2 Table (available at the following link: https://doi.org/10.1371/journal.pone.0224945).

No diagnostic differences were found for samples collected using two different sampling techniques, and therefore they were considered as one case in the consecutive statistical analysis. Cytologically, 37/78 cases were diagnosed as neoplastic (47.44%) and 41/78 as non-neoplastic (52.56%). All cases diagnosed as neoplastic were classified as malignant, thus no benign neoplasms were cytologically observed.

Histologically, 56/78 cases (71.79%) were neoplastic (S1 Table) and 22/78 cases (28.21%) were non-neoplastic. Malignant tumors were 51 (51/56 tumors, 91.07%) and 5 were benign. The prevalence of each tumor type is reported in Table 1. No malignant neoplasm not otherwise specified (MNNOS) were observed among evaluated cases.

Table 1. Prevalence, agreement levels, and sensitivity of cytology in the diagnosis of each neoplastic subcategory.

	Prevalence	TP cases with complete agreement	TP cases with partial agreement	FN cases	Sensitivity	Confidence Interval (95%)
TOTAL	71.79% (56/78)	42.86% (24/56)	21.43% (12/56)	35.71% (20/56)	64.29%	50.36% - 76.64%
HES	28.57% (16/56)	68.75% (11/16)	6.25% (1/16)	25% (4/16)	75%	47.62% - 92.73%
LYM	28.57% (16/56)	37.50% (2/16)	12.50% (6/16)	50% (8/16)	50%	24.65% – 75.35%
STS	12.50% (7/56)	42.86% (3/7)	28.57% (2/7)	28.57% (2/7)	71.43%	29.04% - 96.33%
BSTT	8.93% (5/56)	0% (0/5)	0% (0/5)	100% (5/5)	0%	0.00% - 52.18%

	Prevalence	TP cases with complete agreement	TP cases with partial agreement	FN cases	Sensitivity	Confidence Interval (95%)
HS	7.14% (4/56)	25% (1/4)	50% (1/4)	25% (1/4)	75%	19.41% - 99.37%
MCT	7.14% (4/56)	100% (4/4)	0% (0/4)	0% (0/4)	100%	39.76% - 100%
CARC	5.36%(3/56)	66.67% (2/3)	33.33% (1/3)	0% (0/3)	100%	29.24% - 100%
ORCT	1.79% (1/56)	0% (0/1)	100% (1/1)	0% (0/1)	100%	2.50% - 100%

Legend: BSTT, benign soft tissue tumor including angioma; CARC, carcinoma; FN, false negative; HES, angiosarcoma; HS, histiocytic neoplasm (including hemophagocytic syndrome); LYM, lymphoma; MCT, mast cell tumor (MCT); ORCT, other round cell tumor; STS, soft tissue sarcoma other than angiosarcoma; TP, true positive.

Of the total 78 splenic lesions, 60 were nodular (76.92%), and 17 were diffuse (21.79%), while no information regarding the distribution pattern was available for 1 case (1.28%). Among the 56 neoplastic lesions, 43/56 cases (76.79%) were nodular and 12 cases (21.43%) were diffuse. The case for which distribution pattern was not provided was a liposarcoma (1.79%). This case was excluded from the evaluation of cytology sensitivity according to neoplasm distribution pattern. The proportion of cases with nodular or diffuse pattern for each tumor type are reported in Table 2.

Table 2. Prevalence, agreement levels and sensitivity of cytology in the diagnosis of each neoplastic subcategory on the basis of distribution pattern.

	Nodular				Diffuse			
	# of cases	TP	FN	Sensitivity (95% CI)	# of cases	TP	FN	Sensitivity (95% CI)
TOTAL	43/56 (76.79%)	26/43	17/43	60.47% (44.41% - 75.02%)	12/56 (21.43%)	9/12	3/12	75% (42.81% - 94.51%)
HES	14/16 (87.50%)	10/14	4/14	71.43% (41.90% - 91.61%)	2/16 (12.50%)	2/2	0/2	100% (15.81% - 100%)
LYM	10/16 (62.50%)	5/10	5/10	50% (18.71% - 81.29%)	6/16 (37.50%)	3/6	3/6	50% (11.81% - 88.19%)

	Nodular				Diffuse			
	# of cases	**TP**	**FN**	**Sensitivity (95% CI)**	**# of cases**	**TP**	**FN**	**Sensitivity (95% CI)**
MCT	1/4 (25%)	1/1	0/1	100% (2.50% - 100%)	3/4 (75%)	3/3	0/3	100% (29.24% - 100%)

Legend: CI, confidence interval; FN, false negative; HES, angiosarcoma; LYM, lymphoma; MCT, mast cell tumor (MCT); TP, true positive.

Cyto-histological correlation

Following the tabulation of cytological and histological diagnoses (S1 Table), 36 cases (46.15%) were classified as TP, 21 (26.92%) as TN, 20 (25.64%) as FN, and 1 (1.28%) as FP (Table 3).

Table 3. Cytological-histological correlation categories.

Diagnosis	Histology: neoplastic	Histology: non-neoplastic	Total
Cytology: neoplastic	36 (TP)	1 (FP)	37
Cytology: non-neoplastic	20 (FN)	21 (TN)	41
Total	56	22	78

Legend: FN, false negative; FP, false positive; TN, true negative; TP, true positive.

The FP case had a cytological diagnosis of lymphoma that corresponded histologically to a purulent bacterial splenitis (in this case the full spleen was available for analysis and no tumor was found; however, severe marginal zone hyperplasia was present).

Neoplastic and non-neoplastic lesions were correctly identified in 57/78 cases (Table 3), thus overall accuracy of cytology was 73.08% (Table 4). Sensitivity of cytology in the diagnosis of splenic neoplasms was 64.29%, specificity was 95.45%, positive predictive value was 97.30%, and negative predictive value was 51.22% (Table 4). Positive and negative likelihood ratios were 14.14 and 0.37, respectively (Table 4).

Table 4. Prevalence of neoplastic lesions, with point estimate and 95% confidence interval of diagnostic accuracy indexes, likelihood ratios and Cohen's k.

Diagnostic accuracy index	Value	Confidence Interval (95%)
Prevalence	71.79%	60.47% – 81.41%
Overall accuracy	73.08%	61.84% - 82.50%
Sensitivity	64.29%	50.36% - 76.64%
Specificity	95.45%	77.16% - 99.88%
PPV	97.30%	84.01% – 99.60%
NPV	51.22%	42.21% - 60.15%
PLR	14.14	2.06 – 96.94
NLR	0.37	0.26 – 0.54
K value	0.473	0.304 – 0.643

Legend: NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value.

According to Cohen's test, the level of agreement was considered as "moderate", with a κ value of 0.473 corresponding to a standard error of 0.086.

The distribution of TP and FN cases for each tumor type is reported in Table 1. The sensitivity of cytology in the diagnosis of each neoplastic subcategory was 100% for MCT, CARC, and ORCT, 75% for HES, 75% for HS, 71.43% for STS, 50% for LYM, and 0% for BSTT included in the study. Further details regarding complete and partial agreement between cytological and histological diagnoses as well as confidence intervals of sensitivity value for each tumor type are listed in Table 1. Chi-square analysis of cytological sensitivity was applicable only to HES, HS, STS, and LYM. No statistically significant sensitivity differences were observed (*p*-value range: 0.1506 - 1.0).

The proportion of TP and FN cases with nodular or diffuse pattern for each tumor type are reported in Table 2. Sensitivity of cytology in the diagnosis of neoplastic lesions in general according to their distribution pattern was 60.47% for nodular and 75% for diffuse neoplasms, with no statistically significant difference between the two values ($p = 0.3593$). For some tumor types, the sensitivity of cytology on the basis of the distribution pattern was not calculated, given that only nodular (BSTT, STS, HS,

CARC) or diffuse (ORCT) neoplastic lesions were represented in these categories. Sensitivity in the diagnosis of nodular and diffuse lymphomas was for both 50%, with no statistically significant difference between the two values (p = 1.0). Sensitivity for nodular angiosarcomas was 71.43% and 100% for diffuse angiosarcomas, while sensitivity for both nodular and diffuse mast cell tumors was 100%. Considering these results, Chi-square analysis of sensitivity on the basis of the distribution pattern was not performed for angiosarcomas and mast cell tumors.

DISCUSSION

In this study we reported overall accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of cytology for the diagnosis of canine splenic neoplasms. Similar previous studies [9,30,40,44,48,58,148,232] limited the evaluation of cytological diagnostic accuracy to overall agreement with histopathology, hampering comparison with our results. Our study evidenced a moderate overall accuracy of cytology. Specifically, although this technique had a very high specificity and positive predictive value for the diagnosis of splenic neoplasia, sensitivity and negative predictive value were low, indicating that cytological diagnosis of splenic neoplasia is reliable, but a negative result cannot be used to exclude the possibility of splenic neoplasia. Specifically, low sensitivity and negative predictive value of this study indicate that a cytology negative for neoplasia should prompt further investigations to confirm a dog to be truly free from neoplastic disease. This contrasts with our initial hypothesis that cytology may represent a useful tool to avoid unnecessary splenectomy. Nontheless, the very high specificity and positive predictive value identified cytology as a good and reliable tool to rule in the diagnosis of splenic neoplasia with a high degree of confidence. In practical terms, a cytology positive for neoplasia may lead to a faster surgical treatment, avoiding lag times and higher costs associated with application of diagnostic imaging techniques such as contrast-enhanced ultrasound and computed tomography (CT).[45,109] Our results are in line with studies evaluating diagnostic accuracy of cytology applied to various organs in dogs,[24,36,48,58,71] with sensitivity and negative predictive value generally lower than specificity and positive predictive value, respectively.

According to overall accuracy and Cohen's k values observed in the current study, cytology is not a reliable alternative to histopathology in the definitive diagnosis of splenic tumors in most cases, as previously reported for the evalaution of nodal metastasis in canine malignant solid tumors.[71]

To allow comparison of our findings with data previously reported in the literature, the overall accuracy (defined as the sum of complete and partial diagnostic agreements) was calculated from the raw data of previously published caseloads [9,30,40,44,48,58,148,232] when not explicitally reported in the corresponding manuscript. It should be specified that the overall accuracy was calculated considering those cytological cases with an "equivocal" diagnosis (i.e. reporting more than 1 differentials) as in agreement with the corresponding histology if the right diagnosis was cytologically hypothesized. Results of this comparison are illustrated in Table 5.

Table 5. Comparison of the current study with those previously published which simultaneously reported cytological and histological diagnosis of the same lesions.

	Overall accuracy	# cases	Inconclusive samples	Prevalence of neoplasia	Sampling technique	Study design	Factors limiting comparison
Eich et al. (2000)[58]	38%	13	0/13 cases	ND	Intraoperative FNA	P	No distinction dog VS cat
Braun et al. (2007)[30]	69.7%	34	4/34 cases	61.76%	ND	R	ND
Christensen et al. (2009)[40]	88%	17	0/17 cases	52.94%	US-guided FNA (+ 1 squash preparation)	R	ND
Cohen et al. (2003)[48]	60%	5	0/5 cases	ND	ND	R	No species distinction
Watson et al. (2011)[232]	60%	40	5/40 cases	45.71%	US-guided FNA	P	Histology on NCBs
Cleveland et al. (2016)[44]	66.67%	12	0/12 cases	ND	US-guided FNA	R	ND
Ballegeer et al. (2007)[9]	83.87%	32	1/32 cases	54.84%	US-guided FNA	R	No distinction dog VS cat
O'Keefe et al. (1987)[148]	100%	12	0/12 cases	50%	FNA	R	ND
Current study	73.08%	92	14/92 cases	71.79%	FNA, SCRA, TIC	R	ND

Inconclusive cases were excluded from the calculation of the overall accuracy. Prevalence of neoplasia exclusively refers to those cases for which both cytology and histology were available. Legend: FNA, fine needle aspiration; NCBs, needle-core biopsies (not the real reference standard); ND, not determined/not available/nor reported; P, prospective; R, retrospective; SCRA, scraping smearing; TIC, touch imprinting; US-guided, ultrasonography-guided; # cases, number of lesions evaluated for which both cytology and histology were available.

When compared with previous studies, our overall accuracy value (73.08%) laid in between the higher range of 83.87-100% [9,40,148] and the lower 38-69.7% range [30,44,48,58,232] reported in other studies. None of the previous studies reported specificity, sensitivity, positive predictive values, and negative predictive value for the cytological evaluation of of canine splenic neoplasm. Unfortunately, according to the marked differences in terms of diagnostic categories and criteria applied, it was not possible to extrapolate data for the recalcualtion of these diagnositc accuracy indexes from previous studies.

The values of our diagnostic accuracy indexes might have been biased by the exclusion of inconclusive cases from the statistical analysis. Although this approach is allowed by STARD guidelines,[27] the sampling error represents an important source of the overall error when the goal is to determine the diagnostic accuracy of cytology in a clinical setting.[48] The higher number of inconclusive cases observed in the current study compared to previously published studies might be the consequence of our sensibly larger caseload. In a clinical setting, sampling of inconclusive cases might be avoided immediately evaluating smear quality. In a report focusing on cytological evaluation of malignant tumors of dogs and cats,[128] cytological sampling was repeated up to 3 times until an adequately cellular specimen was obtained. Accordingly, this control procedure should be implemented as frequently as possible in clinical practice. In this view, clinicians should keep in mind that, whenever immediate preliminary evaluation of the quality of cytological smears is not feasible, histopathology might still be necessary for reaching a definitive diagnosis.

Regarding the reliability of cytology in the diagnosis of specific tumor types, the lack of statistically significant differences between subcategories observed in the current study may be related to an imbalance in the number of investigated cases for each tumor type. Additionally, our results may be influenced by the tumor histogenesis, since exfoliation rate varies substantially between round cell, epithelial and mesenchymal tumors.[36,41,48,229] Specifically, mesenchymal tumors have the lowest tendency to exfoliate [36,41,48,229] explaining the observed lack of sensitivity in diagnosing benign mesenchymal tumors. Moreover, identification of vascular tumors (i.e. angiomas and HES) among false negative cases is not surprising since the

architecture of these tumors often leads to significant peripheral blood contamination in aspirates (Figure 1 – A and B).[17,41,58,189] In this context, if a clinical concern of HES is posed, correlation with signalment, results from blood analysis, coagulation values, and ultrasound appearance may increase diagnostic accuracy of cytology.[17] However, even though HES are usually challenging to diagnose cytologically due to the abundant hemodilution, the presence of few, scattered, severely atypical spindle to polygonal cells, often arranged in small groups, having cytoplasmic vacuolations and fading cytoplasmic borders, should warranted a diagnosis of a suspected spindle cell neoplasia.[17] Additional findings such as non-degenerate neutrophilis, eosinophilis, increased erythrophagocytosis, EMH, and apoptotic leukocytes may further support the suspect of HES.[17]

The low sensitivity of cytology in the diagnosis of splenic lymphomas might be correlated to the specific distribution of tumor types in the spleen, where indolent nodular lymphomas (i.e. mantle cell lymphoma and marginal zone lymphoma) are frequent as it was in this caseload. These are nodular lymphomas composed of small to medium sized cells with minimal atypia and a low mitotic rate.[217,219] Thus, mantle cell lymphoma and marginal zone lymphoma can be easily misinterpreted as reactive lymphoid hyperplasia on cytology, and histopathology is often necessary considered that the definitive diagnosis relies on the evaluation of tumor architecture.[219]

Although not statistically significant, our results paralleled those of previous reports identifying higher cytological accuracy in the diagnosis of diffuse compared to focal lesions.[40,58,189]

One false positive diagnosis of neoplasia (i.e. lymphoma) was included in this study. In this case, a cytological diagnosis of lymphoma corresponded to a neutrophilic bacterial splenitis in association with severe marginal zone hyperplasia in histology (Figure 1 – E and F). In general, a false positive diagnosis of lymphoma is extremely rare especially in those tissues where lymphocytes do not represent a normal resident population or in tissues where nodular low grade B cell lymphomas are rare, such as in the mediastinum.[223,227] However, in the spleen of dogs the situation overturns and the cytological diagnosis of lymphoma should be reached with more caution. Indeed, splenic marginal zone hyperplasia is a common finding in

dogs,[219,223] and cytological sampling from these areas may result in a monomorphic specimen mimicking marginal zone lymphoma. This is a risk that pathologists have to bear in mind; thus the diagnosis of nodular low-grade lymphoma should be supported by histological evaluation of architectural changes, especially in dogs. A recent report has demonstrated a high overall concordance between histopathology, immunohistochemistry and PCR for antigen receptor rearrangement (PARR) in the diagnosis of marginal zone lymphoma, mantle cell lymphoma and lymphoid or complex nodular hyperplasia.[178] Therefore, further development of combined methods also applicable to cytological specimens may provide a less invasive and more valuable diagnostic approach to the diagnosis of splenic nodular lymphoid lesions.

Although histopathology is generally considered the diagnostic reference standard,[9,30,36,48,71] several limitations should also be considered for this technique in the diagnosis of splenic tumors. Specifically, the diagnosis of splenic hematomas and hemangiosarcomas is considered difficult, especially if spleens are not submitted entirely and if adequate samples from the margin of the lesion are not collected.[57,91,198,199] Noteworthy, splenic hematomas and HES may not be grossly distinguishable,[40,57,198,199] with the first possibly representing a component of the latter.[91,154] In this regards, it is noteworthy that during the publication process of the current study, another work was published, which suggested that at least 5 sections trimmed by an expert operator should be evaluated to avoid misdiagnosis of HES as a non neopalstic lesion.[91] Although this specific guideline was not included in the currrent work due to timing, it should be remembered that our inclusion criteria specified that for nodular lesions, the review of at least one slide containing at least one margin between the nodule and the adjacent splenic parenchyma was required to consider histological diagnosis as reliable.

The current study is characterized by several limitations, mainly due to its retrospective nature. One major limit was the inclusion of specimens obtained by different sampling techniques, with a high number of impression smears collected from both surgical biopsies and necropsies. Additionally, the inclusion of splenic cytological samples from necropsies and the University setting of this work may have

further biased the study toward cases with a more aggressive behavior and with features of malignancy easier to diagnose, as it might be also suggested by oru higher prevalence of neoplastic conditions compared to previously publsihed studies (Table 5). This may not reflect daily clinical practice in which FNA is the most common sampling technique to pre-operatively assess splenic lesions. Furthermore, as previously observed,[48,229] different sampling methods may have resulted in an improvement of sensitivity of cytology in this study, especially for those neoplasms characterized by low exfoliation rate. On the other hand, this observation can be viewed also in positive terms. Indeed, in a practical setting the preliminary evaluation of surgical biopsies or entire spleens by cytology prior to fixation could be implemented to facilitate the diagnosis and to reduce turnaround time. Additionally, this approach can provide pathologists with material useful not only for a preliminary diagnosis, but also for immunocytochemistry and for PARR on fresh specimens.

Despite this caseload being larger than previously reported ones, the small number of cases evaluated may explain the relatively wide confidence intervals observed around point estimates of our diagnostic accuracy indexes. Results might have been further biased by the inclusion criteria applied in the current study, leading to the exclusion of more than 90% cytological samples of canine spleen in our archives.

Unfortunately, full agreement with STARD guidelines could not be obtained in the current study since the type of treatment administered between cytological and histological sampling (item 22 of guidelines), and the incidence of adverse events following splenic sampling (item 25 of guidelines), could not be retrieved from our electronic archives.[27]

In conclusion, to the best of our knowledge, this is the first study conjunctively reporting overall accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of cytology in the diagnosis of canine splenic neoplasms compared to histopathology. Diagnostic accuracy indexes identified limitations of negative cytological results in excluding a dog to be truly free from neoplasia; however, high specificity and positive predictive value still highlighted cytology as a valuable tool in the diagnostic approach to splenic neoplasms. Nevertheless, our

findingss need to be further validated by prospective studies with sampling standardization.

CHAPTER 2 - INTEROBSERVER AGREEMENT AND DIAGNOSTIC ACCURACY OF CYTOLOGY IN THE PREDICTION OF FELINE LYMPHOMA IMMUNOPHENOTYPE

SCIENTIFIC PRODUCTION OF THE PH.D. CANDIDATE IN THE FIELD

The study described in the current chapter has been published as a Short Paper in February 2021 in the "Journal of Comparative Pathology", with the Ph.D. candidate as the first Author:

- Gambini M, Martini V, Bernardi S, Caniatti M, Gelain ME, Roccabianca P, Comazzi S (2021) Cytology of feline nodal lymphoma: low inter-observer agreement and variable accuracy in lymphoma immunophenotype prediction. *Journal of Comparative Pathology*, 184:1-6.

The manuscript can be found at the following link:
https://www.sciencedirect.com/science/article/abs/pii/S002199752100013X

INTRODUCTION AND AIMS

Lymphoma is the most common hematopoietic neoplasm in both dogs and cats.[214,216] According to the numerous common features shared with the human counterpart,[118] including the vast heterogeneity in subtypes and biological behaviors of non-Hodgkin forms, canine lymphomas has been subjected to numerous investigations up to date.[227] Conversely, literature regarding lymphomas in cats is pretty fragmentary despite the high incidence of this tumor in feline species.[138] Therefore, lymphoma in cats still represents a challenge in daily clinical practice.[214]

In both dogs and cats, lymphomas can be classified anatomically, histologically, immunophenotypically, and according to cell morphology.[216] Despite apparently highly time-consuming, this complicate algorhytm is fundamental considered that these features, in association with and in the frame of the WHO clinical staging, are correlated with response to therapy and clinical outcome.[216,217]

Most dogs affected by lymphoma develop often asymptomatic multicentric forms which are characterized by systemic lymphadenomegaly.[216,242] On the other hand, on the basis of the anatomical location, feline lymphomas can be distinguished in alimentary, mediastinal, extranodal (i.e. involving kidney, skin, central nervous system, eye, nose), multicentric (mainly involving spleen, liver, and peripheral or visceral lymph nodes), and leukemic forms.[32,74,88,138,182,216] Alimentary form followed by mediastinal and extranodal forms, are the most common clinical presentations of feline lymphoma,[62,74,113,138,182,238] and may further correlate with the age and the retroviral status of the cat.[113,227] In this context, conversely to dogs, it is noteworthy that multicentric lymphoma in cats is rarely correlated with systemic lymphadenomegaly which is in contrast mostly associated with reactive lymphadenopathy.[3,23,32,88,138]

Nodal lymphoma is most often diagnosed by cytology in dogs, being the most frequent clinical presentation of this tumor in this species.[216,242] In dogs, cytology is considered a reliable technique to diagnose lymphoma given the remarkable prevalence of high-grade cases, ranging from 73% to 75%.[70,160,183] Still, the classification of these tumours based on their cytological features is characterized by a variable inter-observer agreement ranging from fair to almost perfect, depending on the classification system applied.[208] Additionally, cytology investigations cannot evaluate tumor architecture, which represents a key feature in the establishment of the final diagnsosis.[160,218] Therefore, further laboratory analyses such as immunohistochemistry and/or flow cytometry, are generally required to confirm the diagnosis and determine immunophenotype and lymphoma subtype.[32] However, some specific cytological features have been described in dogs as suggestive of T or B cell origin including cells size, cytoplasmic colour and granules, nuclear shape, chromatin pattern, and number, size, and distribution of nucleoli.[70,160]

Differently from canine species, cytological diagnosis of lymphoma seems to be challenging in cats [23,32,214] and this may be particularly true for nodal lymphomas likely because specific diagnostic criteria have been poorly described in the literature due to the low prevalence of this disease presentation.[74,88,138] Additionally, cytological classification of feline lymphomas has been applied only on relatively small

caseloads of feline lymphomas,[39] with occasional association with the prognosis.[182] The low number of studies investigating cytological features of feline lymphomas might be further correlated to the high prevalence of abdominal and mediastinal forms and the frequent need of patient sedation, both factors that complicate the collection of samples.[106,122,214] The issue is further complicated by the fact that small cell lymphomas might mimic reactive lymphadenopathy, with the opposite evenience being reported too, posing a serious diagnostic challenge which may affect diagnostic accuracy of cytological evaluation.[23,32,106,165,189,214] In this context, histology and immunohistochemistry (IHC) still represent the reference standard for the diagnosis of lymphoma in cats.[32,214,228,238] Histological features are used to further classify feline lymphomas in subtypes,[228,238,239] with enteric T-cell lymphomas and diffuse large B cell lymphomas representing the vast majority of tumors.[227,239]

In dogs, lymphoma morphotype and immunophenotype has been clearly correlated with response to therapy and with the clinical otucome.[161,183,217] In this context, B-cell lymphomas are most common than T-cell ones, and the incidence of high grade cases overwhelms that of low-grade ones.[32,217,227]

Conversely, only few studies investigated correlations between lymphoma immunophenotype and clinical outcome in cats, reporting a variable prognosis according to immunophenotype and other features, mostly depending on the anatomical site investigated.[68,141,169,179,182,239] Most studies focused on alimentary lymphomas, reporting a worse prognosis for large granular lymphomas and B cell lymphomas compared to small cell T lymphomas.[68,141,169,239] Additionally, a worse prognosis was reported in general for high grade lymphomas compared to low grade ones.[182]

To the best of our knowledge, specific cytologic criteria similar to those described for dogs have not been applied to predict immunophenotype of feline nodal lymphomas, and no data are available on inter-observer variability in the assessment of cytological features in this species. In this context, the aim of this study was to evaluate the diagnostic performances of cytology in predicting phenotype of feline nodal and mediastinal lymphomas, both in terms of reproducibility and diagnostic accuracy. Specifically, reproducibility was assessed by calculating the inter-observer

agreement regarding immunophenotype prediction (i.e. B or T) as well as different cytological features which might be useful in the diagnosis. Furthermore, diagnostic accuracy was calculated for each examiner, using the results of histopathology and immunohistochemistry as the reference standard.

MATERIAL AND METHODS

Case selection

Databases of the Veterinary Teaching Hospital (University of Milan) and of the Department of Comparative Biomedicine and Food Science (University of Padua) were retrospectively mined for histopathological samples of feline lymphomas collected between January 2010 and January 2019. Other inclusion criteria besides definitive diagnosis of nodal or mediastinal lymphoma, were the availability of at least one Formalin-Fixed Paraffin-Embedded (FFPE) tissue block of the corresponding lesion and the availability of at least one good-quality matching cytological smear. The current study included the investigation of mediastinal masses since lymphomas in this site may also arise from mediastinal or sternal lymph nodes.[62]

For each histology-cytology pair included in the study, signalment information of the cats and anatomical location of the investigated lymphomas were recorded.

Immunohistochemistry

From each FFPE tissue block, at least 5 sections were cut: one was stained with Haematoxylin and Eosin (HE) and four sections were used for immunohistochemistry (IHC) utilizing primary antibodies directed against the following markers:

- CD20 - marker of mature B cells (epitope-specific rabbit antibody; Thermo Fisher Scientific Inc., Cheshire, UK; diluted 1:800);
- CD79 – marker of all maturation stages of B cells (monoclonal mouse anti-human, clone HM57; Dako Atlanta, Georgia, USA; diluted 1:100);
- CD3 – marker of T cells (mouse monoclonal, clone F7.2.38; Dako, Atlanta, Georgia, USA; diluted 1: 100);

- CD5 – marker of T cells (monoclonal mouse anti-human, clone SP19; Abcam, Cambridge, UK; prediluted - ready to use).

IHC was performed with an automatic immunostainer (Ventana Benchmark XT; Roche-Diagnostics, Monza, Italy). All reagents were dispensed automatically except for the primary antibody, which was manually dispensed.

Case evaluation

The diagnosis of lymphoma was confirmed in all cases by a board-certified pathologist (PR) based on routine HE histopathology and results of IHC, according to criteria reported in the literature.[222–224,228,238,239] IHC was further used to establish the immunophenotype of each lymphoma, which was then classified as of B-cell or T-cell origin. Those cases for which the immunophenotype could not be determined by IHC, were excluded from the study.

All cytological specimens had been air-dried and manually stained with May-Grünwald Giemsa (Merck KGaA, Frankfurt, Germany). Cytological samples included in the study were blindly evaluated by four cytologists with different experience: a board-certified clinical pathologist (evaluator 1), a board-certified pathologist (evaluator 3), and their respective Ph.D. students (evaluator 2 and 4, respectively; among which was included the Ph.D. candidate). All evaluators were aware of the final diagnosis of nodal or mediastinal lymphoma, but not of the subtype nor of the immunophenotype of the neoplasms. Before the beginning of the study, the evaluators conferred to standardize the description and the corresponding categorization of the morphological features that had to be evaluated for each cytological specimen. Morphological features evaluated were based on those used to evaluate canine lymphomas [70,160] with the addenda of other criteria such as cell homogeneity, presence of vacuoles and/or perinuclear halo, and presence of accessory non-neoplastic cells (Table 1). Despite the lack of supposed link with phenotype, the remaining morphological features were considered in the study since they are commonly included in cytological reports in daily laboratory practice.

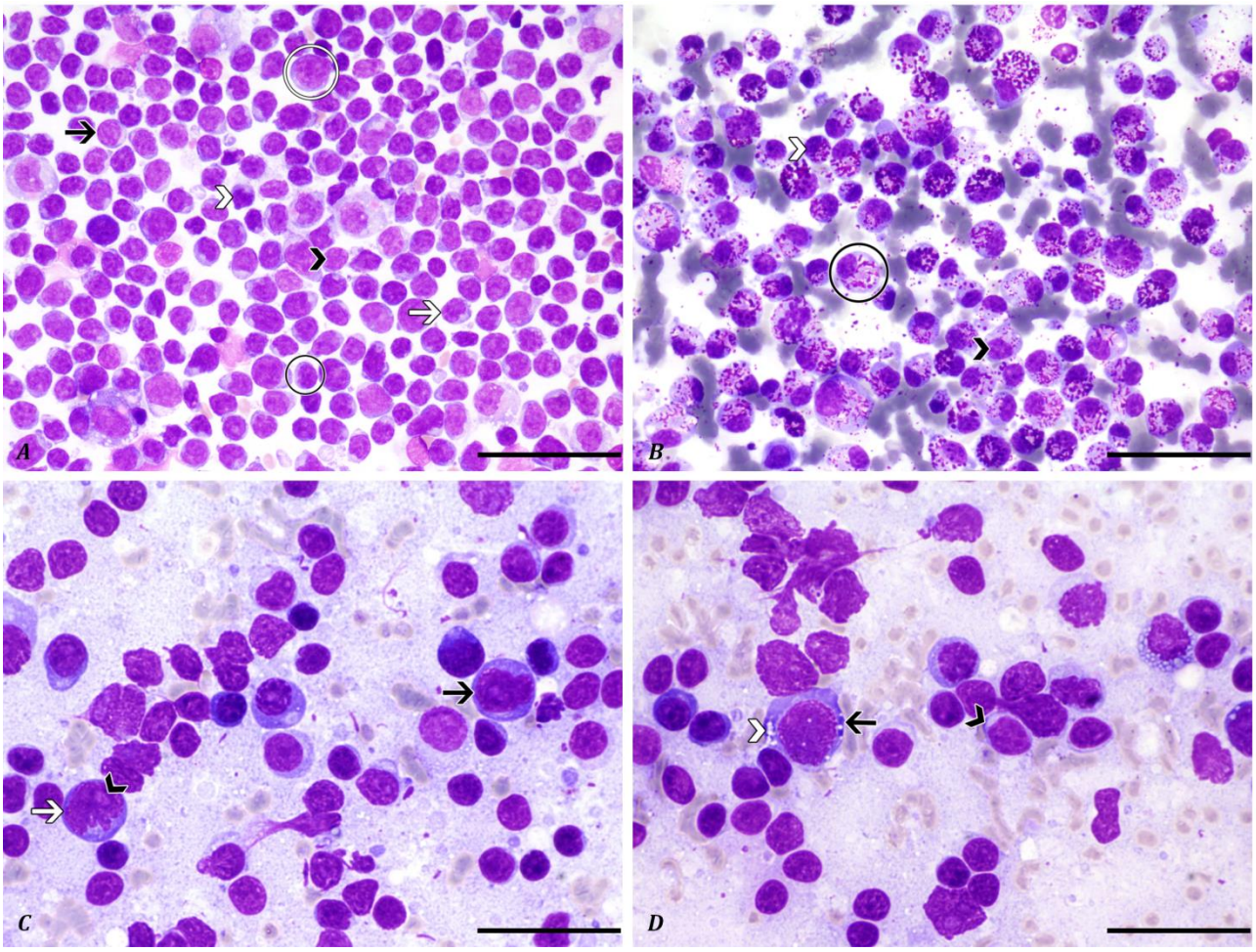
Table 1. Criteria used to describe cytological features of neoplastic cells in 25 samples of feline nodal lymphoma.

Cytological feature	Categories
Cell homogeneity	Homogeneous, heterogeneous *
Cell size	Small, medium, large **
Cytoplasm color	Slightly basophilic, deeply basophilic
Vacuoles	Present, absent #
Granules	Present, absent #
Perinuclear halo	Present, absent
Nuclear shape	Round, indented, irregular, convoluted
Chromatin	Homogeneous, partially clumped, clumped
Nucleolus	Visible, not visible
Number of nucleoli	Not determined, single, multiple
Mitoses (5 HPF)	None, 1 to 3, ≥4
Plasma cells (5 HPF)	None, 1 or 2, 3 or 4, ≥5
Eosinophils (5 HPF)	None, 1 or 2, 3 or 4, ≥5
Macrophages (5 HPF)	None, 1 or 2, 3 or 4, ≥5
Supposed phenotype	B, T

If not differently specified, all the features of the neoplastic cells were considered as present if noted in >50% neoplastic cells. Legend: HPF, high power field (i.e. 400x); *, cell homogeneity was defined as homogeneous if ≥80% cells were neoplastic; **, neoplastic cells were defined as small (nucleus <2 red blood cells), medium (nucleus between 2 and 3 red blood cells), or large (nucleus >3 red blood cells); #, vacuoles and granules were considered as present if noted in >20% neoplastic cells.

In general, according to current literature,[165] deeply bluish cytoplasm, perinuclear halo and round nucleus with visible nucleoli were considered as hints of B-cell phenotype, whereas slightly basophilic cytoplasm, cytoplasmic granules, indented, convoluted or irregular nucleus without nucleoli, and the presence of high numbers of plasma cells and eosinophils were considered as hints of T-cell phenotype. In those cases in which hints of both B-cell and T-cell phenotype were present, each evaluator was left free to decide upon the supposed phenotype on the basis of the current literature and of her/his own experience. Examples of morphological features evaluated in the current study according to the consensus reached by the evaluators before the beginning of the study, are illustrated in Figure 1.

Figure 1. Cytological smears of feline nodal and mediastinal lymphomas.



A) T-cell lymphoma. Homogeneous cell population composed by small (black circle) to medium (white circle) neoplastic cells, characterized by round (black arrow) to indented (white arrow) nucleus with homogeneous (black arrowhead) or partially clumped (white arrowhead) chromatin. May-Grünwald Giemsa (MGG). Bar, 50 µm. **B)** T-cell lymphoma. Homogeneous cell population composed by small to medium neoplastic cells, characterized by abundant slightly basophilic cytoplasm with variably abundant intracytoplasmic magenta granules (black circle), and partially clumped (black arrowhead) or clumped (white arrowhead) chromatin. MGG. Bar, 50 µm. **C)** B-cell lymphoma. Heterogeneous cell population including large neoplastic cells characterized by an irregular (black arrow) or convoluted nucleus (white arrow) containing prominent nucleolus (black arrohead). MGG. Bar, 33.5 µm. **D)** B-cell lymphoma. Heterogeneous cell population including medium to large neoplastic cells characterized by scant to moderately abundant, deeply basophilic cytoplasm (black arrow), perinuclear halo (black arrowhead), and intracytoplasmic clear vacuoles (white arrowhead). MGG. Bar, 33.5 µm.

Statistical analysis

Overall inter-observer agreement among the 4 evaluators for immunophenotype prediction and for each morphological feature evaluated (Table 1) was investigated calcualting free-marginal kappa coefficient using an online calculator (http://justusrandolph.net/kappa/). Kappa coefficients were interpreted according to

literature [107] as follows: ≤0, no agreement; >0.00 and ≤0.20, slight agreement; ≥0.21 and ≤0.40, fair agreement; ≤0.41 and ≥0.60, moderate agreement; ≥0.61 and ≤0.80, substantial agreement; ≥0.81 and ≤1.00, almost perfect agreement.

Considered that the different level of expertise of the evaluators might have influenced our results, the agreement between the two board-certified evaluators, between the two PhD students, and between each board-certified evaluator and the respective PhD student, was investigated calculating Cohen's Kappa coefficient with an online calculator (https://www.graphpad.com/quickcalcs/kappa1/). Pairwise agreement was investigated exclusively calculating Cohen's kappa (and not weighted kappa where applicable) to allow comparison of results for each morphological feature. Cohen's kappa coefficients were interpreted as described above.

Finally, the level of accuracy of each evaluator in immunophenotype prediction was investigated. Overall accuracy in correctly predicting lymphoma immunophenotype was calculated as follows: (number of correctly identified cases / total number of cases) * 100. Additionally, for each evaluator specific accuracy in diagnosing B-cell or T-cell lymphoma (i.e. sensitivity) was calculated as follows: [number of correctly identified (B cell OR T cell) cases / total number of (B cell OR T cell) cases) * 100. For overall accuracy and sensitivity values, 95% confidence intervals were calculated with an online software (https://www.medcalc.org/calc/diagnostic_test.php).

RESULTS

Included animals and tumors

Among cases retrieved from the archives, 36 were considered as eligible in the study. Among eligible cases, 5 were further excluded because at least one evaluator considered insufficient the quality of the cytological specimen. Additional 6 cases were excluded because the immunophenotype could not be determined based on IHC results (due to poor fixation and conservation of FFPE samples) or because the original diagnosis of lymphoma was reformulated as lymphoid hyperplasia. Therefore, 25 samples of nodal and mediastinal lymphomas were finally enrolled in the study. Location of lymphomas investigated in the current study, as well as

signalment information of cats from which they were obtained, are reported in Table 2.

Table 2. Signalment information of cats, and antomical location and immunophenotype of lymphomas investigated.

Case ID	Breed	Sex	Age	Anatomical location	IHC
1	ND	ND	ND	Mesenteric LN	B
2	ND	ND	ND	Mesenteric LN	B
3	ND	ND	ND	Mesenteric LN	B
4	ND	ND	ND	Abdominal LN	B
5	DSH	M	11 yy	L Retromandibular LN	B
6	Chartreux	NM	3 yy	R Retromandibular LN	B
7	Siamese	NF	9 yy	Submandibular LN	B
8	DSH	NM	12 yy	Popliteal LN	B
9	DSH	NF	11 yy	Multicentric (mesenteric and mediastinal LNs)	B
10	ND	ND	ND	LN NOS	B
11	ND	ND	ND	LN NOS	B
12	ND	ND	ND	LN NOS	B
13	ND	M	16 yy	LN NOS	B
14	DSH	NF	16 yy	L Retromandibular LN	T
15	DSH	NF	11 yy	Mesenteric LN	T
16	DSH	NM	13 yy	Mesenteric LN	T
17	DSH	NM	11 yy	Mesenteric LN	T
18	ND	ND	ND	Mesenteric LN	T
19	ND	ND	ND	Mesenteric LN	T
20	DSH	M	10 mm	Mediastinal mass	T
21	DSH	NF	2 yy	L Axillary LN	T
22	Maine Coon	NF	11 yy	Prescapular LN	T
23	DSH	M	10 yy	Popliteal LN	T
24	DSH	NM	ND	Multicentric (mesenteric and mediastinal LNs)	T
25	ND	ND	ND	LN NOS	T

Legend: DSH, Domestic Shorthair; L, left; LN, lymph node; M, intact male; mm, months; ND, not determined/not available; NF, neutered female; NM, neutered male; NOS, lymph node not otherwise specified; R, right; yy, years.

Most cats were Domestic Shorthair (11), other breeds represented by a single patient each were Chartreux, Maine Coon, and Siamese; breed was not available in 11 cases. Six (6) cats were neutered females, 5 neutered males, and 4 intact males; sex was not available for 10 patients. Mean age was 9.77 years (range: 10 months – 16 years; age not available for 11 cats). Most lymphomas were located in mesenteric or abdominal LNs (9), 5 in LNs not otherwise specified, 4 in retromandibular or

submandibular LNs, 2 in popliteal LNs, 2 were multicentric lymphomas simultaneously involving LNs of the abdominal and thoracic cavity, 2 involved an axillary and a prescapular LN each, and 1 lymphoma was identified as a mediastinal mass. The 25 lymphomas investigated in the current study were classified by means of histology and IHC as 13 (52%) of B-cell and 12 (48%) of T-cell origin.

Interrater agreement investigations

Cross tabulation of the results reported by each evaluator for each feature evaluated, by the results of the reference standard, are available upon request.

The results of the analysis of the overall inter-observer agreement on the morphological features of the 25 cases enrolled in the study are listed in Table 3.

Table 3. Inter-observer agreement among four evaluators in the assessment of cytological features of 25 samples of feline nodal lymphoma.

Cytological feature	Free-marginal Fleiss' kappa* (95% CI)
Supposed immunophenotype	-0.04 (-0.21-0.13)
Cell homogeneity	0.44 (0.24-0.64)
Cell size	0.23 (0.08-0.38)
Cytoplasm color	0.39 (0.15-0.63)
Vacuoles	0.87 (0.72-1.00)
Granules	0.91 (0.78-1.00)
Perinuclear halo	0.12 (-0.06-0.30)
Nuclear shape	0.41 (0.24-0.59)
Chromatin	0.14 (0.01-0.27)
Nucleolus	0.25 (0.06-0.46)
Number of nucleoli	0.24 (0.07-0.41)
Mitoses (5 HPF)	0.31 (0.12-0.50)
Plasma cells (5 HPF)	0.45 (0.29-0.61)
Eosinophils (5 HPF)	0.64 (0.49-0.80)
Macrophages (5 HPF)	0.37 (0.19-0.54)

* The coefficients were interpreted according to Landis and Koch (Landis and Koch, 1977): ≤0, no agreement; >0.00 and ≤0.20, slight agreement; ≥0.21 and ≤0.40, fair agreement; ≤0.41 and ≥0.60, moderate agreement; ≥0.61 and ≤0.80, substantial agreement; ≥0.81 and ≤1.00, almost perfect agreement. Highlighted in grey, those parameters for which agreement was judged as lacking or slight. Highlighted in yellow, those parameters for which agreement was judged as substantial or almost perfect. Not highlighted, those parameters for which agreement was judged as fair or moderate. HPF, high power filed (i.e. 400x).

No agreement among the 4 evaluators was observed when considering supposed immunophenotype. Excluding the latter, the level of agreement was fair to moderate for more than a half of parameters considered (9 out of 14, 64.29%), whereas it was almost perfect for the presence of cytoplasmic granules and vacuoles , substantial for the number of eosinophils, and only slight for the presence of a perinuclear halo and for the chromatin pattern.

Results of pairwise inter-rater agreement evluation between the evaluators are shown in Table 4. In general, the level of agreement largely varied among comparisons, being most commonly slight or fair, with the supposed immunphenotype always ranging between no and slight agreement. Specifically, when considering the two board-certified evaluators, only a slight to fair agreement was found for almost all the parameters considered excluding supposed immunophenotype (12 out of 14, 85.71%), whereas the agreement was almost perfect for the presence or absence of cytoplasmic vacuoles. Similar results were obtained when evaluating the agreement between the two PhD students (10 out of 14 parameters with slight to fair agreement, 71.43%), and between evaluator 1 and the corresponding PhD student (11 out of 14 parameters, 78.57%). Conversely, better agreement was obtained when comparing evaluator 3 and the corresponding PhD student, with a moderate to substantial agreement in 10 out of 14 features (71.43%). No agreement for the presence or absence of cytoplasmic granules was constantly observed for all the paired contrasts investigated.

Table 4. Inter-observer agreement among four evaluators in the assessment of cytological features of 25 samples of feline nodal lymphoma, paired contrasts.

Cytological feature	Cohen's kappa (95% CI)			
	Evaluator 1 versus evaluator 3 *(comparison between diplomates)*	***Evaluator 2 versus evaluator 4*** *(comparison between Ph.D. students)*	***Evaluator 1 versus evaluator 2*** *(diplomate – Ph.D. student comparison)*	***Evaluator 3 versus evaluator 4*** *(diplomate – Ph.D. student comparison)*
Supposed immunophenotype	0.030 (-0.258-0.318)	-0.066 (-0.454-0.323)	0.098 (-0.292-0.489)	0.179 (-0.097-0.455)
Cell homogeneity	0.138 (-0.185-0.460)	0.762 (0.516-1.000)	0.603 (0.301-0.906)	0.295 (-0.015-0.604)
Cell size	0.318 (0.066-0.571)	0.062 (-0.139-0.262)	0.293 (0.085-0.502)	0.536 (0.238-0.833)
Cytoplasm color	0.247 (-0.126-0.621)	0.277 (-0.100-0.653)	0.351 (0.001-0.700)	0.675 (0.384-0.967)
Vacuoles	0.816 (0.572-1.000)	0.818 (0.579-1.000)	0.911 (0.741-1.000)	0.715 (0.415-1.000)
Granules	0.000 (0.000-0.000)	-0.042 (-0.099-0.016)	0.000 (0.000-0.000)	-0.042 (-0.099-0.016)
Perinuclear halo	0.000 (0.000-0.000)	0.024 (-0.171-0.220)	0.186 (-0.244-0.616)	0.000 (0.000-0.000)
Nuclear shape	0.284 (0.004-0.564)	0.178 (-0.039-0.395)	0.154 (-0.111-0.418)	0.541 (0.271-0.810)
Chromatin	0.107 (-0.172-0.380)	-0.166 (-0.401-0.070)	0.333 (0.033-0.632)	0.447 (0.162-0.732)
Nucleolus	0.169 (-0.140-0.479)	0.302 (0.062-0.542)	0.149 (-0.206-0.503)	0.661 (0.319-1.000)
Number of nucleoli	0.121 (-0.151-0.394)	0.175 (-0.012-0.363)	0.032 (-0.145-0.209)	0.459 (0.133-0.785)
Mitoses (5 HPF)	0.191 (-0.101-0.483)	0.072 (-0.201-0.345)	0.483 (0.612-0.804)	0.192 (-0.081-0.464)
Plasma cells (5 HPF)	0.291 (0.050-0.531)	0.349 (0.112-0.586)	0.336 (0.086-0.587)	0.490 (0.238-0.742)
Eosinophils (5 HPF)	0.120 (-0.077-0.317)	0.353 (0.002-0.703)	0.052 (-0.052-0.157)	0.557 (0.255-0.859)

Macrophages (5 HPF)	0.230 (-0.094-0.555)	0.366 (0.117-0.616)	0.003 (-0.288-0.295)	0.633 (0.391-0.876)

Evaluators 1 and 3 were board-certified; evaluators 2 and 4 were their respective PhD students. * The coefficients were interpreted according to Landis and Koch (Landis and Koch, 1977): ≤0, no agreement; >0.00 and ≤0.20, slight agreement; ≥0.21 and ≤0.40, fair agreement; ≤0.41 and ≥0.60, moderate agreement; ≥0.61 and ≤0.80, substantial agreement; ≥0.81 and ≤1.00, almost perfect agreement. Highlighted in grey, those parameters for which no agreement was observed. Highlighted in light blue, those parameter for which agreement was judged as moderate. Highlighted in yellow, those parameters for which agreement was judged as substantial or almost perfect. Not highlighted, those parameters for which agreement was judged as slight or fair. HPF, high power filed (i.e. 400x).

Diagnostic accuracy in immunophenotype prediction

Overall accuracy for prediction of immunophenotype and sensitivity in diagnosing B-cell and T-cell lymphomas for each evaluator are reported in Table 5.

Table 5. Overall accuracy ans sensitivity for B-cell and T-cell lymphomas for four evaluators in the assessment of 25 samples of feline nodal lymphomas.

Evaluator	Overall accuracy for immunophenotype prediction (95% CI)	Sensitivity for B-cell lymphomas (95% CI)	Sensitivity for T-cell lymphomas (95% CI)
Evaluator 1 (ECVCP-diplomate)	36% (17.97-57.48)	30.77% (9.09-61.43)	41.67% (15.17-72.33)
Evaluator 2 (Ph.D. student of Evaluator 1)	56% (34.93-75.60)	46.15% (19.22-74.87)	66.67% (34.89-90.08)
Evaluator 3 (ECVP-diplomate)	64% (42.52-82.03)	92.31% (63.97-99.81)	33.33% (9.92-65.11)
Evaluator 4 (Ph.D. student of Evaluator 2)	76% (54.87-90.64)	69.23% (38.57-90.91)	83.33% (51.59-97.91)

Overall accuracy in correctly predicting lymphoma immunophenotype was calculated as: (number of correctly identified cases / total number of cases) * 100. Sensitivity in diagnosing B-cell or T-cell lymphoma was calculated as: [number of correctly identified (B cell OR T cell) cases / total number of (B cell OR T cell) cases) * 100. Highlighted in grey, the lowest value for each index. Highlighted in yellow, the highest value for each parameter. 95% CI, 95% confidence interval.

Results among evaluators were highly discrepant. Specifically, overall accuracy in correctly predicting the immunophenotype ranged from 36% to 76%. Sensitivity in diagnosing B cell lymphomas ranged from 30.77% to 92.31%. Finally, sensitivity in diagnosing T cell lymphomas ranged frm 33.33% to 83.33%. Unexpectedly, all the lowest values were reported for the 2 diplomates, while 2 out of the 3 highest values were reported for one of the Ph.D. students.

DISCUSSION

The current study investigated reproducibility among 4 evaluators and diagnostic accuracy of cytology in immunophenotype prediction of feline nodal and mediastinal lymphomas, revealing no agreement among cytologists and highly variable results of overall accuracy.

High levels of agreement were detected among the 4 evaluators exclusively when considering the number of eosinophils, and the presence or absence of cytoplasmic granules and vacuoles. Results regarding eosinophils and cytoplasmic granules were likely affected by the low prevalence of samples with these features. Indeed, eosinophils were detected in high numbers (≥3 in 5 HPF) only in 2 (8%) samples by three evaluators each. Similarly, granules were detected only in 2 (8%) samples (1 by 2 evaluators, and the other one by a single evaluator). This latter observation was further supported by the results of paired contrasts, which highlighted a constant lack of agreement for the presence or absence of granules. In this regards, it is noteworthy that evaluators 1 and 3 did not report in any case the presence of granules or perinuclear halo, respectively. This likely affected the results of the comparison with other evaluators concerning these features. In addition, the apparent discrepancy between Fleiss' k coefficient and Cohen's k coefficients concerning the presence of granules, might rely on the fact that the latter analysis was not performed for all the possible pairings of evaluators (specifically, both evaluator 2 and evaluator 3 reported the presence of granules in the same sample), as well as on statistical limits implicit in each of these two statistical tests.

Conversely, the almost perfect inter-observer agreement for the presence or absence of cytoplasmic vacuoles among the 4 evaluators was mirrored by the results of paired contrasts, which revealed a substantial to almost pefect agreement among all the pairings investigated. This observation might be the consequence that vacuoles were detected in 9 (36%) samples, and specifically in 6 cases by all 4 evaluators, in cases 2 by 3 evaluators, and 1 case by 2 evaluators. Interestingly, B-cell lymphomas were over-represented in the subset of samples for which cytoplasmic vacuoles were reported by the evaluators (7 cases out of 9, 78%). This finding might suggest the presence of cytoplasmic vacuoles as a feature with a possible role in the prediction of immunophenotype. Nonetheless, further studies on a larger scale are warranted to confirm this hypothesis.

The low kappa values obtained in paired contrasts between the board-certified evaluators and between the PhD students suggest that experience might not be not a leading factor influencing inter-observer agreement. Conversely, teaching centre

seems to be a major influencing factor. Indeed, higher kappa values were obtained between each board-certified evaluator and the respective PhD student rather than between the two board-certified evaluators and between the two Ph.D. students.

The causes for the low agreement values found among the observers need futher elucidations. It should be remembered that all the evaluators reached an initial consensus regarding cut off values to define each morphological feature trying to avoid as possible confounding factors correlated with their variable experience in hemato-oncology. In this context, one possible explanation for our results might lie in the fact that samples from feline lymphomas are often composed by heterogeneous populations of cells. Therefore, the choice of different microscopic areas during the evaluation by each evaluator could have strongly biased the results. This further underlines that cytology alone should be taken with caution to predict feline lymphoma immunophenotype and that laboratory testing remain mandatory for a correct final immunophenotyping.[32]

Our results suggest that the overall accuracy of immunophenotype prediction might be evaluator-dependent, varying from less than 40% to more than 75%. This observation is further supported by the highly discrepant results among evaluators when sensitivity was specifically calculated for the diagnosis of B-cell and T-cell lymphomas. Although infered from a very small caseload, these considerations highlight that the morphological features generally used in the dog for tentative immunophenotype prediction do not apply to feline nodal lymphoma, even if amended and implemented with additional criteria. Interestingly, an unexpected finding was that PhD students had a higher accuracy in immunophenotype prediction than their respective tutors. Young cytologists may be more prone to absorb knowledge from different schools, limiting in this way the bias determined by their limited experience. Furthermore, evaluators with a longer experience might have been biased by their former evaluation habits, despite the inclusion of an initial phase aiming to create a consesus regarding cut off values to define each morphological feature. Our results might be further influenced by the diffuse practice of not expressing a presumptive immunophenotype for feline lymphomas during routine cytological evaluation. This might have levelled the differences in the degree of

experience among the evaluators concerning this specific task. Anyway, the discrepant overall accuracy values in association with the unsatisfactory degree of interrater agreement observed, suggest that the morphological evaluation of nodal and mediastinal lymphomas in cats might not be useful at all in predicting B or T cell immunophentoype. Therefore, we highly recommend to apply immunophenotyping techniques such as immunohistochemistry or flow cytometry to define the immunophenotype of feline lymphomas. This is considered mandatory even for canine lymphomas,[32] although inter-observer agreement in dogs is higher than that observed in the current study.[208]

The major limitation of the current study was the low number of cases fulfilling the strict inclusion criteria and thus included in the investigations, which likely derived from the low prevalence of nodal lymphomas in cats, and from the lack of a consistent diagnostic approach for feline lymphoma which might have biased our retrieval rate from the archives. Another pitfall was the lack of intra-observer agreement evaluation, differently from previous studies regarding dog lymphomas.[208] Finally, the study was not conducted according to Standard Guidelines for Reporting of Diagnostic Accuracy Studies (STARD).[27] Nevertheless, all items reported in STARD guidelines were applied in our work except for two regarding case selection (item 9 and item 19) and two regarding clinical implications in involved patients (item 22 and item 25).[27]

In conclusion, a high inter-observer variability may affect the evaluation of morphological features of feline nodal lymphomas, thus preventing comparison of results among laboratories and even among different cytologists within the same laboratory. This variability might be even higher in a diagnostic routine setting, considered that in the current study the evaluators conferred before the beginning of the study to standardize the morphological criteria used to describe the cells, possibly enhancing the level of agreement. Also, including possible differential with non-lymphomatous lesions would likely result in lower agreement among evaluators, whereas in this study we only included samples with a final diagnosis of lymphoma. Our results confirm the limitations of cytology in the immunophenotyping of feline lymphoma and the mandatory need of histopathology and IHC as well as of

laboratory testing such as flow cytometry and PCR for antigen receptor rearrangements (PARR) for a correct interpretation of the immunophenotype.[14,32,85] Additionally, our observations should push veterinarians toward discussion and creation of a shared definition and classification of feline nodal lymphomas from a cytological point of view, which is currently lacking differently from what is reported for dogs.

CHAPTER 3 - APPLICATION OF CYTOLOGY IN THE EVALUATION OF NODAL METASTASIS IN CANINE MAST CELL TUMOR

SCIENTIFIC PRODUCTION OF THE PH.D. CANDIDATE IN THE FIELD

The results of preliminary investigations on cytological quantification of nodal mast cells in lymph nodes obtained from both dogs affected by non neoplastic pathological conditions and mast cell tumor-bearing dogs (**Section 1**) have been presented as an oral communication at the *Joint Congress ECVP-ESVP-ECVCP-ESVCP* at Burgers' Zoom in Arnhem, The Netherlands (25-28/09/2019):

- *Gambini M*, Buzzi G, Recordati C, Caniatti M, Giudice C, Pigoli C, Tecilla M. "Should I spend my time counting mast cells?" - preliminary results concerning mast cell quantification in cytological specimens for the evaluation of nodal metastasis in mast cell tumor-bearing dogs. *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019.* (abstract available at the following link: https://onlinelibrary.wiley.com/doi/abs/10.1111/vcp.12798)

The results of the same work will be soon submitted as a "Short Paper" to the "*Journal of Comparative Pathology*" for consideration of suitability for publication.

The results of the study focused on investigation of the interrater agreement and diagnostic accuracy of cytology for the evaluation of lymph nodes obtained from mast cell tumor-bearing dogs (**Section 2**) will be soon submitted as an "Original Research" to the journal "*Veterinary and Comparative Oncology*" for consideration of suitability for publication.

INTRODUCTION

Among canine round cell neoplasms, mast cell tumor (MCT) has a predominant role being the most common cutaneous neoplasm in dogs and representing around 20% of all canine skin neoplasms.[22,112,143,196,231] MCTs are neoplastic proliferations of mast

cells (MCs),[22,112] which are mainly known as some of the first responders to harmful situations for the organism, as well as for their key role in allergic reactions.[192] Nonetheless, the etiology of MCT is still unknown.[22,231]

Although the mean age of MCT-bearing dogs is 8-9 years, this tumor can be found also in younger animals.[112,231] Breed predisposition has been reported for Boxer, Boston Terrier, English Bulldog, Pug dog, Labrador Retriever, Golden Retriever, Cocker Spaniel, Schnauzer, Staffordshire Terrier, Beagle, Rhodesian Ridgeback, Weimaraner, Shar-pei, French Bulldog, and American Pit Bull Terrier.[53,112] On the other hand, sex seems not to be relevant for MCT incidence,[112,231] although neutering might represent a predisposing factor.[158]

MCT generally occur within skin or subcutis, with cutaneous tumors involving trunk and perineal region in 50% cases, hindlimb in 40% cases, and head and neck in the remaining cases.[112] Atypical sites of MCT occurrence include conjunctiva, salivary glands, rhinopharynx, larynx, oral cavity, ureter, and vertebral column.[59,93,112] The vast majority of affected dogs carry a single MCT, although 11-20% animals is characterized by simultaneous, multiple *de novo* tumors which need to be distinguished from locally-recurrent tumors or poorly-differentiated MCTs with satellite nodules.[22,112]

Macroscopic appearance of cutaneous MCT is highly variable, somehow mirroring Patnaik histopathological grading system.[22,112,153] Well-differentiated MCTs are generally single and characterized by slow growth, small size, and non-ulcerated, occasionally alopecic skin.[22,112,231] Poorly-differentiated MCTs have faster growth and increased size, are frequently ulcerated and itchy, and could give rise to satellite lesions in surrounding skin.[22,112] Subcutaneous MCT can range from solid masses to diffuse, ill-demarcated neoplasms, and thus might be misdiagnosed as lipomas at clinical inspection.[22,231] Visceral form of MCT has also been reported, being known as “disseminated mastocitosis” or “systemic mastocitosis”.[22,112] Nontheless, the latter condition is generally preceded by cutaneous or subcutaneous, biologically-aggressive MCTs,[112] although it can also occur independently from the presence of skin MCT.[205]

MCT biological behavior is extremely variable, ranging from cases that can be resolved with a simple surgical excision to others that are characterized by local recurrence and/or nodal and eventual distant metastasis.[22,94,112,196,231] Nonetheless, it is noteworthy that the two described above are just the extremes of a continuum of biological behaviors that characterize MCT.[112] Metastases are most frequently seen to the skin surrounding the primary mass and to the regional lymph node (RLN; i.e. the lymph node (LN) which is supposed to drain a regional lymphatic basin, and generally the closest to the primary mass), and can be followed by spreading to spleen, liver and bone marrow.[22,112,117,119,120,145,231,240]

Several attempts have been tried to find out specific features able to independently predict the biological behavior of canine MCTs.[22,94] The literature reports several factors that can influence MCT prognosis, with the latter defined as an increased risk of developing metastasis and subsequent negative effects on the survival time.[22,94,117,156,159,196] These prognostic factors include:

- clinical features, such as clinical stage, [22,59,93–95,104,108,112,117,145,158,201,230,240] surgical excision margins,[34,56,94,112,131,144,159] breed,[53,158] anatomic location,[22] clinical signs (including tumor size, fast growth, local irritation and inflammation such as “Darier’s sign”, local infiltration and demarcation from adjacent tissues, ulceration, presence of satellite nodules or paraneoplastic signs),[22,142] and clincial presentation (i.e. history of tumor recurrence);[22,94,159,162]

- histological features, including classification according to the 3-tier Patnaik grading system (grades 1, 2, and 3)[153] and to the 2-tier Kiupel grading system (low and high grade)[103] for cutaneous MCT, presence of specific histological characteristics as reported by Thompson et al. for subcutaneous MCT,[210] and mitotic index;[60,171]

- histochemical and immunohistochemical markers including Ki-67, expression pattern of KIT receptor (CD117), and expression of argyrophilic nucleolar organizing regions (AgNORs);[22,94,105,162]

- molecular features, such as c-KIT oncogen mutations status (including internal tandem duplications in exons 8, 9 and 11).[22,94]

Despite the availability of such a high number of prognostic factors, small animal clinicians still have to clash with a challenging decision-making in daily management of canine MCT, including the limitations to treatment imposed by costs that owners could not afford.[94,131,196,231] Indeed, none of the aforementioned prognostic factors is completely predicitive *per se* of MCT biological behavior. [112,196] Additionally, several therapeutic options are available, ranging from simple surgical excision (i.e. the most commonly performed treatment, especially for localized non-metastatic MCTs) to combinations of the latter with local adjuvant radiotherapy, and/or systemic chemotherapy or administration of tyrosine kinase inhibitors (TKI).[5,22,38,120,159,196] With this in mind, the best therapeutic approach should be established on a case-by-case basis, according to all the prognostic information which are available at that moment.[22,120,196]

Complementary to the application of histological and eventual cytological grading,[34,90,103,153,184,196] clinical staging (i.e. the process applied to deeply define the nature and the extent of a neoplastic disease)[22,231] has a key role in treatment planning for canine MCTs.[5,22,104,108,112,117,145,196,201,230,240] The staging system currently used for canine MCT is the one proposed by World Health Organization (WHO) in 1980.[150] This system is based on a TNM (i.e. *Tumor, Node, Metastasis*) scheme and is divided in 4 cathegories:

- I stage: single MCT confined to the dermis, without involvement of the RLN and further distinguished in:
    - Ia: absence of clinical signs;
    - Ib: presence of clinical signs;
- II stage: single MCT confined to the dermis, with involvement of the RLN and further distinguished in:
    - IIa: absence of clinical signs;
    - IIb: presence of clinical signs;

- III stage: multiple cutaneous MCTs or single, large, infiltrating MCT, with or without involvement of the RLN and further distinguished in:
    - IIIa: absence of clinical signs;
    - IIIb: presence of clinical signs;
- IV stage: any primary or recurrent MCT with distant metastasis (including blood and/or bone marrow involvement).

After the observation that the presence of multiple MCTs is not systematically correlated with a worse clinical outcome,[22,94,142,185,231] the original WHO staging system proposed in 1980 [150] has been questioned and amended by some authors.[94,104] Nonetheless, in current literature the terminology “stage II” is still used to indicate the presence of a MCT with confirmed LN metastasis.[65,104,105,119,159,196,201]

Although the extent of appropriate staging should be decided on a case-by-case basis, the RLN should always be checked for metastatic status (i.e. the presence or the absence of MCT metastasis).[22,104,108,159,196,230,234] Indeed, despite some contrasting reports, the presence of LN metastasis is nowadays generally recognized as a negative prognostic factor.[5,6,59,93,95,104,117,119,145,159,185,196,234] Specifically, surgical extirpation or local radiotherapy applied to metastatic LNs have been correlated to an improved clinical outcome [5,6,38,65,95,119,120,129] which could be the consequence of tumor burden reduction through the elimination of the LN as a reservoir for MC proliferation and/or spreading.[6,119] Additionally, considered the general assumption that MCT metastasize to the RLN before spreading towards distant locations or visceral organs,[22,112,117,145,231,240] the absence or presence of local LN infiltration by neoplastic MCs have been proposed as one of the criteria to stop or proceed with further staging, respectively.[22,196,230] The latter include radiographic and ultrasound investigations as well as cytological evaluation of blood smears and liver and spleen fine needle aspirates (FNAs).[67,156,159,196,200]

Excisional biopsy and histological evaluation of the entire LN is still considered the reference test (i.e. the gold standard) to evaluate nodal metastatic status in canine MCT.[6,71,104,108,119,145,150] In 2014, Weishaar et al.[234] proposed a novel classification

system to standardize histological evaluation and interpretation of LN metastasis in canine MCT. The system is based on 4 classes (HN0, HN1, HN2, and HN3) as showed in Table 1 which reports also the corresponding histopathological criteria and proposed interpretations.

Tab.1. Classes, histopathological criteria and corresponding proposed interpretation as reccommended by Weishaar et al. for histological classification of LN metastasis in canine MCT.

Class	Histopathological criteria	Proposed interpretation
HN0	0-3 discrete (isolated) MCs within sinuses (subcapsular, paracortical, or medullary) and/or parenchima per HPF	*Non metastatic*
HN1	>3 discrete (isolated) MCs within sinuses (subcapsular, paracortical, or medullary) and/or parenchima in at least 4 HPF	*Pre metastatic*
HN2	MC aggregates/clusters (≥3 associated MCs) within sinuses (subcapsular, paracortical, or medullary) and/or parenchima, or sheets of MCs within sinuses	*Early metastatic*
HN3	Discrete foci, nodules, sheets, or overt masses composed of MCs, disrupting or effacing the normal nodal architecture	*Overtly metastatic*

Legend: HPF, high power field (i.e. x400); MC, mast cell. Table adapted from Weishaar et al.[234]

Although the system by Weishaar et al. provided standard criteria for the uniformation of histopathological evaluation and showed to be significantly correlated with clinical outcome,[196,234] it is still affected by some limitations. Firstly, no guidelines were provided attempting to standardize the trimming of LN biopsies. Indeed, as reported by Herring et al.[92] in a study regarding maxillo-facial neoplasms other than MCT in dogs and cats, the histological examination of serial sections may improve sensitivity in detecting micrometastasis. Nonetheless, a recent study [66] suggested a protocol for serial sectioning at regular intervals (1.5 mm) of LNs from MCT-bearing dogs, although another study questioned that a similar approach might only arbitrarily allow the detection of individual MCs or arranged in small clusters.[120] Secondly, the criteria proposed by Weishaar et al. might still suffer from a certain degree of subjectvity in the evaluation (e.g. it is not reported how the detection of numerous couples of MCs should be interpreted),[120,234] thus requiring further sudies on interobserver agreement which have not been perfomed up to date. Additionally, as highlighted by the same Authors,[234] the system does not take into account the different microanatomical location (i.e. sinuses or parenchima) of MCs within the LN,

voiding the potential differences in biological behavior of MCs corresponding to different anatomical distribution. Furthermore, the lack of systematical evaluation of toluidine blue- or Giemsa-stained slides besides routine hematoxyllin and eosin, might have negatively influenced MC quantification.[174] Finally, as highlighted by the same Weishaar et al. and by successive investigators,[120,196,234] the validity of the interpretation provided for each class better represents certain histopathological features correlating with clinical outcome rather than a true progression trough the metastatic process. Therefore, future prospective studies with more accurate inclusion criteria including standardization of treatment protocols and follow-up procedures, and a larger number of cases better distributed among the 4 HN classes, are warranted.[234] Despite all the aforementioned issues, the histopathological system by Weishaar et al. is still diffusely accepted and used.[66,71,73,106,119,120,174,204]

Some recent studies suggested that the surgical extirpation of the RLN followed by histological examination is always useful due to lacking association between the LN size or its clinical status (i.e. non-palpable/normal-sized versus palpable/clinically enlarged) and the actual presence of metastasis.[6,65,66,108,145] Nonetheless, the evaluation of LN metastatic status prior to surgical extirpation continues to cover a key role in surgical treatment planning.[6,117,119,196,231] Indeed, LN extirpation is an invasive procedure which may not be feasible in all patients due to costs or adverse effects associated with surgery (such as the extension of anesthesia time).[66,73,108,145,213] The utility of needle core biopsy (NCB) in pre-surgical evaluation of canine and feline LNs was characterized by a low sensitivity, and with remarkable difficulties and potential side-effects correlated with the execution of this sampling technique.[108] Considered the above, the assessment of nodal MCT metastasis is firstly performed usually via cytological examination of FNA samples.[6,71,104,105,108,145,156,196,240] The use of this sampling technique is further supported by its low cost, safety, rapidity and feasibility also in first opinion practice.[71,106,112,117,145,189,235]

Despite the key role covered by cytology in the evaluation of LN metastatic status in MCT-bearing dogs, the evaluation and interpretation of these kind of cytological samples is still a nebulous and controversial subject.[145,196] Indeed, literature has

always been extremely fragmentary and reported highly variable, frequently subjective, and non-specific cytological criteria. These criteria mainly included semiquantitative estimates of individualized or aggregated MCs,[6,33,77,108,117,130,135,201,209,230] with occasional reference to morphometric cellular features.[117] This fragmentation led to diffuse lack of consistency and reproducibility of the data regarding the interpretation of LN involvement through cytology in canine MCT.[104,117] Therefore, in 2009 Krick et al.[104] proposed a cytological interpretative system for the evaluation of LN metastatic status in canine MCT, reporting quite codified cytological criteria and 5 proposed interpretation (normal, reactive lymphoid hyperplasia, possible metastasis, probable metastasis, and certain metastasis) as showed in Table 2.

Tab.2. Cytological criteria and corresponding proposed interpretation as reccommended by Krick et al. for the cytological evalaution of the RLNs in canine MCT classification of LN metastasis in canine MCT.

Cytological criteria	Proposed interpretation
No MCs observed	*Normal*
>50% small lymphocytes with a mixed population of prolymphocytes, lymphoblasts, plasmacells, and/or few to moderate numbers of macrophages, neutrophils, and eosinophils, and/or rare individual MCs	*Reactive lymphoid hyperplasia*
2-3 incidences of MCs aggregated in couples or triplets	*Possible metastasis*
>3 incidences of MCs aggregated in couples or triplets and/or 2-5 aggregates composed by >3 MCs	*Probable metastasis*
Effacement of lymphoid tissue by MCs, and/or aggregated, poorly differentiated MCs (pleomorphism, anisocytosis, anisokaryosis, and/or decreased or variable granulation), and/or >5 aggregates composed by >3 MCs	*Certain metastasis*

Legend: MC, mast cell. Table adapted from Krick et al.[104]

The system by Krick et al. was statistically correlated with the clinical outcome, defined as the median survival time (MST).[104] Specifically, it can be desumed that dogs with “normal” LNs and with “reactive lymphoid hyperplasia” did not show any singnificant difference in MST compared to dogs with “possible metastasis”. On the other hand, dogs with “possible metastasis” had a longer MST than dogs with “certain metastasis”, which in turn did not show any difference in MST compared to dogs with “probable metastasis”. Nonetheless, the correlation of the system by Krick et al. with the clinical outcome might have been affected by the marked variability in

treatment protocols adminsitered to dogs enrolled in the study.[104] This hypothesis is supported by brief communication by Fournier et al.[72] who observed a statistically signifcant difference in the one-year metastasis-free survival rate exclusively between dogs with “reactive lymphoid hyperplasia” and dogs with “certain metastasis”. On the other hand, it is noteworthy that Fournier et al.[72] monitored MCT-bearing dogs only for one year, while in the study by Krick et al.[104] the follow-up information for the patients was available up to 10 years after cytological sampling of LNs. Although the intepretative system proposed by Krick et al.[104] represented a key turning point in the cytological evaluation of nodal samples from MCT-bearing dogs,[196] successive reports did not systematically applied [6,71,106,135,159,201,230,240] or did modify this system,[71,145,174,180] with few exceptions.[105,119,156,204] The main differences in the design of studies reported in the literature and which included the cytological evaluation of LNs sampled from MCT-bearing dogs, are summarized in Table 3.

Tab.3. Summary of the main differences in the study design of studies which included the cytological evaluation of LNs sampled from MCT-bearing dogs.

Study	Study design (# dogs)	RLN or SLN (# LN examined)	Cytological criteria and classes	Associated histological evaluation (application of criteria by Weishaar et al. Y/N)	Cytological sampling technique
Langenbach et al. (2001)[108]	P (7)	RLN (7)	1) NM: all cases excluded from “M” class 2) M: moderate to high numbers of atypical MCs, moderate to high numbers of clusters composed by >3 MCs, high to very high numbers of normal MCs without evidence of hyperplasia/inflammation	yes (N)	FNA
Gieger et al. (2003)[77]	R (24)	RLN (19)	1) NM: all cases excluded from “M” class 2) M: high numbers of well-differentiated MCs, presence of poorly-differentiated MCs	no	FNA
Cahalane et al. (2004)[33]	R (68)	RLN (17)	1) NM: all cases excluded from “M” class 2) M: increased numbers of MCs in >2 consecutive HPF	yes (N)	FNA
Thamm et al. (2006)[209]	R (61)	RLN (28*)	1) NM: single, scattered MCs 2) M: sheets or aggregates of MCs	no	FNA

Study	Study design (# dogs)	RLN or SLN (# LN examined)	Cytological criteria and classes	Associated histological evaluation (application of criteria by Weishaar et al. Y/N)	Cytological sampling technique
Marconato et al. (2008)[117]	P (27)	RLN (27)	1) NM: criteria not specifically reported 2) Inconclusive for M: occasional MCs 3) M: effacement of lymphoid tissue by MCs	no	FNA
Krick et al. (2009)[104]	R (152)	RLN (152)	1) Normal: no MC 2) Reactive lymphoid hyperplasia: scant, single MCs 3) Possible M: 2-3 couples/triplets of MCs 4) Probable M: >3 couples/triplets of MCs or 2-5 aggregates composed by >3 MCs 5) Certain M: effacement of lymphoid tissue by MCs and/or aggregated, poorly differentiated MCs (pleomorphism, anisocytosis, anisokaryosis, and/or decreased or variable granulation), and/or >5 aggregates composed by >3 MCs	no	FNA
Miller et al. (2016)[135]	R (94)	RLN (69)	1) NM: all cases excluded from "M" class 2) M: >3% MCs on the total population or >20 MCs per HPF showing at least 3 atypias, or clusters of MCs	no	FNA
Warland et al. (2014)[230]	R (220)	RLN (119)	1) NM: all cases excluded from "M" class 2) M: increased number of MCs (>3% or >20 total MCs) with at least 1 of the 3 following features: highly variable or generally reduced granularity, moderate to marked anisocytosis, moderate to marked anisokaryosis, increased N:C ratio, nuclear/nucleolar pleomorphism or altered number of nuclei/nucleoli, presence of eosinophils or aggregates of MCs	no	FNA
Worley (2014)[240]	P (19)	RLN (3) + SLN (6)	1) NM: all cases excluded from "M" class 2) M: effacement of lymphoid tissue by MCs	yes (N)	FNA
Baginski et al. (2014)[6]	R (90)	RLN (34)	1) NM: all cases excluded from "M" class 2) M: prevalence (>50%) of MCs, numerous aggregates of MCs, and/or poorly differentiated MCs	no	FNA

Study	Study design (# dogs)	RLN or SLN (# LN examined)	Cytological criteria and classes	Associated histological evaluation (application of criteria by Weishaar et al. Y/N)	Cytological sampling technique
Stefanello et al. (2015)[201]	R (386)	RLN (386)	1) NM: well-differentiated MCs, well-granulated and without atypia 2) M: sheets or aggregates of MCs, high numbers of MCs, atypical MCs	yes (N)	FNA
Ku et al. (2017)[106]	R (50*)	RLN (50*)	• Criteria proposed by Krick et al.**	yes (N)	FNA, SCRA
Krick et al. (2017)[105]	R (64)	RLN (59)	• Criteria proposed by Krick et al.	yes (N)	FNA
Mutz et al. (2017)[145]	P (20)	RLN (23)	1) Normal: as proposed by Krick et al. 2) Lymphoid reactive hyperplasia: as proposed by Krick et al. (2009) 3) Possibile M: as proposed by Krick et al. (2009) 4) Certain M: as proposed by Krick et al. (2009)	yes (N)	FNA
Sapierzynski et al. (2017)[180]	R (40)	RLN (40)	1) Low M involvement: aggregates of >3 MCs 2) Medium M involvement: small groups of neoplastic cells in nearly all microscopic fields 3) Massive M involvement: prevalence (>50%) of MCs	no	FNA
Fournier et al. (2018)[71]	R (110)	RLN (78)	• Criteria proposed by Krick et al.**, ***	yes (Y)***	FNA
Pizzoni et al. (2018)[159]	P (45)	RLN (12)	1) NM: all cases excluded from "M" class 2) M: sheets or aggregates of MCs, high numbers of MCs, atypical MCs	no	FNA
Marconato et al. (2018)[119]	R (152)	RLN (71)	• Criteria proposed by Krick et al.	no	FNA
Sabattini et al. (2018)[174]	P (24)	RLN (28)	• Criteria proposed by Krick et al.***	yes (Y)	FNA
Sulce et al. (2018)[204]	P (38)	RLN (16)	• Criteria proposed by Krick et al.	yes (Y)	FNA
Pecceu et al. (2019)[156]	P (82)	RLN (55)	• Criteria proposed by Krick et al.***	yes (Y)**	FNA
Fournier et al. (2020)[73]	P (59)	SLN (54)	• Criteria proposed by Krick et al.	yes (Y)	FNA

Legend: FNA, fine needle aspiration; HPF, high power field; MC, mast cell; M, metastatic; N, no; NM, non metastatic; N:C ratio, nuclear-cytoplasmic ratio; P, prospective study; R, retrospective study; RLN, regional lymph node; SCRA, scraping; SLN, sentinel lymph node; Y, yes; *, the exact number of dogs or lymph nodes cannot be extrapolated; **, criteria not systematically applied to all cases; ***changes applied to the original system but not reported, or not clear explanation of which interpretative classes were considered as negative and positive for metastasis.

Although Krick et al. suggested to investigate the interrater agreement of their interpretative system,[104] only a single study [174] included the cytological evaluations of three independent observers regarding metastatic LNs (i.e. histologically classified as HN2 and HN3) obtained from MCT-bearing dogs. Nonetheless, Sabattini et al.[174] did not investigated interrater agreement in details, limiting to report that the agreement among evaluators was negatively influenced by the use of a rapid cytological stain in the case of smears characterized by scant, poorly-granulated MCs. The need of systematic interobserver agreement evaluations is further supported by the observation that the criteria by Krick et al. might not be perfectly complete, omitting the interpretation of findings frequently encountered during the cytological evaluation of nodal specimens obtained from MCT-bearing dogs. Specifically, cytological criteria by Krick et al. do not report how to interpret the presence of a single couple or triplet of MCs, of moderately increased numbers of single, well-differentiated MCs, or of scant, non-aggregated atypical MCs. Regarding the latter finding, it is noteworthy that in a normal or reactive lymph node, no atypical MCs should be detected.[112,117] Previous literature referring to dogs has defined atypical MCs as cells characterized by at least one of the following morphological features: variable to scant granularity, anisocytosis, anisokaryosis, bi- or multi-nucleation, nuclear pleomorphism.[34,90,104,117,184] Sulce et al.[204] also observed that compared to IgE-positive MCs, IgE-negative neoplastic MCs can be characterized by low numbers of larger metachromatic granules, interpreting this finding as indicative of immaturity or genetic mutations. Additionally, a considerable lack of consensus among cytologists in the use of modifiers (such as “possible” or “probable”) for cytological diagnoses in general has been reported,[42] especially regarding the numerical percentage probability of an accurate diagnosis implied by each modifier. Similarly, it cannot be excluded that the use of generic, unspecific quantifiers in the criteria proposed by Krick et al.,[104] such as the one referring to “rare” MCs in lymphoid reactive hyperplasia as well as that referring to the “effacement of lymphoid tissue by MCs” in certainly metastatic samples, might lead to a reduced reproducibility of cytological interpretations among different readers. Another critical point included in the system by Krick et al.[104] is the one represented by the required absence of MCs in a cytological sample to define it as “normal”. This criterion seems to be in contrast with

other studies reporting that a variable amount of this cells can be detected in normal lymph nodes,[12,25,112,117,130,145,165,174,230] and might lead to decreased reproducibility of results among different cytologists.

The aforementioned observations regarding the critical points included in the cytological interpretative system proposed by Krick et al.,[104] as well as the complexity of the final categorization scheme in five intepretations, might explain the reason why successive studies did not systematically applied or did modify this system. In their work, Mutz et al.[145] grouped together the classes “probable metastasis” and “certain metastasis” proposed by Krick et al., due to the lack of statistically significant differences between the corresponding survival times. Another strength of the study by Mutz et al. was the quantification of MCs in cytological nodal samples from healthy, allergic, and MCT-bearing dogs.[145] Indeed, no clear consensus exists regarding the quantity of MCs that are regularly present in normal or reactive canine LNs.[12,25,112,117,145,174,230] The studies which focused on the quantification of nodal MCs in dogs without health issues, affected by non-neoplastic conditions such as allergies, or bearing MCT [12,25,117,145,174,204] have been conducted applying different inclusion criteria, counting methods, and statistical analysis, thus limiting the comparison of obtained results. The main differences in the study design of these works are summarized in Table 4.

Tab.4. Summary of the main differences in the study design of studies aiming to quantify MCs in lymph nodes of dogs belonging to different clinical categories.

Study	Study design	Counting method (magnification)	Clinical categories (# dogs/cases)
Bookbinder et al. (1992)[25]	P	#MC/entire slide (100x)	Healthy (56)
Marconato et al. (2008)[117]	P	#MC/2000 cells – expressed as percentage least mean square (400x)	Healthy (4) inf/non-npl (31) MCTDs (27)
Bauer et al. (2011)[12]	P	“Fields”: #MC in 20 fields (200x)	Healthy (30) Allergic (20)
		“Cells”: #MC/500 lymphoid cells (400x)	
		“Total”: #MC/entire slide (100x)	

Study	Study design	Counting method (magnification)	Clinical categories (# dogs/cases)
Mutz et al. (2017)[145]	P	#MC/2000 cells (n/a)	Healthy (20) Allergic (20) MCTDs (20)
Sabattini et al. (2018)[174]	P	"Total": #MC/entire slide - expressed only as % (n/a)	Npl ot MCT (n/a) MCTDs (28)
Sulce et al. (2018)[204]	P	#MC/1000 cells – expressed only as % (n/a)	MCTDs (12)

Legend: Allergic, dogs affected by allergic dermatopathy; inf/non-npl, dogs affected by infectious/inflammatory non neoplastic conditions; MC, mast cells; MCTDs, mast cell tumor-bearing dogs; Npl ot MCT, neoplasms other than mast cell tumor; n/a, non-applicable/non-available; P, prospective; R, retrospective.

Specifically, the studies by Bookbinder et al.[25] and by Bauer et al.[12] reported that a certain amount of MCs can be present in LN obtained from healthy dogs, ranging from 1 to 16 and from 0 to 13, respectively. Additionally, Bauer et al. reported that the number of MC in LNs from allergic dogs was significantly higher than that in LNs from healthy dogs.[12] To the best of our knowledge, only four studies included the quantification of MCs in LNs from MCT-bearing dogs.[117,145,174,204] In the study by Marconato et al.,[117] dogs with definite MCT nodal metastasis had a significantly higher number of nodal MCs than healthy dogs, dogs affected by non-neoplastic infectious-inflammatory conditions, and MCT-bearing dogs without nodal metastasis and with doubtful nodal metastasis. In the study by Mutz et al.,[145] MCT-bearing dogs were more likely to present at least one MC or >0.05% MCs within LNs than healthy dogs, but not when compared with animals affected by allergic dermatopathy. Sabattini et al.[174] applied semi-quantitivative estimation of MCs in FNAs obtained exclusively from LNs belonging to MCT-bearing dogs and histologically classified as HN2 or HN3, finding variable proportions of MCs ranging from <10% to >50%. Finally, in the study by Sulce et al.[204] variable amounts of MCs ranging from 0 to about 89% over 1000 total cells, were oserved in MCT-bearing dogs, well correlating with the number of intranodal MCs estimated with flow-cytometry. Besides these findings, it is noteworthy that in a brief communication Fournier et al.[72] reported that LN samples classified in a new cateogry named "high-risk metastasis" which was characterized by increased MC nuclear-cytoplasmic ratio and the presence of >8% of MCs among all nucleated cells, showed a better correlation with one-year metastasis-free survival rate compared to that observed for the system proposed by Krick et al.[104]

To correctly interpret the results of the studies cited above, it should be remembered that it is not clear whether an increased number of well-differentiated MCs in LNs of MCT-bearing dogs should be attributed to the chemoattractive effect of the primary mass microenvironment on reactive cells or to an effective migration of neoplastic MCs draining from a well differentiated primary mass.[99,104,108,119,201,230] Additionally, besides the aforementioned difficulties in the comparison of studies focusing on the quantification of nodal MCs in different clinical categories of dogs, it is noteworthy that no easily-evaluable morphological features are available to distinguish well-differentiated neoplastic MCs from reactive or inflammatory ones.[105,108,119,120,145,196,201] Indeed, although Marconato et al. reported that nodal MCs from MCT-bearing dogs with doubtful nodal metastasis and with definite nodal metastasis were characterized by specific nuclear morphometric alterations,[117] the application of morphometry in a routine diagnostic set would sensibly increase turnaround time. The issue is further complicated by the observation that nodal macrophages engulfing intracytoplasmic metachromatic granules released by MCs may mimic the presence of the latter cell type,[167] and that scant, hypogranular MCs might be misdiagnosed as histiocytes or large granular lymphocytes.[174]

Although all the considerations above, neither Mutz et al.[145] nor Sulce et al.[204] proposed to include some kind of MC semi quantitative or quantitative evaluation in the interpretative system previously proposed by Krick et al.[104] Still, despite the suggested simplification of the system by Krick et al., in their study Mutz et al. maintained the separation between the two interpretative classes "normal" and "reactive lymph node hyperplasia", despite the absence of a statistically significant difference in the survival time, as can be extrapolated from the original work by Krick et al.[104] In this view, subsequent studies grouped together the first two interpretative classes of the system by Krick et al. in a single category indicative of lack of nodal MCT metastasis.[71,73,174]

Considered the aforementioned issues and the fact that cytology is often used alone to investigate LN metastatic status before the establishment of therapeutic protocols in canine MCT in the clinical practice, potential causes of misinterpretation of cytological specimens leading to tumor over- or under-staging should be

investigated.[104,106,196] In this frame, although the interpretative system proposed by Krick et al. correlated with the clinical outome, comparison of cytological findings with histopathology is warranted, given that the latter represents the reference test.[6,71,84,92,104,105,108,119,145,150,196] Despite the suggestion by the same Krick et al.,[104] only few studies directly or indirectly investigated the agreement among cytological and histological findings in detecting MCT nodal metastasis.[71,73,105,106,108,145,174,201,204,240] Unfortunately, the comparison among the results of these studies is limited by several differences in the study design (see Table 3) since each of them did not meet one or more of the following parameters: systematic application of cytological criteria proposed by Krick et al.,[71,106,108,201,240] systematic application of the histopathological criteria proposed by Weishaar et al.,[71,105,106,108,145,201,240] exposition and clarification of the changes applied to the interpretative cytological system by Krick et al.[71,174] and/or to the histological system by Weishaar et al.[71] The only two studies which systematically applied the cytological interpretative system by Krick et al.[104] and the histological one by Weishaar et al.[234] were those by Sulce et al.[204] and by Fournier et al.,[73] which unfortunately were limited by the inclusion of a relatively low number of LNs (6 and 21, respectively) for which both cytological and histological diagnosis were available. Therefore, further investigations are required to establish more standardized and feasible criteria for LN cytological assessment, that would correspond as much as possible to the histological classification by Weishaar et al.[106,145,230,234] In this view, diagnostic accuracy studies conducted according to the Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines [27] might represent a valid alternative.

The need of validation of standardized cytological criteria for the evaluation of nodal metastatic status in MCT-bearing dogs is further strengthened by the growing attention toward LN mapping techniques in veterinary medicine.[10,13,20,31,66,73,75,84,114,172,197,213,230,240] Indeed, the individuation of the sentinel LN (SNL, i.e. the first LN in a specific lymphatic drainage path actually receiving lymph, and the eventual metastatic spread, from a primary tumor) can provide irreplaceable information about the state of invasion of the lymphatic system by the tumor.[54,73,196,213,230] Only a limited number of studies have focused so far on the exact mapping of lymphatic drainage of different anatomical sites in healthy dogs.[63,203] The

issue is further complicated by the observation that not all RNL can be located or palpated, especially due to their small dimensions or their anatomical site (e.g., prescapular, axillary, inguinal, sacro-iliac LNs), and thus it is not always possible to collect FNAs from these sites.[73,156,172,230] In this context, the common belief that distant MCT metastases have no reason to be suspected without first spreading to the RLN [230] has recently been questioned due to the observation that MCT can metastasize to one or more LNs different from the expected RLN.[66,71,73,84,119,159,195,201,240] Indeed, the neoplastic diffusion can "skip" the expected RLN metastasizing in favor of one or more unpredictable LNs, probably due to tumor-induced lymphangiogenesis, increased resistance to lymphatic flow through some draining paths compared to others, or marked extension of specific lymphatic drainage beds such that associated with the oral cavity.[54,71,73,99,172,195,240] In the light of the above considerations, it is obvious that the efficacy of clinical staging would always be improved evaluating the metastatic status of the actual SLN rather than the expected RLN.[66,73,84,196] Therefore, several studies investigated various techniques for the mapping of SLN in MCT-bearing dogs.[31,66,73,84,114,172,240] Besides the undoubted improvement of clinical staging, the surgical excision of LNs guided by SLN mapping could also avoid the surgical extirpation of the RLN when it is not necessary (i.e. in those cases where the RLN and the SLN do not coincide) and could reduce all post-operative complications that may associate with massive LN excision, especially of intracavitary ones (such as surgical wound infections, seroma formation, postoperative pain, reduced mobility, lymphedema and compromised immune response).[66,71,73,84,117,195,240] In this context, the individuation of the SLN in canine MCT might open new perspectives for the use of cytology in an intraoperative environment also in veterinary medicine. Indeed, similarly to human medicine,[50,96,241] real-time cytological evaluation of surgically extirpated SLNs sampled via touch imprint and/or scraping might be helpful in further reducing the surgical dose and the side effects associated with routine, unguided, and extensive LN extirpation.[71,84,117,195,240] Specifically, several issues associated with the application of SLN mapping techniques might beneficiate from the use of intraoperative cytology as an additional guide for the surgeon. Firstly, Grimes et al.[84] reported that SLNs without histological evidence of metastasis were constantly associated with non

metastatic non-SLN, supporting the hypothesis that the SLN might represent a kind of “gate” for MCT metastasis toward other LNs. In this context, intraoperative cytology might be useful in the choice of removing also the RLN in those cases in which it does not coincide with a SLN cytologically suggestive of the presence of metastasis. Furthermore, a previous study [66] regarding canine MCT reported that multiple SLNs were identified within the same lymphocentrum (i.e. a LN station draining a certain lymphatic basin). Despite the application of dynamic tracking of the contrast medium in this and other studies,[66,73,172] it cannot be completely excluded that this finding might be the consequence of an irregular distribution of the radiotracer and/or of the intraoperative dye, as previously reported in human beings.[29,80,111,188] This observation, together with a desirable application of the “sentinel chain” concept also in veterinary medicine,[37] point out how in the future the use of intraoperative cytological evaluation of each SLN excised from a lymphocentrum might guide the surgeon in the choice of which other LNs extirpate from the same lymphocentrum. Finally, the same study by Ferrari et al.[66] reported that both methylene blue injection and lymphoscintigraphy identified as the SLN a subcutaneous structure lacking the classical features of a LN at histopathology. As previously reported in dogs and rats,[73,194,203] this finding might be the consequence of a rearrangement of the draining lymphatics following previous excision of another tumor, which might work as new metastatic pathways for eventual residual cancer cells.[66,203] In this specific event, cytology might be useful in providing an immediate feedback to the surgeon who could then proceed further with the detection and extirpation of the real SLN or, eventually, of the RLN if a SLN could not be detected. Considered all the aforementioned issues, further studies aiming to validate the reproducibility and accuracy of the cytological evaluation of LNs in MCT-bearing dogs are warranted to investigate the reliability of this diagnostic technique as an additional supporting intraoperative guide for the surgeon.

AIMS

Considered the issues highlighted in the introduction of the current chapter, the work of the Ph.D. candidate in this field was focused on prelimary investigations on the cytological quantification of nodal MCs in LNs obtained from both dogs affected by

non neoplastic pathological conditions and MCT-bearing dogs, as well as on the investigation of the interrater agreement and the diagnostic accuracy of cytology for the evaluation of LNs obtained from MCT-bearing dogs.

Specifically, the preliminary investigations on the cytological quantification of nodal MCs in LNs obtained from both dogs affected by non neoplastic pathological conditions and MCT-bearing dogs, which results are showed in **Section 1**, aimed to verify whether the number of nodal MCs is able to indipendently predict the presence or absence of MCT nodal metastasis in canine cytological samples, disregarding the sampling method applied to obtain cytological specimens. Specifically, the validity of the following hypotheses was investigated:

1) The number of nodal MCs is significantly higher in MCT-bearing dogs compared to dogs affected by non neoplastic conditions;
2) The number of nodal MCs is significantly different among MCT-bearing dogs without LN metastasis, MCT-bearing dogs with possible LN metastasis, and MCT-bearing dogs with certain LN metastasis;
3) The number of nodal MCs is not significantly different among cytological LN specimens obtained from MCT-bearing dogs via fine-needle aspiration, scraping, or touch imprint.

On the other hand, the study focused on the investigation of the interrater agreement and the diagnostic accuracy of cytology for the evaluation of LNs obtained from MCT-bearing dogs, which results are showed in **Section 2**, aimed to verify the following hypotheses:

1) The application of the cytological interpretative system proposed by Krick et al. (Krick 2009) is characterized by inadequate interrater agreement, both when applied according to the 5 classes proposed by the original Authors and when simplified in a 3-classes version;
2) The diagnostic accuracy of a 3-classes simplified version of the cytological interpretative system proposed by Krick et al. is moderate and its morphological criteria might need to be improved;

3) A 3-classes and a 2-classes amendment of the cytological interpretative system by Krick et al., including criteria for nodal MCs quantification, are characterized by improved interrater agreement and diagnostic accuracy compared to the original interpretative system.

<u>SECTION 1</u> - Preliminary investigations on the cytological quantification of nodal mast cells in dogs affected by non neoplastic pathological conditions and mast cell tumor-bearing dogs

MATERIAL AND METHODS

Case selection from the archives and preliminary classification into clinical categories

Canine cytological nodal samples collected between January the 1st, 2016 and December the 31st, 2018 were retrospectively recovered from the electronic database of the Diagnostic Pathology Service of the Department of Veterinary Medicine (DIMEVET) of the University of Milan, obtaining a consecutive series of cytological specimens. All cytological slides used in the current study had been air-dried and manually stained with May-Grünwald-Giemsa (Merck KGaA, Frankfurt, Germany).

Each cytological sample was defined as a single and independent *"case"* on the basis of the sampled LN and of the sampling technique (i.e. sampling of multiple LNs from the same dog and/or sampling of the same LN from the same dog with different techniques were not considered as exclusion criteria).

Retrieved cases previously diagnosed as inconclusive (i.e. acellular, hemodiluted, incorrectly smeared samples, or samples with suboptimal staining or containing only tissues other than lymphoid tissue such as fat, connective tissue, salivary gland, etc.) were excluded from the study. According to the cytological diagnosis reported in the electronic database as well as to the information contained in the medical record of each dog (including eventual histological reports dated not later than one month after the cytological report), cases were classified in one of the following clinical categories:

- Category 1 (Cat.1): cases from non-oncological dogs (NODs), i.e. dogs with reactive, degenerative, infectious-inflammatory, or traumatic conditions, in the absence of MCT or any other neoplasm;

- Category 2 (Cat.2): cases from MCT-bearing dogs (MCTDs), in the absence of any other neoplasm;
- Category 3 (Cat.3): cases from dogs bearing single or multiple neoplasms other than MCT.

Cases were excluded from the study when obtained from dogs lacking detailed clinical history, or simultaneously bearing a MCT and another kind of neoplasm. Cases belonging to NODs (Cat.1) and sampled with techniques other than fine needle aspiration (FNA), as well as all the cases belonging to dogs bearing single or multiple neoplasms other than MCT (Cat.3), were excluded the current work. Cases belonging to MCTDs (Cat.2) were then further classified in the following subcategories according to the sampling technique:

- Subcategory 2.1 (Subcat.2.1): cases from MCTDs sampled with FNA;
- Subcategory 2.2 (Subcat.2.2): cases from MCTDs sampled with scraping smearing;
- Subcategory 2.3 (Subcat.2.3): cases from MCTDs sampled with touch imprinting.

Sub-selection of eligible cases

Each previously selected case was then considered as eligible in the study only if at least 1 cytological slide was available for review, and if at least 20 adequately cellular high-power fields (i.e. 400x) were available. Each microscopic high-power field (HPF evaluated with routine light microscopy (Olympus BX51TF; Olympus Corporation, Tokyo, Japan), was considered as “adequate” when at least 50% of its surface was occupied by well-preserved cells arranged in a single layer without overlapping. When 1 slide was not sufficient to reach a total of 20 adequately cellular HPFs, a maximum of 3 additional slides were evaluated. Among eligible cases, only the 10 most recent ones for each Cat.1 (NODs), Subcat.2.1 (cases from MCTDs sampled with FNA), Subcat.2.2 (cases from MCTDs sampled with scraping smearing), and Subcat.2.3 (cases from MCTDs sampled with touch imprinting) were selected for the following investigations, for a final caseload of 40 cases. For each case included in

the study, details regarding each sampled LN (including location, size, and eventually-available histopathological diagnosis expressed according to Weishaar et al.[234] (Appendix 1 – Supplementary Figure 2) and signalment data of dogs from which they were obtained (i.e. breed, sex, age, and primary pathological process affecting each animal) were collected (Appendix 2 – Supplementary Table 2). All inclusion criteria applied in the current study are summarized in Appendix 2 – Supplementary Figure 2.

For each of the 30 selected cases obtained from MCTDs (Cat.2), the metastatic status (i.e. the presence or the absence of MCT metastasis) was evaluated by one of the investigators (i.e. the Ph.D. candidate) applying the criteria proposed by Krick et al.[104] as further modified in previously published studies[73,145] and by our research group, as shown in Appendix 2 - SupplementaryTable 1. Examples of morphological features evalauted to classify the cytological cases are illustrated in Appendix 1- Supplementary Figure 1. Specifically, the cases included in the current study were classified as follows:

- Subcategory 2.NM (Subcat.2.NM) – "Non-metastatic": including cases classified as "normal" or "reactive lymphoid hyperplasia" according to criteria by Krick et al.;
- Subcategory 2.PM; (Subcat.2.PM) - "Possibly metastatic": including cases classified as "possible metastasis" according to criteria by Krick et al.;
- Subcategory 2.M (Subcat.2.M) – "Metastatic": including cases classified as "probable metastasis" or "certain metastasis" according to criteria by Krick et al.

Criteria applied to microphotograph acqusition and counting

For each investigated case, microphotographs of 20 hot-spot, adequately cellular HPFs were taken with a digital camera Olympus DP70 (Olympus Corporation, Tokyo, Japan) supported by the software Olympus DP-Soft (v.5.0 for Microsoft Windows; Olympus Corporation, Tokyo, Japan). "Hot-spot" HPFs were defined as those characterized by the best balance between cellularity and cell preservation. Each

microphotograph field corresponded to an area of 434.97x327.51 microns, as measured with the aid of a calibration slide.

Each microphotograph was then randomized and blinded for category, subcategory and signalment information. Microphotographs were then uploaded in the Fiji software [187] and manually counted after application of a square grid of 108.74 microns side to facilitate counting (Appendix 2 – Supplementary Figure 1 A). The number of MCs for each evaluated case was then determined by one of the investigators with the help of another investigator (i.e. the Ph.D. candidate) for those more challenging cases, applying 6 different counting methods:

- *Counting "per fields"*: number of MCs counted in 4, 8, and 20 HPFs;
- *Counting "over cells"*: number of MCs over 500, 1000, and 2000 cells counted at 400x.

Naked nuclei and cells in mitosis or only partially depicted in the pictures were not counted (Appendix 2 – Supplementary Figure 1 B). To avoid introduction of bias during counting "over cells", each microphotograph used for this purpose was entirely counted. When cellular counting overwhelmed the threshold of 500, 1000, and 2000 cells, values were recalibrated with a proportion (i.e. number of MCs : number of counted cells = X : 500/1000/2000).

Statistical analysis

For each counting method applied to each category and subcategory (from now on indicated as "groups"), both the median (with the corresponding minimum-maximum range) and the mean (with the corresponding standard deviation) absolute number of MCs were calculated to facilitate comparison of our results with those reported in previous studies. Similarly, both the median and the mean percentage number of MCs were calculated for counting methods over cells applied to each group.

Data distribution for each counting method applied to each group mentioned above was investigated with Shapiro-Wilk normality test. Considered that the majority of datasets did not pass the normality test, non-parametric tests (i.e. Mann-Withney test

for pairwise comparison and Kruskall-Wallis test with Dunn's correction for multiple comparisons) were applied to investigate the difference in the median number of MCs for each counting method among the following groups:

- NODs (Cat.1) versus MCTDs (Cat.2);
- Non metastatic cases (Subcat.2.NM) versus possibly metastatic cases (Subcat.2.PM) versus metastatic cases (Subcat.2.M);
- FNAs (Subcat.2.1) versus scraping smears (Subcat.2.2) versus touch imprints (Subcat.2.3).

Specifically, for those comparisons among 3 groups, pairwise comparisons with Mann-Withney test were investigated in addition to Kruskall-Wallis test with Dunn's correction to improve the sensitivity of the analysis.

Beside MCs quantification, the difference in the mean cellularity among LN specimens grouped together according to the categories and subcategories mentioned above was also investigated. Mean cellularity of each case was defined as the mean cellularity observed in the first 5 microphotopgraphs entirely counted during counting "over cells", which represented the minimum common number of microphotographs entirely counted among all cases (Appendix 2 – Supplementary Table 2). Although all the datasets reporting mean cellularity data passed Shapiro-Wilk normality test, non-parametric tests (i.e. Mann-Withney test for pairwise comparison and Kruskall-Wallis test with Dunn's correction for multiple comparisons) were applied to maintain consistency with the other statistical analyses. Finally, to verify whether cellularity influenced the amount of nodal MCs, Spearman correlation coefficient was calculated to determine the eventual correlation between the number of MCs observed according to each counting method applied to each group on one hand, and the corresponding mean cellularity on the other.

Statistical analyses were performed with GraphPad Prism software (version 9.0.0 for Windows 64-bit, GraphPad Software, San Diego, California USA, www.graphpad.com). Statistical significance was set at p-value ≤ 0.05.

RESULTS

Selection workflow and signalment data for the cases included in the study

Among the 2665 cytological electronical records collected between January the 1st, 2016 and December the 31st, 2018, 1901 referred to cytological samples obtained from dogs. Out of a total of 492 canine lymph node cytological samples, 443 were reported for "regional lymphadenopathy" and 49 for "systemic lymphadenopathy".

Ninety (90) cases were then excluded from the study because classified as "inconclusive", 36 cases because obtained from subjects with not available or insufficient clinical history, and 7 cases because obtained from animals with a history of MCT and another type of neoplasia. Additionally, the 155 diagnostic samples obtained from dogs bearing single or multiple neoplasms other than MCT (Cat.3), were excluded en bloc because not consistent with the aims of the current study study.

Of the cases obtained from NODs (Cat.1), 28 were excluded due to sampling method other than FNA (2), inadequate cellularity for the requirements of the current study (15), and failure to retrieve slides (11). Of the cases obtained from MCTDs (Cat.2) via FNA (Subcat.2.1), 22 were excluded due to inadequate cellularity and 7 for failure to retrieve slides. Of the cases obtained from MCTDs (Cat.2) via scraping smearing (Subcat.2.2) and touch imprinting (Subcat.2.3), 23 and 27 were excluded due to inadequate cellularity, respectively.

In conclusion, 27 cases obtained from NODs (Cat.1), and 13, 25, and 32 cases obtained from MCTDs (Cat.2) via FNA (Subcat.2.1), scraping smearing (Subcat.2.2), and touch imprinting (Subcat.2.3), respectively, were defined as eligible in the study, for a total of 97 cases. The selection workflow applied in the current studies is summarized in Appendix 2 – Supplementary Figure 2.

Details regarding the 40 cases finally included in the current study as well as signalment data of the 25 dogs from which they were obtained, are summarized Appendix 2 – Supplementary Table 2. Among the 25 dogs included in the study, 17 patients provided 1 case each, 6 patients provided 2 cases each, 1 patient provided 3 cases, and 1 patient provided 8 cases. Eight (8) dogs were intact males, 2 intact

females, 4 castrated males, and 11 neutered females. The mean age was 6.98 years (range: 1 year - 13 years and 11 months). Patients belonged to 12 different breeds in addition to 8 mongrels.

The 8 NODs (Cat.1) were affected by the following pathological conditions: rhinitis with dermatitis and reactive histiocytosis (1 patient which provided two cases), necrotizing gingivitis (1), bilateral purulent otitis (1 patient which provided two cases), leishmaniasis (2), hepatic amyloidosis (1), allergic dermatopathy (1), stomatitis (1).

Among the 17 MCTDs (Cat.2), 15 patients had 1 single MCT, 1 patient had two MCTs, and 1 patient had three MCTs. Further data regardign the localization of each primary tumor are reported in Appendix 2 – Supplementary Table 2.

For all cases included in the study, data on the location of the sampled lymph node, as well as the size in cm (when available), are shown in Table 5.

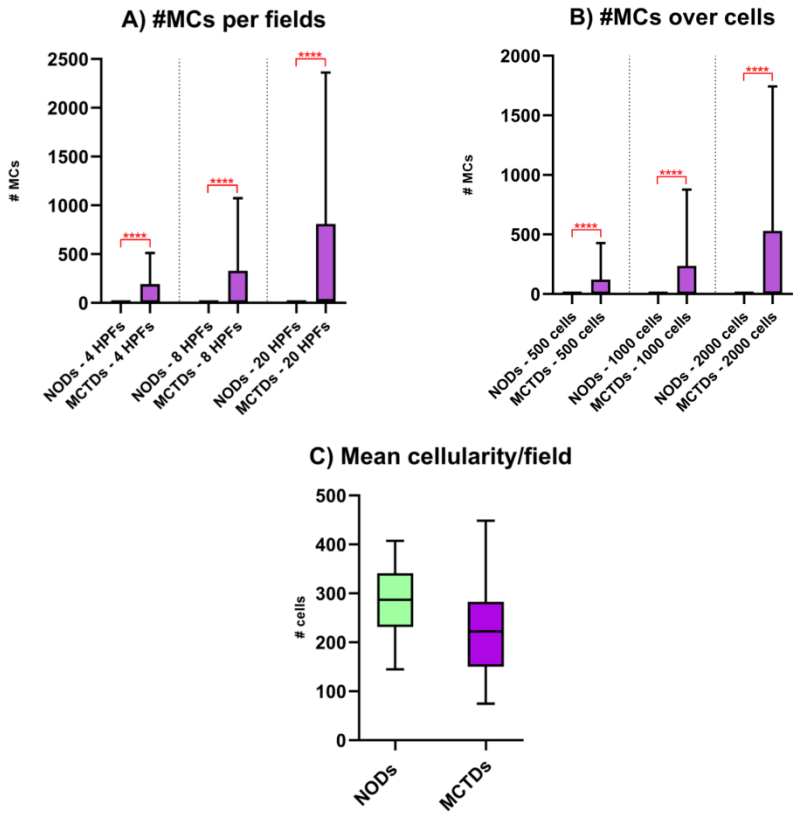
Cytological samples had been submitted by the University Veterinary Hospital of the DIMEVET or by external private veterinary clinics, or had been prepared at the Diagnostic Pathology Service of the DIMEVET from surgical biopsies sent for other investigations or from necroscopic samples. In general, the samples had been sent to assess regional or more rarely systemic lymphadenomegaly conditions, or for neoplasm staging. In this regards, it is notewhorty that most of the samples prepared via scraping smearing and/or touch imprinting (15/20) were obtained from clicnically non-palpable sentinel LNs surgically extirpated from MCTDs. Further details regarding each sampled LN (including location, size, and eventually-available histopathological diagnosis expressed according to Weishaar et al.(Weishaar 2014) are reported in the next paragraphs and/or in Appendix 2 – Supplementary Table 2. The latter includes also the number of slides available for review and the number of slides used to gain 20 adequately cellular HPFs for each examined cases, as well as the number of HPFs entirely-counted during the counting “over cells”.

Nodal mast cells in non-oncological dogs (NODs) and mast cell tumor-bearing dogs (MCTDs)

According to the different counting methods applied, the median range of nodal MC absolute number was 0-0.5 and 3.90-28 for NODs (Cat.1) and MCTDs (Cat.2), respectively.The mean range of nodal MC absolute number was 0-1.7 and 79.22-428.3 for NODs and MCTDs, respectively. The details regarding the median and mean absolute and percentage number of nodal MCs in NODs and MCTDs for each counting method, are reported in Appendix 2 – Supplementary Table 3.

When compared with the 10 NODs, the 30 MCTDs constantly showed a significantly higher number of MCs (p<0.0001), disregarding the counting method applied (Figure 1).

Figure 1. Nodal mast cells and sample cellularity in non-oncological dogs (NODs) and mast cell tumor-bearing dogs (MCTDs).



The box represents the interquartile range, with the internal horizontal line indicating the median value. The tip of the upper whisker indicates the maximum value while the tip of the lower whisker reports the minimum value. Statistically significant differences as determined with Mann-Withney test are indicated by red bars and stars. Legend: HPFs, high power fields; MCs, mast cells; #, number of; ****, p<0.0001.

Despite not statistically significant, mean cellularity was higher in cases obtained from NODs (median: 286.7) than from MCTDs (median: 222.2), as illustrated in Appendix 2 – Supplementary Table 6. Spearman correlation analysis revealed a significant inverse correlation between MC number and cellularity in samples obtained from MCTDs (r range: -0.4112 to -0.5281; p value range: 0.0027 to 0.0240), which was not observed for NODs (r range: 0.03906 to 0.05803; p value range: 0.9212 to >0.9999).

Nodal mast cells in specimens obtained from mast cell tumor-bearing dogs (MCTDs) and classified as "non metastatic" (NM), "possibly metastatic" (PM), and "metastatic" (M)

On the basis of the metastatic status as cytologically determined according to Appendix 2 – Supplementary Table 2, 10 cases out of the 30 obtained from MCTDs were classified as "non metastatic" (Subcat.2.NM), 5 cases as "possibly metastatic" (Subcat.2.PM), and 15 cases as "metastatic" (Subcat.2.M).

Despite beyond the aims of the current study, the histopathological evaluation of the nodal metastatic status was available for 23 out of 30 cases obtained from MCTDs for comparison with the corresponding cytological diagnosis (Appendix 2 – Supplementary Table 2). Specifically, cytological and histological diagnosis completely and partially agreed in 15/23 cases (65.22%) and in 4/23 cases (17.39%), respectively, corresponding to an overall agreement of 82.61%. Conversely, no agreement between cytological and histological diagnosis was observed in in 4/23 cases (17.39%).

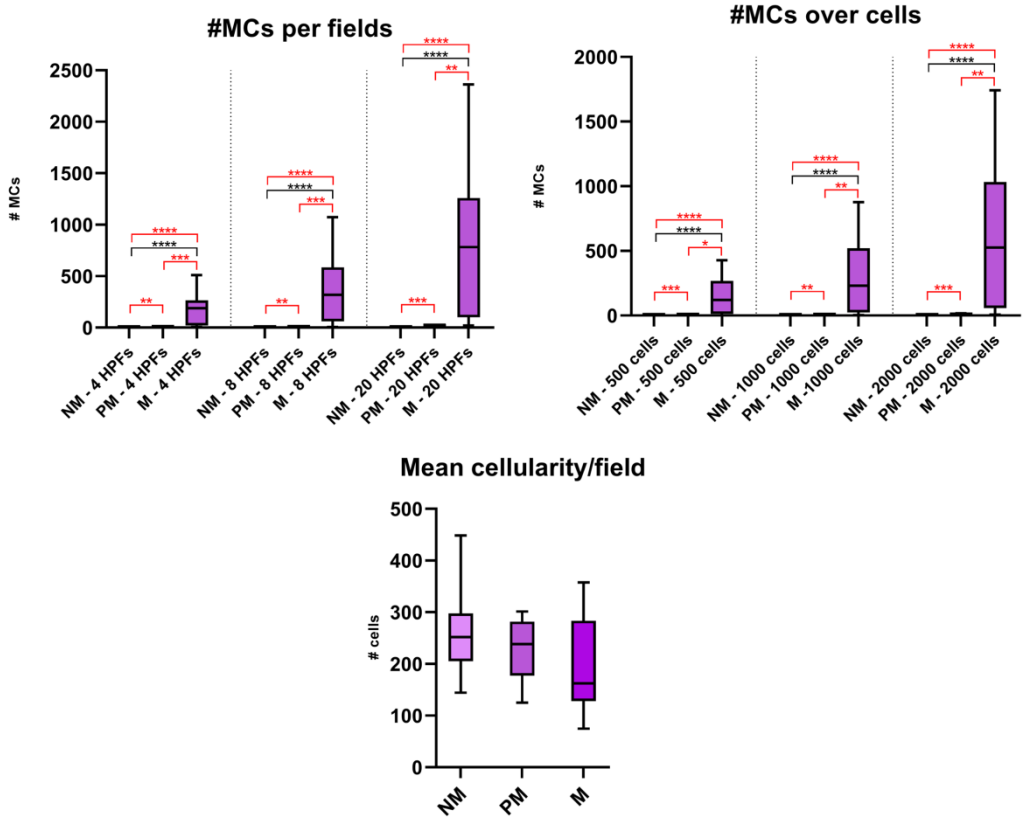
According to the different counting methods applied, the median range of nodal MC absolute number was 0-2 for "non metastatic" cases, 5-28 for "possibly metastatic" cases, and 120-782 for "metastatic" cases, respectively. The mean range of nodal MC absolute number was 0.25-2.40, 4.05-19.40, and 142.7-718.3 for "non metastatic", "possibly metastatic", and "metastatic" cases, respectively. The details regarding the median and mean absolute and percentage number of nodal MCs for

each counting method in “non metastatic”, “possibly metastatic”, and “metastatic cases are reported in Appendix 2 – Supplementary Table 4.

When differences in the amount of nodal MCs in MCTDs were investigated on the basis of the metastatic status with Kruskall-Wallis test with Dunn’s correction for multiple comparisons, cases classified as “metastatic” constantly showed a significantly higher number of MCs (p<0.0001) compared to “non metastatic” cases disregarding the counting method applied, but not compared to “possibly metastatic” cases (Figure 2 – black bars). Addtionally, no significant difference in the number of nodal MCs was observed between “non metastatic” and “possibly metastatic” cases. On the contrary, when Mann-Withney test was applied to pairwise comparisons, “non metastatic” cases constantly showed a signifcantly lower number of MCs compared to both “possibly metastatic” (p ranging from <0.01to <0.001) and “metastatic” cases (p<0.0001), disregarding the counting method applied (Figure 2 – red bars). Additionally, a significantly higher number of MCs (p ranging from <0.05 to <0.001) was observed in “metastatic” cases compared to “possibly metastatic” ones.

Despite beyond the aims of the current study, the statistical significance of the difference in the number of MCs between samples obtained from NODs and each of the subcategories of the samples obtained from MCTDs based on the metastatic status, was also investigated to allow complete comparability of our results with those reported in the literature. Results of this analysis are illustrated in Appendix 2 – Supplementary Figure 3.

Figure 2. Nodal mast cells and sample celularity referring to specimens obtained from mast cell tumor-bearing dogs (MCTDs) and classified as "non metastatic" (NM), "possibly metastatic" (PM), and "metastatic" (M).



The box represents the interquartile range, with the internal horizontal line indicating the median value. The tip of the upper whisker indicates the maximum value while the tip of the lower whisker reports the minimum value. Statistically significant differences as determined with Kruskall-Wallis test with Dunn's correction for multiple comparisons, are indicated by black bars and stars. Statistically significant differences as determined with Mann-Withney test are indicated by red bars and stars. Legend: HPFs, high power fields; MCs, mast cells; #, number of; *, p<0,05; **, p<0,01; ***, p<0,001; ****, p<0.0001.

Despite not statistically significant neither with Kruskall-Wallis test with Dunn's correction for multiple comparisons nor with Mann-Withney test for pairwise comparisons, mean cellularity showed a decreasing trend, being slightly higher in "non metastatic" cases (median: 251.8) compared to "possibly metastatic" ones (median: 238.6), which in turn were characterized by increased cellularity compared to "metastatic" cases (median: 162.2; Appendix 2 – Supplementary Table 6). Disregarding the counting method, Spearman analysis revealed a significant inverse

correlation between MC number and cellularity in cases classified as “metastatic” (r range: -0.5536 to -0.7500; p value range: 0.0014 to 0.0349), which instead was not observed in “non metastatic” (r range: -0.3206 to 0.6272; p value range: 0.0591 to >0.9999) and “possibly metastatic” cases (r range: -0.3000 to 0.4617; p value range: 0.433 to 0.9500).

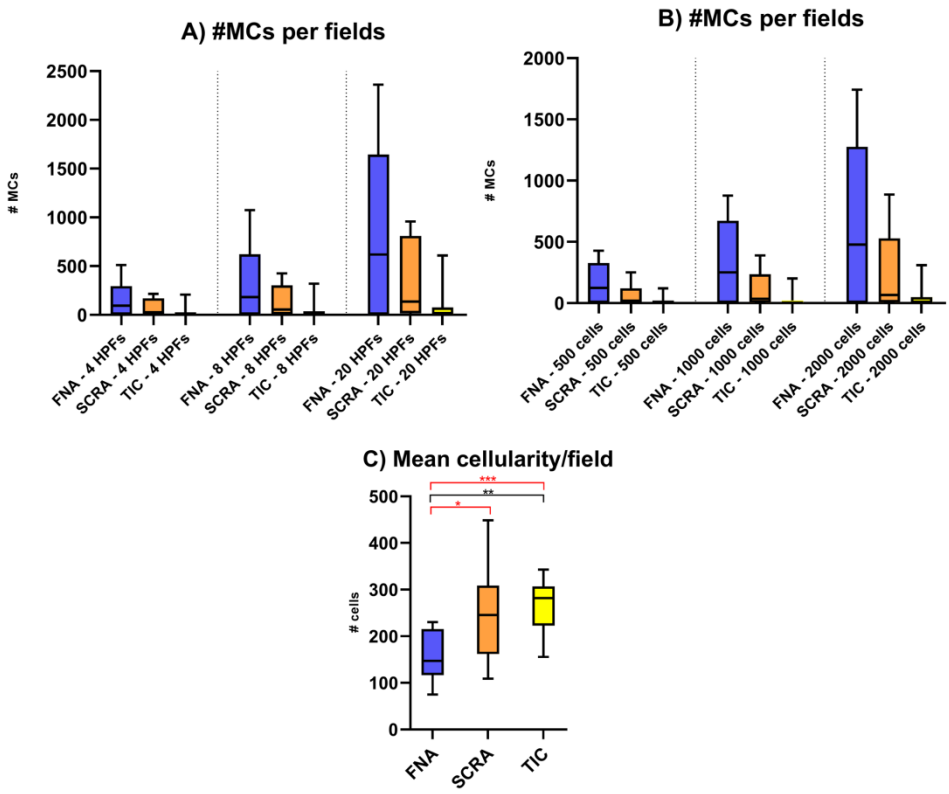
Nodal mast cells in specimens obtained from mast cell tumor-bearing dogs (MCTDs) and sampled via fine-needle aspiration (FNA), scraping smearing (SCRA), and touch imprinting (TIC)

As illustrated in Appendix 2 – Supplementary Table 2, among the 10 cases obtained from MCTDs via FNA (Subcat.2.1), 4 were classified as “non metastatic”, 1 as “possibly metastatic”, and 5 as “metastatic”. Among the 10 cases sampled with scraping smearing (Subcat.2.2), 2 were classified as “non metastatic”, 2 as “possibly metastatic”, and 6 as “metastatic”. Finally, among the 10 cases obtained via touch imprinting (Subcat.2.3), 4 were classified as “non metastatic”, 2 as “possibly metastatic”, and 4 as “metastatic”.

According to the different counting methods applied, the median range of nodal MC absolute number was 94.5-620 for FNAs, 19.10-136.5 for scraping smears, and 2.25-17.50 for touch imprints. The mean range of nodal MC absolute number was 156.3-833.6 for FNAs, 57.83-365.8 for scraping smears, and 14.66-85.60 for touch imprints. The details regarding the median and mean absolute and percentage number of nodal MCs for each counting method in FNAs, scraping smears, and touch imprints are reported in Appendix 2 – Supplementary Table 5.

When differences in the amount of nodal MCs in MCTDs were investigated on the basis of the sampling technique, neither Kruskall-Wallis test with Dunn’s correction for multiple comparisons nor Mann-Withney test for pairwise comparison revealed any statistically significant difference among the groups, disregarding the counting method applied. Nonetheless, a descending trend indicative of higher numbers of MCs in FNAs followed in order by cases obtained by scraping smearing and touch imprinting, was observed.

Figure 3. Nodal mast cells and sample celularity referring to specimens obtained from mast cell tumor-bearing dogs (MCTDs) *and sampled via fine-needle aspiration (FNA), scraping smearing (SCRA), and touch imprinting (TIC)*.



The box represents the interquartile range, with the internal horizontal line indicating the median value. The tip of the upper whisker indicates the maximum value while the tip of the lower whisker reports the minimum value. Statistically significant differences as determined with Kruskall-Wallis test with Dunn's correction for multiple comparisons, are indicated by black bars and stars. Statistically significant differences as determined with Mann-Withney test are indicated by red bars and stars. Legend: HPFs, high power fields; MCs, mast cells; #, number of; *, p<0,05; **, p<0,01; ***, p<0,001; ****, p<0.0001.

Mean cellularity showed an increasing trend, being lower in FNAs (median: 147.2) compared to scraping smears (median: 245.8), which in turn were characterized by slightly lower cellularity compared to touch imprints (median: 281.7; Appendix 2 – Supplementary Table 6). Kruskall-Wallis test with Dunn's correction for multiple comparisons revealed a significant difference in the mean cellularity between FNAs and touch imprints ($p<0.01$), which was also observed in pairwise comparison performed with Mann-Whitney test ($p<0.001$). The application of the latter test also revealed a significant difference in mean cellularity between FNAs and scraping smears ($p<0.05$).

Spearman correlation analysis revealed a significant inverse correlation between MC number and cellularity in scraping smears according to all counting methods (r range: -0.6525 to -0.8182; p value range: 0.0047 to 0.0463) except for counting of MCs in 8 HPFs (r: -0.6364; p value: 0.0544), and between MC number and cellularity in FNAs exclusively for counting over 500 cells (r: -0.6930; p value: 0.0314), although Spearman r coefficient was constantly negative also for the other counting methods (r range: -0.4085 to -0.6930; p value range: 0.0679 to 0.2153). On the contrary, Spearman correlation analysis did not reveal any significant correlation between MC number and cellularity in touch imprints, disregarding the counting method applied for MC quantification (r range: -0.006079 to 0.1376; p value range: 0.7026 to 0.9924).

DISCUSSION

The current study aimed to verify whether the quantification of nodal MCs is useful in the cytological determination of the presence or absence of MCT nodal metastasis in dogs. Specifically, the aim was pursued investigating the differences in the number of nodal MCs between non-oncoligcal dogs (NODs – Cat.1) and MCT-bearing dogs (MCTDs – Cat.2), as well as between cytological LN specimens obtained from MCTDs and classified as “non metastatic” (Subcat.2.NM), “possibly metastatic” (Subcat.2.PM), and “metastatic” (Subcat.2.M). Furthermore, investigations aimed to verify whether the sampling technique (i.e. FNA, scraping smearing, or touch imprinting) influences the number of nodal MCs in cytological specimens obtained from MCTDs.

Cytological LN specimens obtained from MCTDs (Cat.2) were characterized by a significantly higher number of MCs compared to NODs (Cat.1). This finding was in line with previous literature reporting that increased amounts of nodal MCs are suggestive of the presence of metastasis.[6,104,108,117,145,159,201,204,209,230]

On the other hand, only pairwise comparisons performed with Mann-Whitney test revealed a signifcantly increasing trend in “non metastic”, “possibly metastatic”, and “metastatic” samples obtained from MCTDs. Our findings were in line with previous studies evaluating the amount of nodal MCs among different groups of LN cytological specimens obtained from MCTDs, and reporting an increasing trend in the number of

nodal MCs shifting from non-metastatic cases to metastatic ones.[117,204] On the other hand, the lack of significant differences between “non metastic” and “possibly metastatic” cases observed with the application of Kruskall Wallis test with Dunn’s correction for multiple comparisons might be a consequence of the low number of cases classified as “possbily metastatic” (5) as well as of the cytological criteria applied [104,145] to classify the samples. Indeed, the criteria to be met for inclusion in the class "possibly metastatic" are based on the presence of occasional (i.e. ≤3) pairs or triplets of MCs on an entire slide, rather than on quantification of this cell type. In this regards, the uneven distribution of aggregated MCs among microphotographs investigated in the current study might have influenced our results. Despite not completely correct from a methodological point of view, the simultaneous application of statistical analyses for both mutlple and pairwise comparisons in the current study was justified by the preliminary nature of the current study, in view of further investigations aiming to verify the diagnostic accuracy of the inclusion of MC quantitative estimates among cytological criteria for the evaluation of nodal specimens obtained from MCTDs (as described in Section 2 of the current Thesis). In this regards, it is noteworthy that in a brief communication Fournier et al.[72] reported that LN samples classified in a new cateogry named "high-risk metastasis”. which was characterized by increased MC nuclear-cytoplasmic ratio and the presence of >8% MCs among all nucleated cells, showed a better correlation with one-year metastasis-free survival rate compared to that observed when samples were classified according to the system proposed by Krick et al.[104]

As highlighted in the introduction, former studies which focused on the cytolgoical quantification of nodal MCs in dogs without health issue, affected by non-neoplastic conditions such as allergies, or bearing MCT [12,25,117,145,174,204] were characterized by different inclusion criteria, counting methods, and statistical analysis, limiting the comparison of their findings with our results.

Specifically, the studies by Bookbinder et al.[25] and by Bauer et al.[12] reported that a certain amount of MCs can be present in LN obtained from healthy dogs, ranging from 1 to 16 and from 0 to 13, respectively. The comparison of our results with those reported by Bookbinder et al. is not possible because healthy dogs were not included

in the current study. On the other side, Bauer et al. applied different counting methods of nodal MCs observing that the number of MC in LNs from dogs with allergies was significantly higher than that in LNs from healthy dogs.[12] Considered that in the current study we also applied different counting methods of nodal MCs obtained from NODs, including one patient with allergic dermatopathy, the comparison of our results with those reported by Bauer et al.[12] is illustrated in Table 1.

Table 1. Mast cell (MC) cytological estimates in LNs obtained from allergic dogs and investigated by Bauer et al.,[12] and in LNs obtained from non-oncological dogs (NODs) and investigated in the current study.

#dogs and clinical category evaluated by Bauer et al. (2011)	#cases and clinical category evaluated in the current study	Counting method	Bauer et al. (2011)[12]	Current work
20 allergic dogs	10 NODs (Cat.1)	Range of MCs in 20 fields (mean)	0-3 (0.5)	0-11 (1.7)
		Range of MCs over 500 cells (mean)	0-1 (0.2)	0 (0)

Legend: n/a, not applied; #, number of.

Although observed results are similar between the two compared studies, it is noteworthy that the comparison is not entirely appropriate because of the retrospective nature of our study, which forced us to include among NODs patients affected by different pathological non-neoplastic conditions due to the low amounts of LN specimens obtained from allergic dogs retrieved from our archive. With this in mind, it should be noted that the heterogenity in pathological conditions affecting NODs included in the current study (Appendix 2 – Supplementary Table 2) might have determined different morphological patterns of lymph node reactivity secondary to different immune response patterns, which in turn are not necessarily associated with a marked increase in the number of nodal MCs, thus potentially biasing our results.

To the best of our knowledge, only 4 studies included the quantification of MCs also in cytological LN samples obtained from MCT-bearing dogs.[117,145,174,204]

Mutz et al.[145] compared MC estimates over 2000 cells in LNs obtained from healthy dogs (Group 1), dogs with allergic dermatopathy (Group 2), and MCT-bearing dogs

(Group 3). The comparison of our results regarding nodal MCs in NODs and MCTDs with those reported by Mutz et al.[145] for allergic and MCT-bearing dogs, respectively, is illustrated in Table 2.

Table 2. Mast cell (MC) cytological estimates in LNs investigated by Mutz et al.,[145] and in LNs investigated in the current study.

#dogs and clinical categories evaluated by Mutz et al. (2011)	#cases and clinical categories evaluated in the current study	Counting method	Mutz et al. (2011)[145]	Current work
20 allergic dogs (Group 2)	10 NODs (Cat.1)	Median% (min-max range%) MCs over 2000 cells	0.05% (0-0,55%)	0% (0-0,1%)
		Median (min-max range) MCs over 2000 cells	1 (0-110)	0 (0-2)
20 MCTDs (Group 3)	30 MCTDs (Cat.2)	Median% (min-max range%) MCs over 2000 cells	0.4% (0-77.4%)	0.64% (0-87.14%)
		Median (min-max range) MCs over 2000 cells	8 (0-1548)	12,8 (0-1743)

Legend: MCTDs, MCT-bearing dogs; Median%, median of the number of mast cells expressed as the percentage over 2000 cells; min-max range, minimum-maximum range referred to the Median; min-max range%, minimum-mximum range referred to the Median%; NODs, non-oncological dogs; n/a, not applied; #, number of.

While results regarding MCTDs almost overlapped between the two compared studies, MC estimates reported by Mutz et al.[145] for allergic dogs were higher than those reported for NODs in the current study. As already mentioned above about the study of Bauer et al.,[12] this difference might be the consequence of the bias introduced by the heterogenity of pathological conditions affecting NODs included in our study. Additionally, the logistic regression applied by Mutz et al.[145] revealed that MCT-bearing dogs were more likely to present at least one nodal MC or >0.05% nodal MCs compared to healthy dogs, but not when compared with animals affected by allergic dermatopathy. Conversely, in the current study MCTDs showed a significantly higher number of nodal MCs compared to NODs. Again, our results might be affected by the inclusion criteria applied to NODs iincluded in the current study.

Marconato et al.[117] evalauted LN cytological specimens obtained from healthy dogs (Group 1), dogs with infectious or inflammatory diseases (Group 2), and MCT-

bearing dogs (Group 3). The latter were subdivided into 3 subgroups: MCT-bearing dogs without LN metastases (Subgroup 3.1), MCT-bearing dogs with LN "inconclusive" for metastasis (i.e. LN cytological samples with only occasional mast cells; Subgroup 3.2), and MCT-bearing dogs with LN metastasis (i.e. LN cytological samples with replacement of normal lymphoid cells by MCs; Subgroup 3.3). Considered that in the current study LN samples obtained from MCTDs were similarly classified as "non metastatic", "possibly metastatic", and "metastatic" according to Appendix 2 – Supplementary Table 1, the comparison of our results with those reported by Marconato et al.[117] is illustrated in Table 3.

Table 3. Mast cell (MC) cytological estimates in LNs investigated by Marconato et al.,[117] and in LNs investigated in the current study.

#dogs and clinical categories evaluated by Marconato et al. (2008)	#cases and clinical categories evaluated in the current study	Marconato et al. (2008)[117] Mean%* (min-max range*) MCs over 2000 cells	Current work Mean% (min-max range) MCs over 2000 cells
31 dogs with Inf/non-npl (Group 2)	10 NODs (Cat.1)	0,01% (0-0,1)	0,01% (0-2)
5 MCTDs without LN metastasis (Subgroup 3.1)	10 "non metastatic" nodal samples from MCTDs (Subcat.2.NM)	0,07% (0,1-0,3)	0,05% (0-2.91)
4 MCTDs with LN inconclusive for metastasis** (Subgroup 3.2)	5 "possibly metastatic" nodal samples from MCTDs (Subcat.2.PM)	2,4% (0.8-2,8)	0,47% (4.84-16.82)
18 MCTDs with LN metastasis (Subgroup 3.3)	15 "metastatic" nodal samples from MCTDs (Subcat.2.M)	47,1% (0,4-98,86)	30,60% (6.98-1743)

Legend: inf/non-npl, infectious/inflammatory non neoplastic conditions; LN, lymph node; MCTDs, MCT-bearing dogs; Mean%, mean of the number of mast cells expressed as the percentage over 2000 cells; min-max range, minimum-maximum range; NODs, non-oncological dogs; n/a, not applied; std dev, standard deviation of the mean; #, number of; *reported as the least square mean; **, defined as LN cytological samples with only occasional mast cells.

Nodal MC esitimates observed in the current study were in line with those reported by Marconato et al.[117] for NODs, MCTDs without LN metastasis, and MCTDs with LN metastasis. On the other hand, we observed lower numbers of nodal MCs in MCTDs

with “possibly metastatic” LN compared to MCTDs with LN inconclusive for metastasis evaluated by Marconato et al. This difference might be the consequence of the different cytological evaluation criteria applied for the determination of the LN metastatic status between the two different studies, highlighting that the comparisons of the results from the two studies is not completely appropriate. Indeed, while Marconato et al. determined the LN metastatic status mainly on the basis of MC amounts, in the current study we applied a cytological interpretative system which was mainly based on the evaluation of MC tendency to aggregation (Appendix 2 – Supplementary Table 1). Nonehteless, Marconato et al. also reported that MCT-bearing dogs with LN metastasis had a significantly higher number of nodal MCs compared to healthy dogs, dogs affected by non-neoplastic infectious-inflammatory conditions, and MCT-bearing dogs without LN metastasis and with LN inconclsuive for metastasis. No significant differences emerged between any of the other groups and subgroups. Similar results were observed also in the current study when Kruskall Wallis test with Dunn’s correction for multiple comparisons was applied, as illustrated in Figure 2 and Appendix 2 – Supplementary Figure 3. Anyway, it is noteworthy that the additional differences among evaluated groups observed in our study might be the consequence of the aplication of Mann-Whitney test for pairwise comparisons, which is characterized by an increased power compared to statistical tests for multiple comparisons (such as ANOVA with Tukey’s correction for multiple comparison used in the study by Marconato et al.).

In their study, Sulce et al.[204] investigated the amount of nodal MCs over 1000 cells in 12 LNs obtained from MCTDs and classified according to the cytological interpretative system proposed by Krick et al.[104] The comparison of our results with those reported by Sulce et al.[204] is illustrated in Table 4.

Table 4. Mast cell (MC) cytological estimates in LNs investigated by Sulce et al.,[204] and in LNs investigated in the current study.

#samples according to metastatic status evaluated by Sulce et al. (2018)	#cases according to metastatic status evaluated in the current study	Sulce et al. (2018)[204] Min-max range% of MCs over 1000 cells	Current work Min-max range% of MCs over 1000 cells
3 LNs with "reactive lymphoid hyperplasia"* from MCTDs	10 "non metastatic" nodal samples from MCTDs (Subcat.2.NM)	0%-0.2%	0%-0.26%
2 LNs with "possible metastasis"* from MCTDs	5 "possibly metastatic" nodal samples from MCTDs (Subcat.2.PM)	2%-6%	0.20%-0.75%
7 LNs with "probable metastasis"* or "certain metastasis"* from MCTDs	15 "metastatic" nodal samples from MCTDs (Subcat.2.M)	1.6%-88.8%	0.43%-87.71%

Legend: LN, lymph node; MCTDs, MCT-bearing dogs; Min-max range%, minimum-maximum percentage number of MCs over 1000 cells; #, number of; *, determined according to criteria by Krick et al.[104]

While MC estimates reported by Sulce et al.[204] for LN samples classified as "reactive lymphoid hyperplasia" and "probable" or "certain metastasis" overlapped almost perfectly with those reported in the current study, our results for "possibly metastatic" LNs were markedly lower than those observed by Sulce et al. This latter finding might be the consequence of the different number of cases evaluated in each study, or of the difference in sampling techniques applied (4/5 specimens included in the current studies were sampled via scraping smearing or touch imprinting while those from Sulce et al. were all FNAs). Alternatively, our lower number of MCs might be due to the fact that 4/5 LNs classified as "possibly metastatic" in the current study were clinically non-palpable/normal-sized SLN, suggesting that the metaststic process might have been in its first pahses. Nonetheless, no data reagrding the size of LNs investigated by Sulce et al. were avaialable to verify this hypothesis.

Finally, Sabattini et al.[174] applied semi-quantitivative estimation of MCs in 28 LN cytological samples obtained from MCT-bearing dogs and histologically classified as HN2 or HN3, finding variable proportions of MCs ranging from <10% to >50%.

Unfortunately, MC estimates were evaluated on the entire slides and only aggregated data were reported in the paper, thus making not possible the comparison with our results.

To the best of our knowledge, this is the first study quantifying MCs in cytological nodal specimens from MCTDs sampled with different techniques. The differences in the number of MCs among the different sampling techniques were not statistically significant neither with the application of Kruskall Wallis test with Dunn's correction for multiple comparisons nor with Mann-Whitney test for pairwise comparisons. This observation suggests that the sampling technique applied do not have influence on the cytological evaluation of LN metastatic status in MCT-bearing dogs, thus opening the road to future investigations on the intraoperative use of cytology for the evaluation of surgically extirpated SLNs. Nontheless, the fact that none of the lymph nodes from MCTDs had been sampled according to all three different methods may had markedly limited the reliability of the results related to the effects of the sampling technique. Additionally, it is noteworthy that the cytological diagnosis for the scraping smear and the touch imprint obtained from the same lymph node was not consistent in 3 cases (coupled case IDs #4-#5, #25-#26, and #12-#13). For all these cases, the histopathological diagnosis was available for comparison, revealing that in two cases the correct diagnosis was moderately to significantly underestimated in scraping smears (coupled case IDs #25-#26, and #12-#13) while in the other case a partially incorrect diagnosis was provided for the touch imprint (coupled case IDs #4-#5). This finding underlyes the need of further studies aiming to validate the diagnostic accuracy of cytology and to clarify which factors influence the cytological diagnosis of LNs obtained from MCT-bearing dogs.

The observed decreasing trend in the number of MCs from FNAs to touch imprints might be explained by potential correlations between the size of sampled LNs and their metastatic status. Specifically, at least 8/10 FNAs obtained from MCTDs were reported as megalic, of which 1 was classified as "possibly metastatic" and 3 as "metastatic". It is therefore likely that in these cases, lymphdenomegaly which led to clinical sampling with FNA was determined by the invasion of the LNs by metastasis, with a consequent marked increase in the number of MCs. On the contrary, most of

the cases sampled via scraping smearing or touch imprinting were obtained from clinically non-palpable/normal-sized SLNs (15/20), disregarding their metastatic status. This observation might be the consequence that the majority of the LNs sampled via scraping smearing or touch imprinting and classified as “metastatic” had been surgically extirpated at an earlier stage of metastatic invasion, prior to lymphadenomegaly development.

When the mean cellularity was investigated among the groups evaluated in the current study, the only significant differences were observed between FNAs and touch imprints and, with Mann-Whitney test, also between FNAs and scraping smears. The lower cellularity of scraping smears compared to touch imprints could have resulted from a greater number of cells broken as a result of the traumatic action on the neoplastic cells during smear preparation, although the literature reports that the scraping smears ensures an increased cellularity compared to FNA.[48,229] Additionally, when the correlation between estimates of nodal MCs and mean cellularity of each group was investigated, a significant inverse correlation was almost constantly observed for MCTDs in general, for cases obtained from MCTDs and classififed as “metastatic”, and for scraping smears, disregarding the counting method applied. These signifcant correlations observed might be correlated to the number of MCs in the lymph node and their bigger size compared to lymphocytes. Indeed, MCs are large cells which occupy more space than lymphocytes within each HPF, thus reducing the space for other cells and leading to a lower mean cellularity. This hypothesis was further supported by the observation that a significant Spearman correlation was revealed for almost all counting methods applied in scraping smears and for counting over 500 cells in FNAs, which were the groups including the highest numbers of cases classsified as metastatic (6 and 5, respectively). The analysis of MC estimates and mean cellularity after re-grouping of cases obtained from MCTDs according to both the sampling tehcnique and the metastatic status would have further clarified the reasons of these correlations. Unfortunately, this regrouping of cases would have reduced too much the number of cases for each group, not allowing reliable statistical analyses.

It is noteworthy that the statistically significant differences observed among the evaluated groups, were maintained disregarding the counting method applied for MC quantification. Counting methods “over cells” are still considered more reliable because are based on a proportionality ratio that, at least theoretically, is not affected by the extremely variable cellularity that can characterize cytological samples.[117,145] Nonetheless, our findings open the road to the application of counting methods “per field” in further investigations aiming to verify the diagnostic accuracy of the inclusion of MC quantitative estimates among cytological criteria for the evaluation of nodal specimens obtained from MCTDs (as described in Section 2 of the current Thesis). Indeed, counting “per fields” might markedly reduce the time needed for MC quantification compared to counting methods “over cells”. This hypothesis might be supported by the wide range of microphotographs used in the current study to gain the 2000 cell cut-off (5-19 microphotographs; Appendix 2 – Supplementary Table 2). Other strengths of the current study are the application of numerous different counting methods, which could improve the comparison of our results with those from potential future studies by other investigators, as well as the quantification of MCs in microphotographs rather than directly at the microscope, which could improve the standardization of future interrater agreement investigations.

The lack of systematical comparison of the cytological evaluation of the LN metastatic status with the corresponding histological one represents the major limitation of the current work, given that the latter still represents the reference diagnostic standard.[6,71,104,108,119,145,150] This limitation is supported by the observation that the agreement between cytological and histological diagnosis was not perfect for those cases evaluated with both techniques (82.61%), and by the fact that in 3 cases the cytological diagnosis for the scraping smear and the touch imprint obtained from the same lymph node was not consistent. These discrepances highlight the presence of some diagnostic bias which could have limited the reliability of our findings.

The retrospective nature of the study and the experimental design determined a number of further limitations. Firstly, the disparity in the number of cases assigned to each of the distinct subcategories according to the cytological diagnosis may have influenced the statistical significance of differences observed among different groups.

This limitation is related to the inclusion criteria which were applied in this form to avoid bias resulting from subjective selection of cases. Nonetheless, many subjects have been excluded due to inaccurate information regarding their clinical history, thus introducing a certain degree of selection bias. Another limitation that may have affected the results by introducing a fair degree of variability, is the different anatomical location of evalauted LNs. Nonetheless, to the best of our knowledge, the literature does not report whether there are differences in the amount of MC normally present in dog LNs according to the anatomical location.

In conclusion, the current preliminary study revealed that MC quantification in LN cytological samples obtained form MCTDs might be useful to determine the nodal metastatic status. This conclusion is particularly true in view of the inclusion of nodal MC quantitative estimation in the current system for cytological evaluation of LNs proposed by Krick et al.,[104] which is mainly based on morphological and aggregation criteria rather than on the quantification of MCs. However, this proposal needs to be verified in more detailed studies of diagnostic accuracy (such as the one illustrated in the Section 2 of the current Thesis), since in the present work histological diagnosis was available only for a limited subset of cases. Although limited by the lack of the application of the 3 different sampling techniques on the same LN, our findings suggest that a specific sampling technique does not influence the number of nodal MCs. Therefore, future studies aiming to test the diagnostic accuracy of cytology could include specimens obtained with a different sampling technique, without theoretically biasing the results. Finally, it is noteworthy that all the differences found in this study were always maintained between the different groups investigated regardless of the counting method applied. This observation opens the road to the use of counting methods “per fields” in future studies, which could significantly speed up cytological evaluation times compared to counting methods “over cells”.

SECTION 2 - Investigations on the interrater agreement and the diagnostic accuracy of cytology in the evaluation of lymph nodes obtained from mast cell tumor-bearing dogs

MATERIAL AND METHODS

Case selection from the archives

The current study has been conducted according to STARD guidelines.[27]

Canine cytological nodal samples collected between January the 1st, 2008 and July the 31st, 2019 were retrospectively recovered from the electronic database of the Diagnostic Pathology Service of the Department of Veterinary Medicine (DIMEVET) of the University of Milan, using the keywords “lymph node” and “dog”. A consecutive series of cytological specimens was thus obtained. According to the medical record of each selected dog, those cytological nodal specimens obtained from patients lacking a clinical history of MCT, not bearing a MCT at the time of submission, simultaneously bearing a MCT and another kind of neoplasm at the time of submission, or lacking a detailed clinical history, were excluded from the study (Appendix 3 – Supplementary Figure 1 – category A).

Each cytological sample was defined as a single and independent *“case”* on the basis of the sampled LN and of the sampling technique (i.e. sampling of multiple LNs from the same dog and/or sampling of the same LN from the same dog with different techniques were not considered as exclusion criteria).

The electronic database of the Diagnostic Pathology Service of the DIMEVET was then mined to recover histopathological samples matching to the same LNs cytologically sampled. Those cytological cases for which an histopathological LN specimen was not available in general (category B.1) and those for which histopathological specimens were obtained from LNs other than the ones cytologically sampled (category B.2), were excluded from the study. Additionally, all those cases for which >2 months elapsed between cytological and histopathological

evaluation (category B.3) as well as those whose perfectly matching histopathological specimens were obtained through incisonal biopsy or needle-core biopsy (category B.4), were also excluded from the study. Cases fulfilling the above mentioned inclusion criteria were considered as eligible in the study.

Those cytological cases for which at least one cytological and one histological slide were not available for review (category C.1 and C.2, respectively), were excluded from the study. All cytological slides used in the current study had been air-dried and manually stained with May-Grünwald-Giemsa (Merck KGaA, Frankfurt, Germany). All the samples for histopathology had been fixed in 10% buffered formalin for at least 24 hours, followed by routine processing and inclusion in paraffin. For each formalin-fixed paraffin-embedded (FFPE) sample, at least two 5-micron sections were cut and stained with hematoxylin-eosin and with Giemsa, respectively.

All cytological and histopathological slides were reviewed at low- and intermediate-power magnification (i.e. 10x and 20x, respectively) to assess specimen adequacy. Those cytological cases which corresponding histopathology revealed the presence of a neoplasm other than a MCT (e.g. lymphoma) or the absence of tissue consistent with lymph node, were excluded from the study (category D). Similarly, also those cytological cases assessed as inconclusive (i.e. acellular, hemodiluted, incorrectly smeared samples, or samples with suboptimal staining or containing only tissues other than lymphoid tissue such as fat, connective tissue, salivary gland, etc.), were excluded from the study (category E). At the end of the review process, only the best cytological slide in terms of cellularity and lack of overlapping of cells among the available ones for each case, was selected for further investigations. As an example, each touch imprint cytological slide should have included at least two imprints.

For each case included in the study, details regarding each sampled LN (including location, size, classification as regional or sentinel LN, and sampling technique applied for collection of cytological specimens – i.e. FNA, scraping smearing, or touch imprinting) and signalment data referring to dogs from which they were obtained (i.e. breed, sex, age, and number and location of primary MCT) were collected whenever available (Appendix 3 – Supplementary Table 1). All inclusion

criteria applied in the current study are summarized in Appendix 3 – Supplementary Figure 1.

Independent evaluation of cytological and histopathological specimens

Each cytological case was then randomized and blinded for signalment data and details regarding the sampled LN. Cases were then independently evaluated by five investigators: an undergraduate student routinely involved in the Cytopathological Diagnostic Service of the DIMEVET (Reader 1), a first year PhD student routinely involved in the Hematology Diagnostic Service of the DIMEVET (Reader 2), a second year PhD student and ECVP trainee routinely involved in the Cytopathological Diagnostic Service of the DIMEVET (Reader 3, the Ph.D. candidate), a second year PhD student and ECVP trainee occasionally involved in the Cytopathological Diagnostic Service of the DIMEVET (Reader 4), and an ECVP-diplomate who is the Diagnsotic Leader of the Cytopathological Diagnostic Service of the DIMEVET (Reader 5) . Each Reader was exclusively aware of the fact that all cytological specimens were obtained from dogs with a clinical history of MCT or bearing a MCT at the time of submission. Each Reader was provided with an Excel table (Microsoft Corporation, 2018) with specific features that had to be evaluated (Excel tables containing the results of the evaluation by each Reader available upon request), which included the evaluation/quantification of:

- Sample cellularity (over the entire slide): semiquantitative evaluation to be expressed as +, ++, or +++;
- effacement of lymphoid tissue (i.e. replacement of at least 50% normal lymphoid cells by MCs over the entire slide): to be expressed as yes or no (Appendix 1 – Supplementary Figure 1);
- number of aggregates of MCs composed by >3 MCs (over the entire slide): to be quantified in detail from 0 to 5, or as >5 (Appendix 1 – Supplementary Figure 1);
- number of couples of MCs (over the entire slide): to be quantified in detail from 0 to 3, or as >3 (Appendix 1 – Supplementary Figure 1);

- number of triplets of MCs (over the entire slide): to be quantified in detail from 0 to 3, or as >3 (Appendix 1 – Supplementary Figure 1);
- number of discrete, morphologically normal (i.e. without cellular atypies) MCs (over 8 “hot spot-choosen” HPFs): to be quantified in detail for each HPF from 0 to 8, or as >8;
- number of discrete, atypical MCs (over 8 “hot spot-choosen” HPFs): to be quantified in detail for each HPF from 0 to 8, or as >8;
- cellular atypies observed in atypical MCs (i.e. bi- and multinucleation, nuclear pleomorphism, nuclear-cytoplasmic ratio >1:1, nuclear-cytoplasmic ratio <1:1,anisokaryosis, anisocytosis, variable/reduced cytoplasmic granulation): to be expressed as yes or no for each type;
- eosinophils (over the entire slide): semiquantitative evaluation to be expressed as +, ++, or +++;
- background granularity (over the entire slide; exclusively considering metachromatic granules of MC origin): semiquantitative evaluation to be expressed as +, ++, or +++;
- presence/absence of reactive lymphadenopathy (i.e. increased number of medium- and large-size lymphocytes, plasma cells, etc, according to Raskin et al.[166]): to be expressed as yes or no, disregarding all the aformentioned parameters.

Specifically, “hot spot-choosen” HPFs were defined as 8 fields at 400x magnification, not necessarily consecutive, and characterized by the best balance among highest possible cellularity, lack of cell overlapping, best cell preservation, and highest possible amount of MCs. In view of saving time, each Reader was left free to use the same 8 HPFs for counting morphologically normal and atypical MCs. For touch imprints, each Reader was asked to evaluate all the imprints to express a judgement for those parameters to be evaluated “over the entire slide”, and was still left free to choose the 8 HPFs to quantify morphologically normal and atypical MCs. Examples of morphological features evaluated to classify the cytological cases are illustrated in Appendix 1.

Histopathological samples were collaboratively reviewed by 3 anatomical pathologists (2 ECVP-diplomates and 1 non ECVP-diplomated) routinely involved in the Histopathological Diagnostic Service of the DIMEVET, and classified according to the criteria proposed by Weishaar et al.,[234] as illustrated in Chapter 3 Introduction – Table 1 and in Appendix 1 – Supplementary Figure 2. Specifically, the 3 anatomical pathologists were exclusively aware of the fact that all specimens were obtained from dogs with a clinical history of MCT or bearing a MCT at the time of submission.

Expression of cytological diagnosis according to interpretative system proposed by Krick et al. [104] *and to an arbitrarily amended version of the same interpretative system ("Amendment 1")*

On the basis of data reported in the Excel table, each Reader was asked to establish the cytological diagnosis for each case according to the interpretative criteria previously reported by Krick et al. (Chapter 3 Introduction – Table 2),[104] and according to an arbitrarily amended version of the system by Krick et al., proposed by our research group and named "Amended 1" system (AM1 system), as illustrated in Table 1 below.

Table 1. Criteria of an arbitrarily amended system (AM1 system) for the cytological evaluation of nodal metastasis in MCT-bearing dogs proposed by our research group.

Amended 1 (AM1) system – Cytological criteria	Diagnosis
• 0-1 couple or triplet of MCs / entire slide • 0-3 discrete, atypical MCs (bi- or mutlinucleation, nuclear pleomorphism, altered N:C ratio, anisokaryosis, anisocytosis, variable/reduced cytoplasmic granulation) / 8 HPFs • 0-3 discrete, normal MCs/HPF/8 HPFs (alternatively, 0-24 normal MCs/8 HPFs)	*Non metastatic* (NM)
• 1 aggregate of >3 MCs / entire slide • 2-3 couples or triplets of MCs / entire slide • 4-7 discrete, atypical MCs / 8 HPFs • 4-7 discrete, normal MCs/HPF/8 HPFs (alternatively, 25-56 normal MCs/8 HPFs)	*Possibly metastatic* (PM)

Amended 1 (AM1) system – Cytological criteria	Diagnosis
• Effacement of lymphoid tissue by MCs (i.e. >50% MCs on the total cellular amount/entire slide) • ≥2 aggregates of >3 MCs / entire slide • ≥4 couples or triplets of MCs / entire slide • ≥8 discrete, atypical MCs / 8 HPFs • ≥ 8 discrete, normal MCs/HPF/8 HPFs (alternatively, ≥57 normal MCs/8 HPFs)	*Metastatic* (M)

To express a specific diagnosis, the fulfilling of at least one of the criterion was considered necessary and sufficient. Legend: HPF, high power field (i.e. x400); MCs, mast cells; N:C, nuclear cytoplasmic ratio.

The rationale applied to set criteria included in AM1 is reported in Appendix 3 – Supplementary Section 1.

After each of the Readers had established the cytological diagnosis according to all the cytological intepretative systems included in the current study, two evaluators (Reader 1 and Reader 3) jointly verified that the criteria reported in the different systems had been strictly applied. Any discrepancies between each Reader and the two evaluators in the choice of the class to be assigned to each cytological case, were discussed collectively until an agreement upon the final diagnosis to assign was reached.

Data regarding sample cellularity, cellular atypies observed in atypical MCs, eosinophils, and background granularity were not evaluated in details in the current work thesis as their evaluation was not included in any of the cytological interpretative systems applied in the current study. Conversely, according to the criteria reported in the system by Krick et al.,[104] attempts to take into account the presence or absence of reactive lymphadenopathy were performed. However, during revision of diagnoses expressed by each Reader, several discrepancies regarding the assignment of “normal” and “reactive lymphoid hyperplasia” diagnosis were observed, due to the fact that the system by Krick et al. do not contemplate the possibility of observing few scattered, morphologically normal MCs in a non-reactive nodal specimen. Therefore, to standardize as much as possible the diagnosis expressed by the Readers, a diagnosis of “normal” was assigned in all those cases not falling within “possible metastasis”, “probable metastasis”, or “certain metastasis” diagnoses, and characterized by complete absence of discrete, normal MCs. On the

other hand, a diagnosis of “reactive lymphoid hyperplasia” was assigned in all those cases not falling within “possible metastasis”, “probable metastasis”, or “certain metastasis” diagnoses, and characterized by the presence of discrete, normal MCs.

Interrater agreement evaluation

The interrater agreement among the 5 Readers when the diagnosis was expressed according to the 5 interpretative classes proposed by Krick et al., was investigated calculating free-marginal kappa coefficient with an online calculator (http://justusrandolph.net/kappa/).

According to the preliminary study performed by our research group on MC quantification in canine nodal cytological specimens illustrated in Section 1 of the current Thesis, free-marginal kappa coefficient was calculated also reducing the system proposed by Krick et al.[104] into a 3-classes simplified version, as described in Appendix 2 – Supplementary Table 1. Specifically, “normal” and “reactive lymphoid hyperplasia” classes proposed by Krick et al. were grouped together and renamed “non-metastatic” (NM), “possible metastasis” was renamed as “possibly metastatic” (PM), and “probable metastasis” and “certain metastasis” classes were grouped together and renamed “metastatic” (M). No changes to the cytological interpretative criteria proposed by Krick et al.[104] were applied.

Finally, free-marginal kappa coefficient was calculated to evaluate the interrater agreement after establishment of the cytological diagnosis applying criteria proposed by our research group in the AM1 system.

To investigate the interrater agreement in further details, weighted kappa coefficient was also calculated for pairwise comparison of each couple of Readers using GraphPad QuickCalcs website (https://www.graphpad.com/quickcalcs/kappa1/).

According to previous literature,[126] results of kappa coefficient calculation were interpreted as follows: 0-.20, no agreement; .21-.39, minimal agreement; .40-.59, weak agreement; .60-.79, moderate agreement; .80-.90, strong agreement; >.90, almost perfect agreement.

Statistical analysis for diagnostic accuracy evaluation

For each case, both cytological diagnoses expressed according to the criteria proposed by Krick et al.[104] and to the AM1 system by each Reader, were tabulated with the corresponding histopathological diagnosis for comparison. Specifically, for the cytological diagnosis expressed according to the criteria by Krick et al.,[104] the 3-classes simplified version illustrated in Appendix 2 – Supplementary Table 1 was used to allow and facilitate the comparison with the histopathological diagnosis. Similarly, to allow and facilitate the comparison of the cytological diagnoses with the histopathological ones, HN2 and HN3 classes were grouped together in the class HN2/3, according to previous literature widely recognizing the corresponding histological findings as indicative of nodal metastasis in MCT-bearing dogs.[65,119,156,174,204,234]

When not explicitly stated, the statistical analyses listed and described in the following paragraphs have been performed for the cytological diagnosis expressed according to both the 3-classes simplified version of the system by Krick et al.,[104] and the AM1 system.

Overall percentage agreement between cytology and histopathology was then evaluated for each Reader. Cases were classified as “in complete agreement” (CA) with histopathology when a NM, PM, or M cytological diagnosis corresponded to a HN0, HN1, or HN2/3 histological diagnosis, respectively. All the other combinations of cytological and histological diagnosis were classififed as “not in agreement” (NA). This choice was based on the observation of some kind of overlapping in morphological features between cytological cases with “possible metastasis” and histological cases classified as HN1, both characterized by a slight increase in nodal MCs which is not sufficient *per se* to definitely diagnose a LN specimen as metastatic. Percentage agreement with histopathology was also calculated grouping the cases on the basis of the corresponding histological diagnosis or of the sampling technique through which they were obtained. Finally, median percentage agreement among the Readers was calculated for all the aforementioned comparisons between cytological and histological diagnosis. Furthermore, for each Reader and for the median among the 5 Readers, Chi-squared test for the comparison of proportions

was applied using Medcalc online calculator (https://www.medcalc.org/calc/comparison_of_proportions.php) to investigate the presence of statistically significant differences in the percentage agreement grouping cytological cases on the basis of the corresponding histological diagnosis or of the sampling technique. Statistical significance was set at $p<.05$.

For those cytological cases sampled from the same LN applying different techniques (e.g. touch imprinting and scraping smearing), the concordance between the cytological diagnoses of the different samples was further investigated on the basis of each applied cytological interpretative system. Specifically, complete intersampling coherence was defined when the diagnoses for all the 2 or 3 sampling techniques were in agreement among each other. Complete intersampling coherence was further distinguished on the basis of the correctness of the diagnoses compared to the histological reference standard. Partial intersampling coherence was defined exclusively for those cases sampled with all the 3 different techniques, and was further distinguished on the basis of the number of techniques which were in agreement with histology (i.e. 2 or 1 out of 3). Finally, intersampling inchoerence could have been observed exclusively for those cases sampled with 2 different techniques, among which one diagnosis was right and the other was wrong.

The agreement between cytology and histopathology was further investigated as previously proposed [207,233] tabulating data in a 3x3 contingency table and calculating weighted kappa coefficient using GraphPad QuickCalcs website (https://www.graphpad.com/quickcalcs/kappa1/). After calculating the median of cases for each cell of the 3x3 contingency tables, the median weighted kappa coefficient among the 5 readers was also calculated. Resulting kappa coefficients were interpreted as described above.

To allow calculation of diagnostic accuracy indexes (i.e. overall accuracy, sensitivity, specificity, positive predictive value, and negative predictive value), HN1 histopathological diagnoses were considered as not actually indicative of metastasis according to previous literature,[65,73,84,120,156,204,234] and were therefore considered as “negative for metastasis” being grouped with HN0 diagnoses in the class HN0/1. Conversely, no clear consensus exists in the literature whether those cases classified

as “possible metastasis” according by Krick et al., should be considered as actually indicative of metastasis or not.[73,104] To maintain consistency with the aforementioned choice of considering cytological cases classified as PM in perfect agreement with the paired FFPE specimen when histological diagnosis was HN1, in the current work cases classified as “possibly metastatic” were considered as negative for metastasis (“best case scenario”). Nonetheless, given the issue of the correct biological interpretation of those cytological samples diagnosed with “possible metastasis”, diagnostic accuracy indexes were also calculated alternatively considering those cytological specimens as positive for metastasis in view of further studies regarding this topic (“worst case scenario”). These data are provided exclusively as supplementary material in Appendix 3 – Supplementary Table 9, and will not be taken into account neither in the results nor in the discussion of this Section.

For both cytological interpretative system (the one proposed by Krick et al. and the AM1 system), each cytological diagnosis of each Reader was tabulated with the corresponding histological one. Then, considering histological diagnosis as the reference standard, cytological cases were classified as True Positive (TP) when correctly diagnosed as positive for metastasis, True Negative (TN) when correctly diagnosed as negative for metastasis, False Positive (FP) when diagnosed as positive for metastasis with a corresponding histological diagnosis indicative of lack of metastasis, and False Negative (FN) when diagnosed as negative for metastasis with a corresponding histological diagnosis indicative of actual metastasis. Results of comparison were then included in a 2x2 contingency table and used to perform McNemar’s test to verify whether a test bias (i.e. a significant difference in the proportion of positive responses to each test) was present.[233] McNemar’s test was performed through an online calculator (https://www2.ccrb.cuhk.edu.hk/stat/confidence%20interval/McNemar%20Test.htm), setting statistical significance at $p<.05$.

Diagnostic accuracy indexes with their relative 95% confidence interval (CI) were calculated using Medcalc online software (https://www.medcalc.org/calc/diagnostic_test.php). Briefly, accuracy was defined as the ability of cytology to correctly diagnose a case and was calculated as

(TP+TN)/(TP+FN+FP+TN). Sensitivity was defined as the probability of cytology to correctly detect the presence of metastasis and was calculated as TP/(TP+FN). Specificity was defined as the probability of cytology to correctly detect the absence of metastasis and was calculated as TN/(FP+TN). Positive predictive value was defined as the probability of actually having a metastasis with a cytological positive diagnosis and was calculated as TP/(TP+FP). Negative predictive value was defined as the probability of actually being free from metastasis with a cytological negative diagnosis and was calculated as TN/(FN+TN).

Finally, after calculation of the median number of TP, TN, FP, and FN, McNemar's test and diagnostic accuracy index calculation were repeated to determine the median value of these statistical parameters among the five readers.

According to previous literature,[24,207] each percentage value regarding both percentage agreement and diagnostic accuracy indexes was evalauted as follows: <70%, low; ≥70% - <80%, moderate; ≥80% - <90%, high; ≥90%, very high.

Setting of additional amended cytological interpretative systems with ROC curves and evaluation of the corresponding diagnostic accuracy indexes.

For each of the quantifiable parameters included in AM1 system (i.e. aggregates composed by >3 MCs, couples and triplets of MCs, discrete atypical mast cells, and discrete normal MCs), the median value for each case among the 5 Readers was calculated, and a receiver operating characteristic (ROC) curve was then designed. Specifically, 2 different sets of ROC curves were built with the SPSS software (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp).

The first set of ROC curves aimed to establish the optimal cut-off for each quantifiable parameter associated with the shifting from HN0 diagnosis to HN1 diagnosis, and with the shifting from HN1 diagnosis to HN2/3 diagnosis. Therefore, ROC curves were designed twice. In the first test, a ROC curve for each quantifiable parameter was built taking into account only cases with a corresponding histological

diagnosis of HN0 and HN1, considering HN1 diagnosis as positive. In the second test, a ROC curve for each quantifiable parameter was built taking into account only cases with a corresponding histological diagnosis of HN1 and HN2/3, considering HN1 diagnosis as positive. Resulting optimal cut-offs were then used to create an additional 3-classes based amendment of the AM1 system (named as “Amendment 1.1” – AM1.1).

The second set of ROC curves aimed to establish the optimal cut-off for each quantifiable parameter associated with the shifting from HN0/1 diagnosis to HN2/3 diagnosis. Therefore, a ROC curve for each quantifiable parameter was built taking into account all the cases, considering HN2/3 diagnosis as positive. Resulting optimal cut-offs were then used to create a 2-classes based amendment of the AM1 system (named as “Amendment 1.2” – AM1.2).

For each set of ROC curves, the optimal cut-off was established exclusively for those quantifiable parameters which p value was <.05. The area under the curve (AUC) was additionally used as an indicator of the accuracy of the analysis, being interpreted as follows: ≤0.50, test fail; 0.51-0.69, poor test; 0.70-0.79, fair test; 0.80-0.89, good test; 0.90-0.99, excellent test; 1.0, perfect test.[87] The optimal cut-off for the selected parameters which passed significance test was then chosen on the basis of the maximum Youden index (corresponding to the maximized sensitivity and specificity). The value of the optimal cut-off was then used to amend cytological interpretative system exclusively if reliable from a biological point of view according to previous literature.

After establishment of AM1.1 and AM1.2 systems, each Reader was asked again to establish a diagnosis according to each of the 2 amended systems on the basis of the data included in the corresponding Excel evaluation table. After revision of the diagnoses by two of the investigators (Reader 1 and Reader 3), the diagnostic accuracy indexes of both AM1.1 and AM1.2 system were calculated for each Reader and according to the median number of TP, TN, FP, and FN among the 5 Readers. If any of the median diagnostic accuracy indexes of AM1.1 and AM1.2 was evaluated as low, or if its evalaution was reduced twice compared to the corresponding index referring to the 3-classes simplified version of the system by Krick et al. (e.g.

specificity of the system by Krick et al. evaluated very high, and specificity of AM.1.1 evaluated as moderate), the entire amended system was considered unacceptable and discharged from further investigations. Conversely, for those amended systems considered acceptable, all the investigations described in the paragraphs “interrater agreement evaluation” and “statistical analysis for diagnostic accuracy” were applied to allow comparison of the perfomance with the 3-classes simplified version of the system proposed by Krick et al. and with AM.1 system. To investigate the interrater agreement in further details, Cohen’s kappa coefficient was also calculated for pairwise comparison of each couple of Readers. The switch from weighted kappa to Cohen’s kappa was determined by the binary otuput (i.e. NM or M) of the AM1.2 system compared to all the other systems, which were characterized by a non-binary output (e.g. NM, PM, and M). Cohen’s kappa was calculated using GraphPad QuickCalcs website (https://www.graphpad.com/quickcalcs/kappa1/).

RESULTS

The results and the discussion sections will take into account exclusively the median value among the 5 Readers of each investigated parameter. Nevertheless, the results for each Reader are reported in tables for completeness.

Selection workflow and signalment data for the cases included in the study

Among the 12602 cytological electronical records collected between January the 1st, 2008 and July the 31st, 2019, 8988 were referred to cytological samples obtained from dogs. Among the latter, 1400 records referred to at least one LN cytological sample.

Among the selected records, 1220 were excluded because referring to dogs lacking a clinical history of MCT, not bearing a MCT at the time of submission, simultaneously bearing a MCT and another kind of neoplasm at the time of submission, or lacking a detailed clinical history (category A). Finally, 180 records referring to 313 LN cytological samples obtained from MCTDs were selected. Among selected LNs, 146 were excluded due to lack of a corresponding histological sample.

Specifically, for 116 cases no histological sample of lymph node origin was available (category B.1), while for 30 cases only histological samples obtained from lymph nodes other than those cytologically sampled (category B.2) were available. Among the 167 selected cytological samples, only one was excluded due to unavaliable cytological slide for review (category C.1). Four (4) additional cytological samples were excluded because the associated histology was indicative of the presence of a neoplasm other than a MCT or of the absence of tissue consistent with lymph node (category D). Finally, 29 cases were excluded because the revised cytological slides were inconclusive (category E).

In conclusion, 133 cytological cases obtained from LNs of MCTDs were included in the study (133/162 cases; retrieval rate: 82.61%): 15 FNAs (11.28%), 52 scraping smears (39.10%), and 66 touch imprints (49.62%). The selection workflow applied in the current studies is summarized in Appendix 3 – Supplementary Figure 1.

The 133 cases included in the study were obtained from 84 different LNs, which in turn had been sampled from 47 MCTDs.

Of the 47 dogs included in the study, 43 (91.49%) were treated by the DIMEVET University Veterinary Hospital (OVU). Among included patients, 4 were intact females, 18 neutered females, 16 intact males, and 9 castrated males. The average age recorded was 7.8 years (range: 1-13 years). In addition to 11 mongrels, patients belonged to the following pure breeds: Labrador Retriever (7 dogs), Setter (6), Golden Retriever (5), Boxer (4), Pug Dog (3), and 11 other breeds represented by a single dog. Among the 47 dogs included, 38 had a single MCT, 4 dogs had 2 MCTs, and 5 had multiple MCTs (≥3), for a total of at least 66 MCTs. All reported MCTs were dermal and/or subcutaneous with the exception of one muco-cutaneous case (located in the upper lip). Twenty-six (26) MCTDs provided 1 lymph node each, 11 patients provided 2 lymph nodes, 8 patients provided 3 lymph nodes, and only 2 dogs provided 6 lymph nodes each.

Fifty-four (54) out of the 84 LNs included in the current study were identified as SLN through the application of LN mapping techniques before and/or after cytological sampling, while the remaining 30 were considered as RLN. Of the 84 LNs

investigated in the study, 12 were sampled via FNA alone, 20 were sampled via touch imprinting alone, and 6 via scraping smearing alone. Of the remaining 46 LNs, 3 were sampled using all the 3 techniques, and 43 were sampled via both touch imprinting and scraping smearing. Cytological samples had been submitted by the University Veterinary Hospital of the DIMEVET or by external private veterinary clinics, or had been prepared at the Diagnostic Pathology Service of the DIMEVET from surgical biopsies sent for other investigations or from necroscopic samples. In general, the samples had been sent to assess lymphadenomegaly conditions, or for neoplasm staging. In this regards, it is notewhorty that most of the samples prepared via scraping smearing and/or touch imprinting (94/133, 70.68%) were obtained from clinically non-palpable sentinel LNs surgically extirpated from MCTDs. The material for the cytological samples prepared via touch imprinting and/or scraping smearing at the DIMEVET Pathological Anatomy Department was obtained in most cases by sectioning the lymph node along the median longitudinal plane. The corresponding histological samples were prepared by surgical biopsies sent by the DIMEVET University Veterinary Hospital or by external veterinarians, or were collected during necropsies. The histological diagnoses of the 84 LNs included in the current study, expressed according to Weishaar et al. (Weishaar 2014), were the following: 29 HN0 LNs, 13 HN1, 24 HN2, and 18 HN3, corresponding to a metastasis prevalence of 50%. Considering that 46 LNs were cytologically sampled through 2 or 3 different sampling techniques, the 133 histological diagnoses corresponding to the 133 cytological cases included in the study were distributed as follows: 47 HN0 cases, 16 HN1, and 70 HN2/3 (of which 42 HN2 and 28 HN3), corresponding to a metastasis prevalence of 52.63%.

Details regarding the 133 cases finally included in the current study as well as signalment data of the 25 dogs from which they were obtained, are summarized in Appendix 2 – Supplementary Table 2.

Given the retrospective nature of the study, eventual correlations occurring between sex and age of investigated dogs, MCT location, LN location and size on one hand, and the diagnostic accuracy indexes calculated for each of the cytological interpretative system investigated, falled beyond the aims of the current study.

Nevertheless, further details regarding dogs, MCTs and LNs investigated in the current study are provided in Appendix 3 – Supplementary Table 1 for completeness.

Interrater agreement evaluation for the cytological interpretative system proposed by Krick et al.[104] *and for the AM1 system*

All cytological diagnoses expressed by each Reader according to each interpretative system investigated in the current study are tabulated togetehr with the corresponding histological diagnosis in Appendix 3 – Supplementary Table 2.

Specifically, the sum of each type of diagnosis expressed by each Reader according to the 5-classes system proposed by Krick et al.,[104] to the 3-classes simplified version of the system by Krick et al. (as illustrated in Appendix 2 – Supplementary Table 1), and to the AM1 system, is summarized in Table 2.

Table 2. Sum of each type of diagnosis expressed by each Reader according to all the cytological interpretative systems investigated in the current study.

Krick's system – 5 classes						Krick's system – 3 classes						AM1						AM1.2					
Dia.	**R1**	**R2**	**R3**	**R4**	**R5**	**Dia.**	**R1**	**R2**	**R3**	**R4**	**R5**	**Dia.**	**R1**	**R2**	**R3**	**R4**	**R5**	**Dia.**	**R1**	**R2**	**R3**	**R4**	**R5**
"1"	8	13	4	12	26	**NM**	84	67	77	82	76	**NM**	62	46	66	70	66	**NM**	79	60	79	86	80
"2"	76	54	73	70	50																		
"3"	2	7	4	6	7	**PM**	2	7	4	6	7	**PM**	15	14	9	13	11						
"4"	24	27	25	16	19													**M**	54	73	54	47	53
"5"	23	32	27	29	31	**M**	47	59	52	45	50	**M**	56	73	58	50	56						

Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system – 5 classes, cytological interpretative system proposed by Krick et al.[104] in its orginal 5-classes version; Krick's system – 3 classes, 3-classes simplified version of the system by Krick et al. as illustrated in Appendix 2 – Supplementary Table 1; M. metastatic; NM, non metastatic; PM, possibly metastatic; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; 1, "normal"; 2, "reactive lymphoid hyperplasia"; 3, "possible metastasis"; 4, "probable metastasis"; 5, "certain metastasis".

The interrater agreement among the 5 Readers for each of the investigated interpretative system is illustrated in Table 3.

Table 3. Free-marginal kappa for interrater agreement evaluation among the 5 Readers for each of the cytological interpretative systems investigated in the current study.

	Krick's system - 5 classes	Krick's system - 3 classes	AM1	AM1.2
Free-marginal K	0.62	0.82	0.71	0.80
(95% C.I.)	(0.56 - 0.67)	(0.77 - 0.87)	(0.63 - 0.78)	(0.74-0.87)

Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system – 5 classes, cytological interpretative system proposed by Krick et al.[104] in its orginal 5-classes version; Krick's system – 3 classes, 3-classes simplified version of the system by Krick et al. as illustrated in Appendix 2 – Supplementary Table 1; 95% C.I., 95% confidence interval.

Specifically, the interrater agreement for the 5-classes original version of the system by Krick et al. [104] was moderate while that for the 3-classes simplified version of the same system proposed by our research group was strong. On the contrary, the AM1 system proposed by our research group was moderate.

When interrater agreement was investigated calculating weighted kappa for each couple of Readers (results illustrated in details in Appendix 3 Supplementary Table 3), the 5-classes original version of the system by Krick et al.[104] showed in most of the cases a moderate agreement (range of K: 0.681-0.795), with the exception of a pairwise comparison (Reader 1 vs Reader 3) for which interrater agreement was strong (K: 0.809). On the other hand, the 3-classes simplified version of the system by Krick et al. showed a strong agreement in most of the pairwise comparisons (range of K: 0.825-0.889), with the exception of 3 couples (Reader 1 vs Reader 2, Reader 2 vs Reader 4, and Reader 2 vs Reader 5) for which interrater agreement was moderate (range of K: 0.756-0.778). Finally, interrater agremeent for AM1 system was moderate in most of the cases (range of K: 0.624-0.787), with the exception of a strong agreement for 3 couples of Readers (Reader 3 vs Reader 4, and Reader 3 vs Reader 4; range of K: 0.817-0.863).

Diagnostic accuracy evaluation of the cytological interpretative system proposed by Krick et al.[104] *and of the AM1 system*

All cytological diagnoses expressed by each Reader according to each interpretative system investigated in the current study are tabulated together with the corresponding histological diagnosis in Appendix 3 – Supplementary Table 2.

Additionally, the sum of each type of diagnosis expressed by each Reader according to the 3-classes simplified version of the system by Krick et al. (as illustrated in Appendix 2 – Supplementary Table 1), and according to the AM1 system, is summarized in Table 2.

The Reader-specific and median overall percentage agreement between cytology and histopathology is illustrated in Table 4. Cases were classified as “in complete agreement” (CA) with histopathology when a NM, PM, or M cytological diagnosis corresponded to a HN0, HN1, or HN2/3 histological diagnosis, respectively. All the other combinations of cytological and histological diagnosis were classififed as “not in agreement” (NA).

Table 4. Reader-specific and median percentage overall agreement for all the cytological intepretative systems investigated in the current study.

		R1	R2	R3	R4	R5	Median
Krick's system	**CA**	65.41 (-)	66.92 (-)	69.17 (-)	66.17 (-)	65.41 (-)	66.17 (-)
	NA	34,59	33.08	30.83	33.83	34.59	33.83
AM1	**CA**	66.17 (-)	61.65 (-)	69.92 (-)	63.91 (-)	69.92 (-)	66.17 (-)
	NA	33.83	38.35	30.08	36.09	30.08	33.83
AM1.2	**CA**	**80.45 (++)**	**75.19 (+)**	**80.45 (++)**	**78.20 (+)**	**81.20 (++)**	**80.45 (++)**
	NA	19.55	24.81	19.55	21.80	18.80	19.55

Values are expressed in percentages. In bold and highlighted in yellow, the highest value among the interpretative system evaluated. In green, those values of AM1 and/or AM1.2 systems which corresponding intepretation improved compared to the Krick’s system. In red, those values of AM1 and/or AM1.2 systems which corresponding intepretation worsened compared to the Krick’s system. Interpretation classes: (-), low; (+), moderate; (++), high; (+++), very high. Legend: AM1, amended system 1; AM1.2, amended system 1.2; CA, complete agreement (i.e. proprotion of “non metastatic”, “possibly metastatic”, or “metastatic” diagnoses respectively corresponding to HN0, HN1, or HN2/3 diagnosis); Krick’s system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; NA, no agreement; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

The median overall agreement calculated among the 5 Readers was intepreted as low for both the 3-classes simplified version of the system by Krick et al. and the AM1 system, highlighting that both systems should be improved.

The Reader-specific and median percentage agreement between cytology and histopathology grouping the cytological cases on the basis of the corresponding histological diagnosis is illustrated in Table 5. Specifically, the median percentage agreement for cytological cases which corresponding histological diagnosis was HN0, was judjed very high for the system by Krick et al.[104] (91.49%) and moderate

for the AM1 system (78.72%), suggesting a tendency of the latter system to overestimate the cytological diagnosis. The median percentage agreement for cytological cases which corresponding histological diagnsosis was HN1, was judged low for both the system by Krick et al. (6.25%) and the AM1 system (18.75%), highlighting that both systems are inadequate in the evaluation of this kind of specimens. Finally, the median percentage agreement for cytological cases which corresponding histological diagnsosis was HN2/3, was judjed low for the system by Krick et al. (65.71%) and moderate for the AM1 system (71.43%), suggesting that the latter system is more precise in the evaluation of this cytological samples despite still not very accurate. Chi-squared test confirmed these observations (results illustrated in details in Appendix 3 - Supplementary Table 4). Indeed, when the system by Krick et al. was applied, a significantly higher median percentage agreement in the diagnosis of cases with a corresponding HN0 histology compared to both cases with a corresponding histology diagnosed as HN1 or HN2/3 was observed, as well as a significantly lower median percentage agreement for cases with a corresponding HN1 histology compared to those with a corresponding HN2/3 histology. Similarly, when the AM1 system was applied, a significantly lower median percentage agreement in the diagnosis of cases with a corresponding HN1 histology compared to both cases with a corresponding histology diagnosed as HN1 or HN2/3 was observed. On the other hand, no significant differences in the median percentage agreement between cases with a corresponding histology of HN0 and HN2/3 was observed, probably due to the slight increase in the accuracy of the AM1 system in the diagnosis of the latter group of cases compared to the system by Krick et al.

The Reader-specific and median percentage agreement between cytology and histopathology grouping the cytological cases on the basis of the corresponding sampling technique is illustrated in Table 6. For both the system by Krick et al.[104] and the AM1 system, percentage agreement was constantly judged low, disregarding the sampling technique investigated. Nevertheless, it is notehworthy that the percentage agreement for cytological cases sampled via touch imprinting and scraping smearing approached the threshold of 70% to be judged as moderate (69.70% and 68.18% for touch imprints, and 69.23% and 67.31% for scraping smears, respectively applying the system by Krick et al. and the AM1 system). On the other hand, FNAs showed a

lower percentage agreement compared to the other sampling technique, which was slighlty higher for the AM1 system (53.33%) compared to the system by Krick et al. (40%). These observations were confirmed by chi-squared test, which revealed a signifcantly lower percentage agreement of FNAs compared to both touch imprints and scraping smears applying the system by Krick et al., but not applying the AM1 system.

When the concordance between the cytological diagnoses of the cases referring to the same LN but sampled with different techniques was investigated (Appendix 3 – Supplementary Table 6), the system by Krick et al.[104] showed a higher number of cases with complete intersampling coherence (86.67%) compared to AM1 system (80.00%). In addition, the AM1 system showed a lower number of cases with complete intersampling coherence correctly identifying the metastatic status (55.56% compared to 62.22% of the system by Krick et al.), and a higher number of cases with intersampling incoherence (20.00% compared to 11.11% of the system by Krick et al.). This observation might be indicative of a lower interchangeability of the AM1 system compared to that proposed by Krick et al. among different sampling techniques. Nevertheless, both the systems showed a 24.44% of cases with complete intersampling agreement but expressing the wrong diagnosis, suggesting that a certain amount of wrong diagnoses was the consequence of the cytological sampling in general compared to the histological reference standard.

Further investigation of the general agreement between cytology and histopathology with weighted kappa statistics revealed a weak agreement with histology for both the interpretative system by Krcik et al. (K: 0.520) and the AM1 system (K: 0.514).

Table 5. Reader-specific and median percentage agreement grouping cytological cases according to the corresponding histological diagnosis for all the cytological intepretative systems investigated in the current study.

		R1			R2			R3			R4			R5			Median		
		HNO	NH1	HN2/3	HNO	NH1	HN2/3	HNO	NH1	HN2/3	HNO	NH1	HN2/3	HNO	NH1	HN2/3	HNO	NH1	HN2/3
Krick's system	CA	**93.62 (+++)**	0 (-)	61.43 (-)	**80.85 (++)**	6.25 (-)	71.43 (+)	**91.49 (+++)**	12.5 (-)	67.14 (-)	**95.74 (+++)**	6.25 (-)	60.00 (-)	87.23 (++)	0.00 (-)	65.71 (-)	**91.49 (+++)**	6.25 (-)	65.71 (-)
	NA	6.38	100.00	38.57	19.15	93.75	28.57	8.51	87.50	32.86	4.26	93.75	40.00	12.77	100.00	34.29	8.51	93.75	34.29
AM1	CA	74.47 (+)	18.75 (-)	**71.43 (+)**	57.45 (-)	6.25 (-)	77.14 (+)	82.98 (++)	25 (-)	**71.43 (+)**	78.72 (+)	12.5 (-)	**65.71 (-)**	82.98 (++)	25.00 (-)	**71.43 (+)**	78.72 (+)	18.75 (-)	**71.43 (+)**
	NA	25.53	81.25	28.57	42.55	93.75	22.86	17.02	75.00	28.57	21.28	87.50	34.29	17.02	75.00	28.57	21.28	81.25	28.57
AM1.2	CA	91.49 (+++)	**93.75 (+++)**	70.00 (+)	68.09 (+)	**75.00 (+)**	**80.00 (++)**	**91.49 (+++)**	**93.75 (+++)**	70.00 (+)	**95.74 (+++)**	**93.75 (+++)**	62.86 (-)	**93.62 (+++)**	**93.75 (+++)**	70.00 (+)	**91.49 (+++)**	**93.75 (+++)**	70.00 (+)
	NA	8.51	6.25	30.00	31.91	25.00	20.00	8.51	6.25	30.00	4.26	6.25	37.14	6.38	6.25	30.00	8.51	6.25	30.00

Values are expressed in percentages. In bold and highlighted in yellow, the highest value among the interpretative system evaluated. In green, those values of AM1 and/or AM1.2 systems which corresponding intepretation improved compared to the Krick's system. In red, those values of AM1 and/or AM1.2 systems which corresponding intepretation worsened compared to the Krick's system. Interpretation classes: (-), low; (+), moderate; (++), high; (+++), very high. Legend: AM1, amended system 1; AM1.2, amended system 1.2; CA, complete agreement (i.e. proprotion of "non metastatic", "possibly metastatic", or "metastatic" diagnoses respectively corresponding to HN0, HN1, or HN2/3 diagnosis); HNO, "non-metastatic" according to Weishaar et al.;[234] HN1, "pre-metastatic" according to Weishaar et al.; HN2, "early metastasis" according to Weishaar et al.; HN3, "overt metastasis" according to Weishaar et al.; Krick's system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; NA, no agreement; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

Table 6. Reader-specific and median percentage agreement grouping cytological cases according to the corresponding sampling technique for all the cytological interpretative systems investigated in the current study.

		R1			R2			R3			R4			R5			Median		
		FNA	TIC	SCRA	FNA	TIC	SCRA	FNA	TIC	SCRA	FNA	TIC	SCRA	FNA	TIC	SCRA	FNA	TIC	SCRA
Krick's system	CA	40.00 (-)	69.7 (-)	67.31 (-)	46.67 (-)	69.7 (-)	**69.23 (-)**	46.67 (-)	71.21 (+)	73.08 (+)	40.00 (-)	69.7 (-)	69.23 (-)	40.00 (-)	69.7 (-)	67.31 (-)	40.00 (-)	69.70 (-)	69.23 (-)
	NA	60.00	30.30	32.69	53.33	30.30	30.77	53.33	28.79	26.92	60.00	30.30	30.77	60.00	30.30	32.69	60.00	30.30	30.77
AM1	CA	53.33 (-)	68.18 (-)	67.31 (-)	53.33 (-)	65.15 (-)	59.62 (-)	53.33 (-)	71.21 (+)	73.08 (+)	46.67 (-)	66.67 (-)	65.38 (-)	60.00 (-)	71.21 (+)	71.15 (+)	53.33 (-)	68.18 (-)	67.31 (-)
	NA	46.67	31.82	32.69	46.67	34.85	40.38	46.67	28.79	26.92	53.33	33.33	34.62	40.00	28.79	28.85	46.67	31.82	32.69
AM1.2	CA	**66.67 (-)**	**83.33 (++)**	**80.77 (++)**	**80.00 (++)**	**81.82 (++)**	65.38 (-)	**66.67 (-)**	**81.82 (++)**	**82.69 (++)**	**66.67 (-)**	**81.82 (++)**	**76.92 (+)**	**73.33 (+)**	**81.82 (++)**	**82.69 (++)**	**66.67 (-)**	**81.82 (++)**	**80.77 (++)**
	NA	33.33	16.67	19.23	20.00	18.18	34.61	33.33	18.18	17.31	33.33	18.18	23.08	26.67	18.18	17.31	33.33	18.18	19.23

Values are expressed in percentages. In bold and highlighted in yellow, the highest value among the interpretative system evaluated. In green, those values of AM1 and/or AM1.2 systems which corresponding interpretation improved compared to the Krick's system. In red, those values of AM1 and/or AM1.2 systems which corresponding interpretation worsened compared to the Krick's system. Interpretation classes: (-), low; (+), moderate; (++), high; (+++), very high. Legend: AM1, amended system 1; AM1.2, amended system 1.2; CA, complete agreement (i.e. proprotion of "non metastatic", "possibly metastatic", or "metastatic" diagnoses respectively corresponding to HN0, HN1, or HN2/3 diagnosis); FNA, fine needel aspirates; Krick's system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; NA, no agreement; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; SCRA, scraping smears; TIC, touch imprints.

Details regarding the number of cases classified as TP, FP, FN, and TN for each interpretative system investigated in the current study and included in 2x2 contingency tables for McNemar's test and for calculation of diagnostic accuracy indexes, are reported in Appendix 3 – Supplementary Table 8.

In the best case scenario (i.e. when "possibly metastatic" cases were considered as negative for metastastasis), McNemar's test reached a p value <.05 for both the system by Krick et al. (p = 0.0003) and the AM1 system (p = 0.0108), revealing a significantly different number of cases diagnosed as positive between cytology and histology, and thus the presence of a test bias. Details regarding McNemar's test results for each Reader and in the worst case scenario (i.e. when "possibly metastatic" cases were considered as positive for metastasis) are reported in Appendix 3 – Supplementary Table 9.

Diagnostic accuracy indexes in the best case scenario for the system by Krick et al. and for the AM1 system are summarized in Table 7. Details regarding diagnostic accuracy indexes in the worst case scenario and 95% confidence interval of each index are reported in Appendix 3 – Supplementary Table 10.

Table 7. Diagnostic accuracy indexes in the best case scenario for all the cytological interpretative systems investigated in the current study.

Evaluator	Interpretative system	ACCURACY	SENSITIVITY	SPECIFICITY	POSITIVE PREDICTIVE VALUE	NEGATIVE PREDICTIVE VALUE
R1	**Krick's system**	76.69 (+)	61.43 (-)	**93.65 (+++)**	**91.49 (+++)**	68.6 (-)
	AM1	**80.45 (++)**	**71.43 (+)**	90.48 (+++)	89.29 (++)	74.03 (+)
	AM1.2	**80.45 (++)**	70 (+)	92.06 (+++)	90.74 (+++)	**73.42 (+)**
R2	**Krick's system**	**75.94 (+)**	68.49 (-)	**85 (++)**	**84.75 (++)**	68.92 (-)
	AM1	73.68 (+)	77.14 (+)	69.84 (-)	73.97 (+)	73.33 (+)
	AM1.2	75.19 (+)	**78.57 (+)**	71.43 (+)	75.34 (+)	**75 (+)**
R3	**Krick's system**	78.95 (+)	67.14 (-)	**92.06 (+++)**	**90.38 (+++)**	71.6 (+)
	AM1	78.95 (+)	**71.43 (+)**	87.3 (++)	86.21 (++)	73.33 (+)
	AM1.2	**80.45 (++)**	70 (+)	**92.06 (+++)**	90.74 (+++)	**73.42 (+)**

Evaluator	Interpretative system	ACCURACY	SENSITIVITY	SPECIFICITY	POSITIVE PREDICTIVE VALUE	NEGATIVE PREDICTIVE VALUE
R4	**Krick's system**	76.69 (+)	60 (-)	**95.24 (+++)**	93.33 (+++)	68.18 (-)
	AM1	**78.95 (+)**	**65.71 (-)**	93.65 (+++)	92 (+++)	**71.08 (+)**
	AM1.2	78.2 (+)	62.86 (-)	**95.24 (+++)**	**93.62 (+++)**	69.77 (-)
R5	**Krick's system**	78.95 (+)	65.71 (-)	**93.65 (+++)**	92 (+++)	71.08 (+)
	AM1	80.45 (++)	**71.43 (+)**	90.48 (+++)	89.29 (++)	**74.03 (+)**
	AM1.2	**81.2 (++)**	70 (+)	**93.65 (+++)**	**92.45 (+++)**	73.75 (+)
Median	**Krick's system**	78.95 (+)	65.71 (-)	**93.65 (+++)**	**92.00 (+++)**	71.08 (+)
	AM1	**80.45 (++)**	**71.43 (+)**	90.48 (+++)	89.29 (++)	**74.03 (+)**
	AM1.2	**80.45 (++)**	70 (+)	92.06 (+++)	90.74 (+++)	73.42 (+)

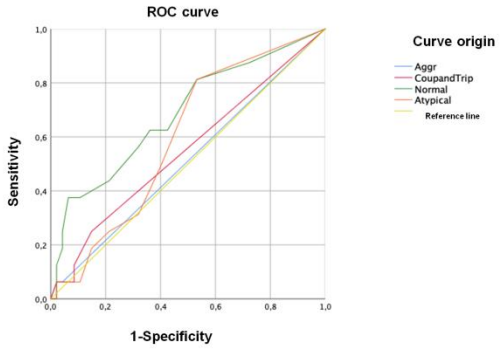
Values are expressed in percentages. In bold and highlighted in yellow, the highest value among the interpretative system evaluated. In green, those values of AM1 and/or AM1.2 systems which corresponding intepretation improved compared to the Krick's system. In red, those values of AM1 and/or AM1.2 systems which corresponding interpretation worsened compared to the Krick's system. Interpretation classes: (-), low; (+), moderate; (++), high; (+++), very high. Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

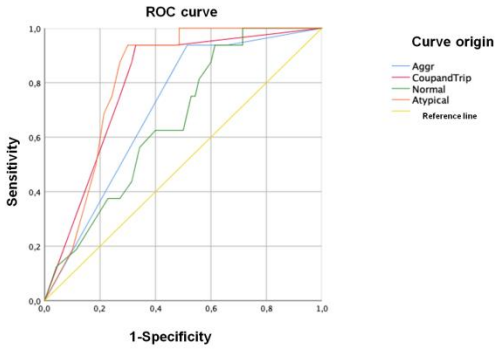
The system by Krick et al.[104] showed a moderate median accuracy (78.95%; C.I. 71,03-85,53%), a low median sensitivity (65.71%; C.I. 53,40-76,65%), a very high median specificity (93.65%; C.I. 84,53-98,24%) and positive predictive value (92.00%; C.I. 81,44-96,79%), and a moderate negative predictive value (71.08%; C.I. 63,85-77,38%). The AM1 system was characterized by a high median accuracy (80.45%; C.I. 72,68-86,81%), a moderate median sensitivity (71.43%; C.I. 59,38-81,60%), a very high median specificity (90.48%; C.I. 80,41-96,42%), a high median positive predictive value (89.29%; C.I. 79,33-94,76%), and a moderate negative predictive value (74.03%; C.I. 66,11-80,63%). When compared to the interpretative system by Krick et al., the AM1 system showed a sensible increase in accuracy and sensitivity, in association with a sensibly decreased positive predictive value.

Setting of additional amended cytological interpretative systems with ROC curves.

ROC curves designed to set the 3-classes AM1.1 system, as well as the corresponding AUC and p values, are illustrated in Figure 1.

Figure 1. Receiver operating characteristic (ROC) curves for optimal cut-off establishment of the 3-classes AM1.1 cytological interpretative system.





A) ROC curves for quantifiable parameters considering HN0 as negative and HN1 as positive

Quantifiable parameter	AUC	Standard error	p value	95% C.I.	
				Lower limit	Upper limit
Aggr	0,511	0,085	0,899	0,344	0,677
CoupandTrip	0,549	0,086	0,564	0,380	0,717
Normal	0,685	0,080	0,028	0,528	0,842
Atypical	0,597	0,076	0,249	0,448	0,747

B) AUC and p value for quantifiable parameters considering HN0 as negative and HN1 as positive

C) ROC curves for quantifiable parameters considering HN1 as negative and HN2/3 as positive

Quantifiable parameter	AUC	Standard error	p value	95% C.I.	
				Lower limit	Upper limit
Aggr	0,707	0,063	0,010	0,583	0,831
CoupandTrip	0,796	0,055	0,000	0,688	0,903
Normal	0,660	0,066	0,046	0,530	0,790
Atypical	0,814	0,046	0,000	0,724	0,904

D) AUC and p value for quantifiable parameters considering HN1 as negative and HN2/3 as positive

Legend: Aggr, aggregates composed by >3 MCs; Atypical, morphologically atypical MCs; AUC, area under the curve; CoupandTrip, couples + triplets of MCs; HNO, "non-metastatic" according to Weishaar et al.;[234] HN1, "pre-metastatic" according to Weishaar et al.;[234] HN2/3, "early metastasis" or "overt metastasis" according to Weishaar et al.; MCs, mast cells; Normal, morphologically normal MCs; 95% C.I., 95% confidence interval.

The first subset of ROC curves to establish the optimal cut-off to shift from a corresponding HN0 histological diagnosis to a HN1 one, revealed that the only significant parameter was the number of discrete, morphologically normal MCs, which was considered a poor test (AUC = 0.685; p value = 0.028). The corresponding optimal cut-off was >7 normal MCs / 8 HPFs. The second subset of ROC curves to establish the optimal cut-off to shift from a corresponding HN1 histological diagnosis to a HN2/3 one, revealed that all investigated parameters were significant. Specifically, the number of aggregates of >3 MCs was a fair test (AUC = 0.707; p value = 0.010), with an optimal cut-off of ≥1 aggregate / entire slide. Also the number of couples and triplets of MCs was a fair test (AUC = 0.796; p value < 0.001), with an optimal cut-off of ≥3 couples and triplets / entire slide. The number of normal MCs was a poor test (AUC = 0.660; p value = 0.046), with an optimal cut-off of ≥14 normal MCs / 8 HPFs. Finally, the number of discrete, atypical MCs was a good test (AUC = 0.814; p value < 0.001), with an optimal cut-off of ≥7 atypical MCs / 8 HPFs.

According to the optimal cut-offs established with ROC curves, the AM1.1 interpretative system was designed as illustrated in Table 8.

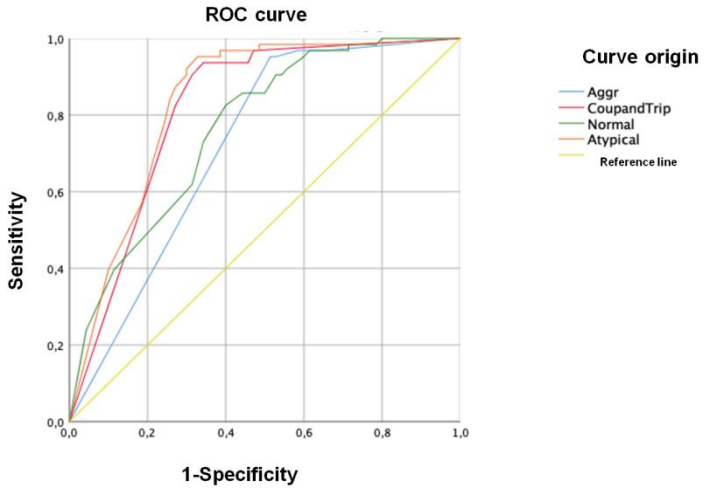
Table 8. Criteria of the 3-classes AM1.1 system created designing two subsets of ROC curves for quantifiable parameters evaluated in the current study.

Amended 1.1 (AM1.1) system – Cytological criteria	Diagnosis
• 0-1 couple or triplet of MCs / entire slide • 0-7 discrete, normal MCs / 8 HPFs	*Non metastatic* (NM)
• 2 couples or triplets of MCs / entire slide • 8-13 discrete, normal MCs / 8 HPFs	*Possibly metastatic* (PM)
• Effacement of lymphoid tissue by MCs (i.e. >50% MCs on the total cellular amount/entire slide) • ≥1 aggregate of >3 MCs / entire slide • ≥3 couples or triplets of MCs / entire slide • ≥7 discrete, atypical MCs (bi- or mutlinucleation, nuclear pleomorphism, altered N:C ratio, anisokaryosis, anisocytosis, variable/reduced cytoplasmic granulation) / 8 HPFs • ≥ 14 discrete, normal MCs / 8 HPFs	*Metastatic* (M)

To express a specific diagnosis, the fulfilling of at least one of the criterion was considered necessary and sufficient. Legend: HPFs, high power fields; MCs, mast cells; N:C, nuclear cytoplasmic ratio.

ROC curves designed to set the 2-classes AM1.2 system, as well as the corresponding AUC and p values, are illustrated in Figure 2.

Figure 2. Receiver operating characteristic (ROC) curves for optimal cut-off establishment of the 2-classes AM1.2 cytological interpretative system.



A) ROC curves for quantifiable parameters considering HN0/1 as negative and HN2/3 as positive

Quantifiable parameter	AUC	Standard error	p value	95% C.I.	
				Lower limit	Upper limit
Aggr	0,720	0,045	0,000	0,633	0,807
CoupandTrip	0,816	0,038	0,000	0,742	0,891
Normal	0,764	0,040	0,000	0,685	0,844
Atypical	0,836	0,036	0,000	0,765	0,906

B) AUC and p value for quantifiable parameters considering HN0/1 as negative and HN2/3 as positive

Legend: Aggr, aggregates composed by >3 MCs; Atypical, morphologically atypical MCs; AUC, area under the curve; CoupandTripl, couples + triplets of MCs; HNO/1, non-metastatic" or "pre-metastatic" according to Weishaar et al.;[234] HN2/3, "early metastasis" or "overt metastasis" according to Weishaar et al.; MCs, mast cells; Normal, morphologically normal MCs; 95% C.I., 95% confidence interval.

ROC curves designed to establish the optimal cut-off to shift from a corresponding HN0/1 histological diagnosis to a HN2/3 one, revealed that all investigated parameters were significant. Specifically, the number of aggregates of >3 MCs was a fair test (AUC = 0.720; p value < 0.001), with an optimal cut-off of ≥1 aggregate / entire slide. The number of couples and triplets of MCs was a good test (AUC = 0.816; p value < 0.001), with an optimal cut-off of >3 couples and triplets / entire slide. The number of normal MCs was a fair test (AUC = 0.764; p value < 0.001), with an optimal cut-off of ≥7 normal MCs / 8 HPFs. Finally, the number of atypical MCs was a good test (AUC = 0.836; p value < 0.001), with an optimal cut-off of >9 atypical MCs / 8 HPFs.

According to previous literature [12,25,116,145] and our preliminary study illustrated in Section 1 of the current Thesis, the number of normal MCs observed in LNs obtained from dogs without MCT nodal metastasis might be higher than 7. An optimal cut-off of 7 normal MCs to shift from a cytological diagnosis of "non-metastatic" to "metastatic" would be further in contrast with the cut-off of >9 atypical MCs, considered that in the literature is reported that no atypical MCs should be seen in a non metastatic nodal cytological specimen. Given all these consideration, the cut-off of >7 normal MCs was considered as a missense from a biological point of view and was thus discharged from the final version fo the AM1.2 system, which is illustrated in Table 9.

Table 9. Criteria of the 2-classes AM1.2 system created designing a set of ROC curves for quantifiable parameters evaluated in the current study.

Amended 1.2 (AM1.2) system – Cytological criteria	Diagnosis
• 0-3 couples or triplets of MCs / entire slide • 0-9 discrete, atypical MCs (bi- or mutlinucleation, nuclear pleomorphism, altered N:C ratio, anisokaryosis, anisocytosis, variable/reduced cytoplasmic granulation) / 8 HPFs	*Non metastatic* (NM)
• Effacement of lymphoid tissue by MCs (i.e. >50% MCs on the total cellular amount/entire slide) • ≥1 aggregate of >3 MCs / entire slide • ≥4 couples or triplets of MCs / entire slide • ≥10 discrete, atypical MCs / 8 HPFs	*Metastatic* (M)

To express a specific diagnosis, the fulfilling of at least one of the criterion was considered necessary and sufficient. Legend: HPFs, high power fields; MCs, mast cells; N:C, nuclear cytoplasmic ratio.

Validation of the AM1.1 and AM1.2 systems through calculation of the corresponding diagnostic accuracy indexes and comparison with those of the interpretative system by Krick et al.[104]

All cytological diagnoses expressed by each Reader according to AM1.1 and AM1.2 are tabulated together with the corresponding histological diagnosis in Appendix 3 – Supplementary Table 2.

Details regarding the number of TP, FP, FN, and TN for both AM1.1 and AM1.2 system, are reported in Appendix 3 – Supplementary Table 8, while details regarding

the results of McNemar's test are illustrated in Appendix 3 – Supplementary Table 9. The diagnostic accuracy indexes together with the corresponding 95% confidence interval for both AM1.1 and AM1.2 system, are reported in Appendix 3 – Supplementary Table 10.

In the best case scenario, the AM1.1 system was characterized by a moderate median accuracy (75.94%; C.I. 67.77-82.92%), a moderate median sensitivity (72.86%; C.I. 60.90-82.80%), a moderate median specificity (79.37%; C.I. 67.30-88.53%), a moderate median positive predictive value (79.69%; C.I. 70.31-86.67%), and a moderate negative predictive value (72.46%; C.I. 63.73-79.76%). When compared to the interpretative system by Krick et al., the AM1 system showed a sensible increase in sensitivity, which was on the other hand associated with 2-fold decreased specificity and positive predictive value. Considered this observation and according to the validation criteria illustrated in the material and methods section, the AM1.1 system was considered unacceptable and thus discharged from further investigations.

On the other hand, in the best case scenario the AM1.2 system was characterized by a high median accuracy (80.45%; C.I. 72,68-86,81%), a moderate median sensitivity (70.00%; C.I. 57.87-80.38%), a very high median specificity (92.06%; C.I. 82.44-97.37%), a very high median positive predictive value (90.74%; C.I. 80.65-95.84%), and a moderate negative predictive value (73.42%; C.I. 65.72-79.92%). When compared to the interpretative system by Krick et al., the AM1 system showed a sensible increase in accuracy and sensitivity, without any other substantial change. Considered this observation, AM1.2 system was considered as acceptable and completely comparable with the system by Krick et al. and the AM1 system. In this context, it is noteworthy that diagnostic accuracy indexes for the AM1.2 system were pretty similar to those for the AM1 system, with the notable difference of a sensibly higher positive predictive value.

Interrater agreement evaluation for the AM1.2 cytological interpretative system compared to the one by Krick et al.[104] *and to the AM1 system*

The sum of each type of diagnosis expressed by each Reader according to the AM1.2 system, is summarized in Table 2.

As illustrated in Table 3, the AM1.2 system showed a strong interrater agreement among the 5 Readers (K: 0.80), which was higher to that of the original 5-classes version of the system by Krick et al.[104] (K:0.62) and to that of the AM1 system (K: 0.71). On the other hand, interrater agreement for the AM1.2 system was in line with that of the simplified 3-classes version of the system by Krick et al. (0.82).

When interrater agreement was investigated calculating Cohen's kappa for each couple of Readers (results illustrated in details in Appendix 3 Supplementary Table 3), the 5-classes original version of the system by Krick et al.[104] showed in most of the cases a weak agreement (range of K:0.452-0.558) with the exception of 3 comparisons (Reader 1 vs Reader 3, Reader 1 vs Reader 4, and Reader 3 vs Reader 4) for which interrater agreement was moderate (range of K: 0.608-0.645). On the other hand, the 3-classes simplified version of the system by Krick et al. showed a moderate agreement in most of the pairwise comparisons (range of K: 0.681-0.792), with the exception of the same aforementioned couples (Reader 1 vs Reader 3, Reader 1 vs Reader 4, and Reader 3 vs Reader 4) for which interrater agreement was strong (range of K: 0.816-0.848). Interrater agremeent for the AM1 system was moderate in most of the cases (range of K: 0.612-0.749), with the exception of a weak agreement for 2 couples of Readers (Reader 1 vs Reader 2, and Reader 2 vs Reader 4; range of K: 0.537-0.558), and of a strong agreement for a single couple of Readers (Reader 3 vs Reader 5; K: 0.827). Finally, the interrater agreement for the AM1.2 system was moderate for all the four couples including Reader 2 (range of K: 0.620-0.719), strong for four couples (Reader 1 vs Reader 3, Reader 1 vs Reader 4, Reader 1 vs Reader 5, and Reader 3 vs Reader 4; range of K: 0.857-0.889), and almost perfect for the 2 residual couples (Reader 3 vs Reader 4, and Reader 4 vs Reader 5; range of K: 0.904-0.922).

Evaluation of other diagnostic accuracy parameters for the AM1.2 cytological interpretative system

The Reader-specific and median overall percentage agreement between cytology and histopathology is illustrated in Table 4. The median overall agreement calculated among the 5 Readers for the AM1.2 system was very high (80.45%), highlighting a sensible improvement compared to both the system by Krick et al.[104] (66.17%) and the AM1 system (66.17%).

The Reader-specific and median percentage agreement between cytology and histopathology grouping the cytological cases on the basis of the corresponding histological diagnosis is illustrated in Table 5. The median percentage agreement for cytological cases which corresponding histological diagnosis was HN0, was judged very high for the AM1.2 system, overlapping the one for system by Krick et al.[104] (91.49%) and being sensibly better than that for the AM1 system (78.72%), confirming a certain tendency for the latter system to overestimate the cytological diagnosis. The median percentage agreement for cytological cases which corresponding histological diagnosis was HN1, was very high for the AM1.2 system (93.75%) and sensibly improved compared to that for the system by Krick et al. (6.25%) and for the AM1 system (18.75%). This observation highlighted that AM1.2 system might represent a good alternative for the evaluation of these samples. Finally, the median percentage agreement for cytological cases which corresponding histological diagnosis was HN2/3, was moderate for the AM1.2 system (70.00%), being sensibly higher than that for the system by Krick et al. (65.71%) and similar to that for the AM1 system (71.43%). Chi-squared test confirmed these observations (results illustrated in details in Appendix 3 - Supplementary Table 4), revealing a significant difference exclusively between the percentage agreement for cases with a corresponding HN0 histology and those with a corresponding HN2/3 histology.

The Reader-specific and median percentage agreement between cytology and histopathology grouping the cytological cases on the basis of the corresponding sampling technique is illustrated in Table 6. The AM1.2 system showed a sensible increase in the percentage agreement for touch imprints (81.82%) and scraping smears (80.77%) compared to both the system by Krick et al.[104] (69.70% for touch

imprints and 69.23% for scraping smears) and the AM1 system (68.18% for touch imprints and 67.31% for scraping smears). On the other hand, despite higher than the values reported for the system by Krick et al. (40.00%) and the AM1 system (53.33%), percentage agreement for FNAs applying the AM1.2 system was still low (66.67%). Chi-squared test revealed no signifcantly different percentage agreement among the different sampling techniques when the AM1.2 system was applied.

When the concordance between the cytological diagnoses of the cases referring to the same LN but sampled with different techniques was investigated (Appendix 3 – Supplementary Table 6), the AM1.2 system showed a higher number of cases with complete intersampling coherence (89.36%) compared to that for the system by Krick et al.[104] (86.67%) and for the AM1 system (80.00%). For the AM1.2 system, also the number of cases with complete intersampling coherence correctly identifying the metastatic status was higher (74.47%) compared to that of the system by Krick et al. and the AM1 system (62.22% and 55.56%, respectively). Nevertheless, despite a lower percentage of cases with intersampling incoherence (8.51% compared to 11.11% for the system by Krick et al. and 20% for the AM1 system), the AM1.2 system still showed a 14.89% cases with complete intersampling agreement but expressing the wrong diagnosis. Similarly, when the agreement among the diagnoses expressed according to each cytological interpretative system applied was investigated, 19.08% cases were cytologically diagnosed in the same way applying the 3 different systems, but still no agreement was observed with the corresponding histopathological diagnosis. This observations further strenghtened the hypothesis that a certain amount of wrong diagnoses was the consequence of the cytological sampling in general compared to the histological reference standard. This hypothesis was finally confirmed by the observation that also for the AM1.2 system, McNemar's test showed a signifcant p value ($p = 0.0003$), stressing out the presence of a test bias (Appendix 3 – Supplementary Table 8).

Further investigation of the general agreement between cytology and histopathology with kappa statistics required to calculate Cohen's kappa for all the cytological interpretative system investigated due to the binary output of the AM1.2 system (i.e. NM or M). According to Cohen's kappa coefficients, the general agreement between

cytology and histopathology was weak for both the system by Krick et al.[104] (K: 0.467; C.I: 0351-0583) and the AM1 system (K: 0.448; C.I. 0.328-0.569), and moderate for the AM1.2 system.

Comparison of the diagnostic accuracy indexes for the AM1.2 system with those for the system by Krick et al. and for the AM1 system has been already presented in the paragraphs above. Further details (including diagnostic accuracy indexes in the worst case scenario and the 95% C.I. of each calculated index), are available in Table 7 and Appendix 3 – Supplementary Table 10.

DISCUSSION

In the current study, conducted according to STARD guidelines,[27] interrater agreement and diagnostic accuracy of the cytological interpretative system currently available in literature for the evaluation of the nodal metastatic status in MCTDs,[104] as well as those of 2 amendments of the latter (AM1 and AM1.2 system) proposed by our research group, were investigated using histopathology as the reference standard. According to the observed results, the AM1.2 system could represent a valid alternative to the simplified 3-classes version of the system proposed by Krick et al.,[104] being characterized by an almost overlapping interrater agreement, and sensibly higher accuracy and sensitivity without substantial changes of the other diagnostic accuracy indexes. Conversely, despite an even higher sensitivity compared to the system by Krick et al., the AM1 system was associated with a sensibly decreased positive predictive value and a reduced interrater agreement. These results might be the consequence of the process applied to create the amended systems. Indeed, the AM1 system was created simply adding criteria regarding quantification of morphologically normal and atypical MCs to those originally included in the system by Krick et al., setting cut-offs for MC quantification through an arbitrary and rough adaptation of the quantification criteria included in the system proposed by Weishaar et al. for the histological evaluation of LNs of MCTDs. Conversely, the AM1.2 system was created setting the optimal cut-off for each quantifiable cytological criterion (i.e. number of aggregates composed by >3 MCs, couples and triplets of MCs, discrete atypical mast cells, and discrete normal MCs)

building ROC curves and setting the histopathological reference standard on a binary outcome (i.e. considering HN0/1 and HN2/3 histological diagnoses as negative and positive for metastasis, respectively). This latter and more rigorous approach allowed a deeper revision of the original system proposed by Krick et al., resulting in an amended system characterized by improved accuracy and sensitivity, still maintaining high reproducibility among different Readers.

To the best of our knowledge, the current work is the first study sistematically investigating reproducibility of cytological criteria for the evaluation of the metastatic status of LNs obtained from MCTDs among different Readers. Indeed, the only former study including the evaluation of this kind of nodal cytological specimens by multiple evaluators did not investigate interrater agreement in details nor reported sufficient raw data for calculation of kappa coefficients.[174] Additionally, the same study by Sabattini et al. focused exclusively on the evaluation of cytological samples obtained from LNs with histopathologically confirmed MCT metastasis, differently from the current work which took into account samples obtained from both histologically non metastatic and metastatic LNs. The results of our study confirmed the hypothesis that the original 5-classes system proposed by Krick et al. might be characterized by low reproducibility among different evaluators, and should thus be abandoned in favour of the simplified 3-classes version of the same system (Appendix 2 – Supplementary Table 1) as proposed by our research group according to observations reported in previous literature.[73,104,145] The evaluators of cytological specimens involved in the current study were specifically choosen trying to stratify as much as possible the group of investigators according to different degrees of experience and expertise in the field of diagnostic cytology. This factor was mirrored by the evaluation of the interrater agreement for each couple of Readers, which frequently revealed a higher interrater agreement among those Readers with more experience in diagnostic cytology (i.e. Reader 3, Reader 4, and Reader 5). Additionally, a reduced agreement between Reader 2 and the other Readers was noted. This observation might be consequence of the fact that Reader 5, being the Diagnostic Leader of the Cytopathological Diagnostic Service of the DIMEVET, trained Readers 1, 3, and 4 during diagnostic cytology rounds on a more or less regular basis. Conversely, Reader 2 was routinely involved in the Hematology

Diagnostic Service of the DIMEVET and not in the Cytopathological Diagnostic Service of the same Department, thus being mostly experienced in the evaluation of blood smears and cytological samples prepared from liquid matrices (e.g. cavitary effusions), and not having received training by Reader 5 on a regular basis. These observations, together with the higher sensitivity and lower specificity (Table 7 and Appendix 3 – Supplementary Table 10) of Reader 2 compared to the other Readers disregarding the intepretative system applied, as well as with the general tendency of the same Reader 2 to report higher number of atypical MCs for each case (data not shown, but available upon request), highlighted that all the cytological intepretative systems investigated might be influenced by the evaluator's experience, as previously reported for cytology in general [42] and specifically for the evaluation of cytological nodal specimens from MCTDs.[174] With this in mind, further studies including a more homogeneous group of cytologists are warranted to validate the reliability of our findings.

To the best of our knowledge, twelve previous studies focused on the investigation of the agreement between cytology and histology in detecting MCT nodal metastasis, or included simultaneous cytological and histological evaluation of LNs obtained from MCTDs.[33,71,73,105,106,108,145,156,174,201,204,240] Among these, only the most recent studies applied cytological criteria proposed by Krick et al.[71,73,105,106,145,156,174,204] In three of the cases a modified version of the cytological interpretative system by Krick et al. was used;[71,145,174] however, changes applied to the Krick's system were not stated. Additionally, 2 studies applied the system by Krick et al. exclusively on a subset of specimens.[71,106] Finally, only 4 studies rigorously and sistematically applied the criteria proposed by Krick et al.[73,105,156,204] All the other studies proposed customized cytological criteria for the evaluation of nodal MCT metastasis (see Introduction – Table 3 for further details).[33,108,201,240] The results of previous studies are compared with those observed in the current study in Table 10. If not explicitally reported, diagnostic accuracy indexes were recalculated whenever possible, extrapolating the number of cases classified as TP, FP, FN, and TN. The study by Cahalane et al.[33] was excluded from the comparison due to complete impossibility to extrapolate data for the investigation of the agreement between cytology and histology. Further details

regarding the recalculation of diagnostic accuracy indexes are reported for each specific compared previous study in Table 10 caption.

Table 10. Summary table of studies previously reported in the literature including cytological and histological evaluation of dog LNs with MCT, and comparison with the results of the current study.

	Langhenbach et al. (2001)[108] *	Worley (2012)[240]	Stefanello et al. (2015)[201]	Ku et al. (2017)[106]	Mutz et al. (2017)[145]	Krick et al. (2017)[105] ***	Sabattini et al. (2018)[174] ****	Fournier et al. (2018)[71]	Sulce et al. (2018)[204] $	Pecceu et al. (2019)[156]	Fournier et al. (2020)[73] $	Current sudy – Krick's system $$$	Current study – AM1 system $$$	Current study – AM1.2 system $$$
Acc	85.71%	50%	100%	n/a	100%	78.57%	100%	84.62%	33.33%	90.00%	57.14%	78.95%	80.45%	80.45%
Se	100%	50%	n/a	67.44%	100%	75.00%	n/a	75%	40.00%	85.19%	40.00%	65.71%	71.43%	70.00%
Sp	n/a	n/a	n/a	n/a	100%	83.33%	n/a	91%	n/a	100%	100%	93.65%	90.48%	92.06%
PPV	85.71%	100%	100%	80.56%	100%	85.71%	100%	86%	66.67%	100%	100%	92.00%	89.29%	90.74%
NPV	n/a	n/a	n/a	n/a	100%	71.43%	n/a	84%	n/a	74.47%	40.00%	71.08%	74.03%	73.42%
HN system	NO	NO	NO	PAR	NO	NO	Y	PAR	Y	Y	Y	Y	Y	Y
Histo metachromatic stains	NO	OCC	OCC	NO	NO	ND	Y	OCC	ND	ND	Y	Y	Y	Y
# cases evaluated	7 (6 TP, 1 FP, 0 FN, 0 TN)	4 (2 TP, 0 FP, 2 FN, 0 TN)	50 (50 TP, 0 FP, 0 FN, 0 TN)	50** (29 TP, 7 FP, 14 FN, TN n/a)	7 (6 TP, 0 FP, 0 FN, 1 TN)	14 (6 TP, 1 FP, 2 FN, 5 TN)	28 (28 TP, 0 FP, 0 FN, 0 TN)	78 (25 TP, 4 FP, 8 FN, 41 TN)	21 (2 TP, 1 FP, 3 FN, 0 TN)	40[$$] (23 TP, 0 FP, 4 FN, 13 TN)	21 (6 TP, 0 FP, 9 FN, 6 TN)	133 (46 TP, 4 FP, 24 FN, 59 TN)	133 (50 TP, 6 FP, 20 FN, 57 TN)	133 (49 TP, 5 FP, 21 FN, 58 TN)

The values reported in the table, when not explicictally reported in the bibliographic sources, were recalculated applying the same statistical calculation software used in the current study after extrapolation from available raw data. Legend: Acc, overall diagnostic accuracy or overall agreement between cytological and histological diagnosis; FN, false negatives; FP, false positives; HN system, application of the histological classification by Weishaar et al.;[234] metachromatic stains, toluidine blue or Giemsa stains; ND, not determined; NPV, negative predictive value; n/a, not applicable due to statistical limitations; OCC, occasionally applied; PAR, application of the histological system by Weishaar et al. exclusively on a subset of cases; PPV, positive predictive value; Se, sensitivity; Sp, specificity; TN, true negative; TP, true positive: Y, yes; * number of dogs and cats with MCT not specified; **at least 50 dogs included (number of false negatives not extractable); ***only cytological samples classified as "certain metastasis" considered as metastatic by the original Authors; ****only HN2/3 histological cases included in the study – only May-Grünwald Giemsa-stained cytological samples taken into account for recalculation of diagnostic accuracy indexes – median values among the 3 cytologists; [$]recalculation of diagnsotic accuracy indexes considering the best case scenario (i.e.

“possible metastasis” considered as negative for metastasis); [$$]data for 1 LN not available; [$$$]median values among the 5 Readers – values referring to the best case scenario (i.e. “possibly metastatic” cytological diagnoses considered as negative for metastasis).

Besides the aforementioned differences in the cytological criteria applied among the different studies, comparison of diagnostic accuracy indexes might have been biased by the high variability in the number of investigated cases as well as by the lack of systematical application of the histological classification system by Weishaar et al. [234] and of metachromatic stains for histological evaluation of LNs. Additionally, it is noteworthy that the reliability of the diagnostic accuracy indexes by Ku et al. is undermined by the lack of perfect matching between cytological and histological samples for some of the investigated LNs. According to the improved diagnostic perfomances showed by the AM1.2 system compared to the AM1 system, only the results of the first will be discussed in comparison to values referring to previously published studies. Nonetheless, the diagnostic accuracy indexes of the AM1 system are included in Table 10 for completeness.

The application of the 3-classes simplified version of the system by Krick et al. and of the AM1.2 system in the current study was characterized by a moderate and high accuracy, respectively, falling between the lower range (33.33%-78.57%) reported for some studies [73,105,204,240] and the upper range (84.62%-100%) reported in the other works available for comparison.[71,108,145,156,174,201]

The sensitivity value of the system by Krick et al. observed in the current study, which was judged as “low”, was lower than that reported in most of the previous studies (6 out of 9) available for comparison.[71,105,106,108,145,156] Similarly, the moderate sensitivity value of the AM1.2 system was lower than that reported in most of the previous studies (5 out of 9) available for comparison.[71,105,108,145,156] The classification of some cases as FN, which in statistical terms are to be considered as responsible for the reduction of the sensitivity value in diagnostic accuracy studies, could be justified primarily by intrinsic characteristics of cytological sampling. Indeed, cytology is based on the evaluation of variably small proportion of the LN, which cannot equal the amount of nodal tissue that is examined with histopathology. In this regards, cytological cases classified as FN might be characterized by a focal distribution of metastasis. This consideration is further strenghtened by the observation that touch imprints and scraping smears, which are generally characterized by an increased cellular yield compared to FNAs,[48,229] were still obtained exclusively from two planar

sections of the LN. Nevertheless, the higher sensitivity value observed for the AM1.2 system may be justified by the inclusion in this interpretative system of quantification of discrete atypical MCs, which conversely is not considered in the system by Krick et al.

The application of both the 3-classes simplified version of the system by Krick et al. and the AM1.2 system was characterized by a very high specificity, falling between the lower range (83.33%-91.00%) values reported for two studies [71,105] and the 100% value reported in the other works available for comparison.[73,145,156] Additionally, it is noteworthy that in 3 studies for which it was not possible to derive the value of specificity, a number of FP cases ranging from 1 to 7 was reported. [106,108,204] It is noteworthy that in the current study, at least 4 cases with a corresponding HN0 histology were diagnosed as “metastatic” by the majority of or by all the Readers. The classification of cases as "false positives", which in statistical terms are to be considered as responsible for the reduction of the specificity value in diagnostic accuracy studies, might be found in more than one explanation.

First of all, the counting of morphologically normal and atypical MCs in 8 HPFs together with excessively low cut-offs for their estimation might have led to an increased number of FP cases. This hypothesis is confirmed by the observation that for each Reader, the AM1 system was characterized by an increased number of FP cases compared to both the system by Krick et al. and the AM1.2 system (Appendix 3 – Supplementary table 8). Indeed, it is highly probable that the arbitrary choice of the cut-offs included in the AM1 system for counting single MCs led to overestimate a number of cases that in histological examination actually turned out to be non-metastatic. Conversely, the probability that the number of counted HPFs might have influenced the number of FP cases is pretty low according to the results showed in Section 1 of the current Thesis, which highlighted that neither the number of counted HPFs nor the counting of MCs over predetermined amounts of total cells influenced the number of MCs observed. This finding further justifies our choice of counting normal and atypical MCs in 8 HPFs rather than over predetermined amounts of total cells, in view of reducing the turnaround time for the application of the investigated cytological interpretative systems. The high discrepancy of FP cases among the

Readers (Appendix 3 – Supplementary table 8) might be further correlated with the fact that each Reader was left free to choose the 8 “hot spot-choosen” HPFs, although guidelines had been provided.

Secondly, previous literature reported that routine histological staining with Hematoxylin-Eosin, which is characterized by a poor dye affinity towards metachromatic granules of MCs, is not always able to ensure a correct quantification of MCs present in the nodal histological sections.[34,108,174] However, in our study this source of bias should be considered practically null since for each histological case both a section stained with Hematoxylin-Eosin and a section stained with Giemsa were evaluated.

Thirdly, as previously suggested, it cannot be excluded that cytology was actually more reliable than histology in the evaluation of micrometastasis that were accidentally found only in the sampled material for cytological examination, thus not being evaluable during the subsequent histological examination.[92] In the same study, Herring et al.[92] reported that the histological examination of LN serial sections may improve sensitivity in detecting micrometastasis. To avoid this, in the last few years all LNs samples submitted to the Diagnostic Pathology Service of the DIMEVET underwent a standardized protocol for trimming and subsequent serial sectioning at regular intervals (1.5 mm).[66] Nonetheless, another study questioned that a similar approach might only arbitrarily allow the detection of individual MCs or arranged in small clusters.[120] Additionally, it is also noteworthy that the criteria proposed by Weishaar et al. might still suffer from a certain degree of subjectivity in the evaluation (e.g. criteria do not report how the detection of numerous couples of MCs should be interpreted),[92,120,234] thus requiring further studies on interobserver agreement in histological evaluation, which unfortunately have not been perfomed up to date.

Finally, FP cases might derive from the contamination of the corresponding cytological slides with metastatic cells remained attached to the scalpel used for scraping smears preparation. Nonetheless, in the current study this hypothesis should be considered unreliable, given that most of the scraping smears were prepared by two of the investigators (Reader 3 and Reader 5) which were aware of this risk and thus payed attention in using a single scalpel for each sampled LN.

The application of both the 3-classes simplified version of the system by Krick et al. and the AM1.2 system was characterized by a very high positive predictive value, falling between the lower range (66-67%-85.71%) values reported for some of the studies previously published [71,105,106,108,204] and the 100% value reported in the other works.[73,145,156,174,201,240]

Finally, the application of both the 3-classes simplified version of the system by Krick et al. and the AM1.2 system was characterized by a moderate negative predictive value, which was similar to that reported for the study by Krick et al.[105] and lower than the 100% value reported for Mutz et al.[145] This latter marked difference might be correlated to the very low number of cases (7) reported in the study by Mutz et al.

As reported in human medicine,[127,136,152,206] the intraoperative evaluation of frozen sections of the SLN might represent a technique which accuracy should be evaluated. Nonetheless, the use of this technique has only rarely been reported in veterinary medicine,[15,16,237] probably due to cost and technical limits associated with the requirement of adequate equipment and trained staff.[152] Sulce et al.[204] proposed the use of flow cytometry on FNA samples for the quantification of nodal MCs in MCT-bearing dogs, with the further potential advantage of combining the evaluation of nodal MC immunophenotype with that of MCs from the matched primary mass. Nonetheless, as stated by the same Authors,[204] further studies are warranted to establish correlation between MC estimates via flow cytometry on one hand and cytological and histopathological findings on the other. In this context, as already highlighted in the introduction, the use of cytology instead of frozen sections and flow cytometry in an intraoperative environment is supported by its low cost, safety, rapidity and feasibility also in first opinion practice,[43,71,106,112,117,145,189] as further testified by the diffuse use of this technique in human medicine.[50,96,241] Specifically, as illustrated in details in the introduction of Chapter 3, intraoperative cytology might be useful in those cases in which the SLN does not coincide with the RLN, or in which LN mapping techniques identified tissue other than LNs (e.g. fat) as the SLN. In the first scenario, a cytological diagnosis of metastatis in the SLN might push the surgeon toward extirpation also of the RLN, according to the previous observation that the SLN might represent a kind of “gate” for MCT metastasis toward other LNs,[84]

and that each metastatic LNs can function as reservoirs for MCs, increasing the total neoplasm burden and worsening the prognosis.[119,159] In the second scenario, cytology might be useful in providing an immediate feedback to the surgeon who could then proceed further with the detection and extirpation of the real SLN or, eventually, of the RLN if a SLN could not be detected. In this context, according to the diagnostic accuracy indexes reported in the current study for each investigated cytological interpretative system, the AM1.2 system might be suitable for intraoperative cytologically evaluation of surgically excised SLN. This observation is correlated with the higher accuracy and sensitivity of this system compared to the 3-classes simplified version of the system by Krick et al., and is further strenghtened by the simplicity of interpretation of the 2 possible diagnostic interpretations (i.e. non metastatic and metastatic). Despite not investigated in details in the current study, it is noteworthy that in spite of markedly decreased values of specificity and positive predictive value, the AM1.1 system was characterized by an even higher sensitivity. According to the current belief that selective lymphadenectomy is always useful and characterized by mild and infrequent controindications,[65,66,119,120,129] the AM1.1 system might be the best system for intraoperative evaluation of surgically extirpated SLNs. Nonetheless, further studies focusing on replicating the results of the current study are warranted before expressing a final recommendation on this topic. Additionally, in their study Sabattini et al.[174] discouraged the use of rapid aqueous stains for the cytological evaluation of the metastatic status in MCT-bearing dogs in favor of May-Grünwald Giemsa or other metachromatic stains. Nonetheless, this aspect might be correlated with an undesirable prolongation of turnaround time of SLN cytological evaluation in an intraoperative context, according to the longer time required for methanolic-based stains.[174] Therefore, further investigations are also warranted to establish faster protocols for metachromatic stains.

The common sense would suggest that the reduction of possible diagnoses from the 3 classes of the simplified version of the system by Krick et al. to the 2 classes of the AM1.2 system reduced tha percentage agreement for those cases previously classified as “possibly metastatic”. Conversely, this hypothesis was denied by the observation that the AM1.2 system was characterized by a sensibly higher agreement with histopathology compared to the 3-classes simplifed version of the

system by Krick et al. for those cases with a corresponding HN1 or HN2/3 histological diagnosis. This observation highlights that quantification of discrete MCs, and specifically of atypical ones, represents a key step in cytological evaluation of LNs obtained from MCTDs. Nonetheless, it is noteworthy that ROC curves gave back a biologically paradoxal optimal cut-off for the quantification of normal MCs in the AM1.2 system. Considered this observation, it cannot be excluded that the sensibly reduced specificity and positive predictive value of the AM1.1 system might be the consequence of the inclusion of quantification of normal MCs among the cytological criteria that should have been evaluated. According to these considerations, further investigations are warranted to establish more precise cut-off for the quantification of normal MCs in cytological LN samples obtained from MCTDs.

To the best of our knowledge, the current study is the first taking into account different cytological sampling techniques in the evaluation of the agreement between cytology and histology in the investigation of LN metastasis in canine MCTs. In this regards, the AM1.2 system showed a sensibly higher agreement with histopathology for all the sampling techniques compared to the 3-classes semplified version of the system by Krick et al. The results of the current study unexpectedly showed that system by Krick et al., although originally designed for the evaluation of FNAs, was generally less effective in achieving the correct diagnosis in these kind of samples compared to touch imprints and scraping smears. Conversely, the high agreement between cytology and histology for touch imprints and scraping smears when the AM1.2 system was applied, might be the consequence of the fact that the cut-offs for the same system were established according to quantifiable parameters evaluated in a caseload which was mostly composed by these kind of specimens. The agreement with histology of the AM1.2 system for FNAs was still considered low. According to previous considerations, this observation might be correlated with the very low number of FNAs (15) included in the current study, further highlighting the need of additional investigations to validate and eventually improve the accuracy of the AM1.2 system in the evaluation of this kind of cytological specimens.

When the concordance between the cytological diagnoses of the cases referring to the same LN but sampled with different techniques was investigated, the AM1.2

system showed a higher number of cases with complete intersampling coherence and a lower number of cases with intersampling incoherence compared to that for the 3-classes simplifed version of the system by Krick et al. The persistency of a certain proportion of discrepances in the cytological diagnosis of different specimens obtained from the same LN might be correlated to the need of further refining cytological criteria on the basis of the sampling technique. Nevertheless, this seems to contrast with the almost overlapping percentage agreement between cytology and histology when the cases were grouped according to the sampling technique. Additionally, it should be remembered that in daily practice, touch imprinting and scraping smearing might represent complementary techniques, compensating each other and providing potential material for the application of immunocytochemical and molecular biology investigations. Another interesting observation is that the AM1.2 system still showed a relatively high proportion of cases with complete intersampling agreement but expressing the wrong diagnosis. Similarly, when the agreement among the diagnoses expressed according to each cytological interpretative system applied was investigated, a considerable proprotion of cases were cytologically diagnosed in the same way, but no agreement was observed with the corresponding histopathological diagnosis. These observations were in line with the significant p value for both the 3-classes simplified version of the system by Krick et al. and the AM1.2 system when McNemar's test was applied. This finding finally confirmed the presence of a test bias, supporting the previous observation that histological examination of LNs from MCTDs cannot always be reliably substituted by cytology.[71] Anyway, the fact that only few LNs (3) were sampled applying all the 3 different techniques might have limited the reliability of the results of the current study in this specific topic. This observation might be worsened by the fact that specimens sampled with different techniques form the same LNs were considered as indipendent cases in the current study.

A major strength of the current study was represented by the application of the STARD guidelines [27] that, in addition to increase the standardization and reproducibility of this work, allowed a quick and clear comparison of obtained results with those reported in previous studies that had applied the same guidelines.[71] However, item 22 (report of clinical interventions occurred between the performance

of the index test and the gold standard test) and item 25 (side effects occurred following the performance of the index test and/or the reference standard test) of STARD guidelines wer not fulfilled in the current study due to lack of specifical clinical information for some of the cases investigated.

Limitations other than the aforementioned ones that characterize the current study were largely due to its retrospective nature. To the Authors' opinion, the lack of statistical analysis aiming to validate the AM1.2 system correlating it with the clinical outcome of investigated dogs, represents the major limit of this work. The choice of excluding this investigation from the current study was mainly due to the relatively low number of included dogs in association with the marked differences in applied therapeutic protocols (e.g. some histological specimens were obtained from RLNs while the other ones from SLNs). These conditions might have introduced too many confounding effects in the statistical analysis, biasing the results in an unpredictable and unclear way. Nonetheless, additional samples are currently under collection, allowing further investigations in this direction in the future.

Another limitation of the current study was the lack of analysis aiming to investigate the presence of correlations between nodal metastatic status and the grade of the corresponding primary MCT.[5,196,201] In the last years, an increasing number of histologically low-grade but still biologically aggressive MCTs has been described in the literature.[5,103,196,201] Additionally, it cannot be excluded that cytologically and /or histologically well-differentiated, low grade tumors might be associated with the presence of nodal atypical MCs, representing neoplastic cells that underwent premature mutations associated with a more aggressive biological behavior. According to all these considerations, further and specific studies focusing on this topic are warranted.

Several sources of selection bias strictly correlated with the retrospective nature of the current study, might have undermined the reliability of our results. In this context, it should be remembered that the rationale underlying the setting of the inclusion criteria for our study was to guarantee a perfect matching between cytological and histological samples referring to each LN investigated. Firstly, the retrospective nature of our work imposed a sample size equal to the number of cases in the

archives that met the inclusion criteria, making inapplicable any test of statistical power to determine the number of samples needed to improve reliability of findings. However, it should be noted that, to the best of our knowledge, the current study investigated a higher number of cases compared to other works previously published. Secondly, the prevalence of LN metastasis reported in the current study is noticeable and higher than that available among the works that included the largest number of investigated cases.[71,201] Nevertheless, our value of metastasis prevalence was really similar to that reported in previous study focused on investigation on SLNs,[66,73,240] probably due to the fact that the majority of investigated LNs in the current work were SLNs. Thirdly, the fact that most patients (43 out of 47 dogs) were treated at the DIMEVET University Hospital (OVU) may have affected the severity of clinical presentations. Nonetheless, the Small Animal Surgery and Diagnostic Imaging Services of the DIMEVET represent a reference center for the whole North-West of Italy in the treatment of canine MCT, attracting also numerous dogs affected by de novo MCT presentation. This situation might have also positively influenced the standardization of LN sampling for cytological and histopathological investigations, which was performed by a small group of highly specialized university personnel. Finally, only one cytological case among those evaluated in this study was obtained from the necropsy of a dog dead due to MCT systemic spreading.

The choices applied to allow the calculation of diagnostic accuracy indexes might have further biased the results of the current study. Specifically, the choice of considering HN1 diagnoses as negative for metastasis might have strongly influenced the diagnostic accuracy indexes reported in the current study, despite perfectly consistent with previous literature.[65,73,84,120,156,204,234] Nonetheless, according to the fact that the biological role of LNs classified as HN1 is still unclear,[65,196,234] and in view of future progresses in the knowledge regarding this topic, all raw data of the current study including the cytological diagnoses expressed by each Reader and the corresponding histological diagnosis are available in Appendix 3 – Supplementary Table 2. Similarly, diagnostic accuracy indexes calculated for each investigated cytological interpretative system and including the values for both the best and the worst case scenario (i.e. considering “possibly metastatic” cytological cases as

negative and positive for metastasis, respectively) are available for comaprison with eventual future studies (provided in Appendix 3 – Supplementary Table 10).

In conclusion, the current study conducted according to STARD guidelines,[27] investigated interrater agreement and diagnostic accuracy of the cytological interpretative system currently available in literature for the evaluation of the nodal metastatic status in MCTDs,[104] as well as those of 2 amendments of the latter (AM1 and AM1.2 system) proposed by our research group, using histopathology as the reference standard. According to the observed results, the AM1.2 system could represent a valid alternative to the simplified 3-classes version of the system proposed by Krick et al., being characterized by an almost overlapping interrater agreement, and sensibly higher accuracy and sensitivity without substantial changes of the other diagnostic accuracy indexes. These observations further suggest the use of the AM1.2 system also in an intraoperative environment for the real-time evaluation of surgically-extirpated SLNs. On the other hand, the relatively low number of dogs included in the study, the different techniques applied for LN cytological sampling in association with the low number of included FNAs, the uni-institutional nature of our study, and the lack of investigation of prospective correlations with clinical outcome, push toward considering the results of the current work as preliminary. Therefore, other investigators are kindly invited to replicate this study to verify the reliability of our results, eventually providing more precise cut-offs for MC quantification and additional cytological criteria better fitting with the histopathological diagnosis and the clinical outcome.

REFERENCES

1. Adams VJ, Evans KM, Sampson J, Wood JLN. Methods and mortality results of a health survey of purebred dogs in the UK. *J Small Anim Pract*. 2010;51:512–524.
2. Allen P, Paul A. Gastropexy for Prevention of Gastric Dilatation-Volvulus in Dogs: History and Techniques. *Top Companion Anim Med*. 2014;29:77–80.
3. Amores-Fuster I, Cripps P, Graham P, Marrington AM, Blackwood L. The diagnostic utility of lymph node cytology samples in dogs and cats. *J Small Anim Pract*. 2015;56:125–129.
4. Aronsohn MG, Dubiel B, Roberts B, Powers BE. Prognosis for acute nontraumatic hemoperitoneum in the dog: a retrospective analysis of 60 cases (2003-2006). *J Am Anim Hosp Assoc*. 2009;45:72–77.
5. Bae S, Milovancev M, Bartels C, et al. Histologically low-grade, yet biologically high-grade, canine cutaneous mast cell tumours: A systematic review and meta-analysis of individual participant data. *Vet Comp Oncol*. 2020 Dec;18:580–589.
6. Baginski H, Davis G, Bastian RP. The Prognostic Value of Lymph Node Metastasis with Grade 2 MCTs in Dogs: 55 Cases (2001–2010). *J Am Anim Hosp Assoc*. 2014;50:89–95.
7. Bahr KL, Sharkey LC, Murakami T, Feeney DA. Accuracy of US-Guided FNA of Focal Liver Lesions in Dogs: 140 Cases (2005–2008). *J Am Anim Hosp Assoc*. 2013 May 1;49:190–196.
8. Bairdain S, Litman HJ, Troy M, et al. Twenty-years of splenic preservation at a level 1 pediatric trauma center. *J Pediatr Surg*. 2015 May;50:864–868.
9. Ballegeer EA, Forrest LJ, Dickinson RM, Schutten MM, Delaney FA, Young KM. Correlation of ultrasonographic appearance of lesions and cytologic and histologic diagnoses in splenic aspirates from dogs and cats: 32 cases (2002–2005). *J Am Vet Med Assoc*. 2007 Mar 1;230:690–696.
10. Balogh L, Thuróczy J, Andócs G, et al. Sentinel lymph node detection in canine oncological patients. *Nucl Med Rev*. 2002;5:139–144.
11. Banzato T, Gelain ME, Aresu L, Centelleghe C, Benali SL, Zotti A. Quantitative analysis of ultrasonographic images and cytology in relation to histopathology of canine and feline liver: An ex-vivo study. *Res Vet Sci*. 2015;103:164–169.
12. Bauer, CL; Rakich, PM; Latimer, KS; White, A; Hensel P. Assessment and comparison of reference intervals for mast cells in peripheral lymph nodes of normal dogs and dogs with allergic skin disease. *26th Annu Meet AAVD/ACVD, Galveston, TX Vet Dermatol*. 2011;22:2–4.
13. Beer P, Pozzi A, Rohrer Bley C, Bacon N, Pfammatter NS, Venzin C. The role of sentinel lymph node mapping in small animal veterinary medicine: A comparison with current approaches in human medicine. *Vet Comp Oncol*. 2018;16:178–187.
14. Bernardi S, Martini V, Perfetto S, Cozzi M, Comazzi S. Flow Cytometric Analysis of Mediastinal Masses in Cats: A Retrospective Study. *Front Vet Sci*. 2020;7:444.
15. Bernstein JA, Hodgin EC, Holloway HW, Hedlund CS, Storey ES, Hubert JD. Mohs micrographic surgery: a technique for total margin assessment in veterinary cutaneous oncologic surgery. *Vet Comp Oncol*. 2006 Sep;4:151–160.
16. Bernstein JA, Storey ES, Bauer RW. Moh's micrographic surgery for the management of a periocular mast cell tumor in a dog. *Vet Ophthalmol*. 2013 May;16:234–239.
17. Bertazzolo W, Dell'Orco M, Bonfanti U, et al. Canine angiosarcoma: cytologic, histologic, and immunohistochemical correlations. *Vet Clin Pathol*. 2005;34:28–34.
18. Bertazzolo W, Didier M, Gelain ME, et al. Accuracy of cytology in distinguishing adrenocortical tumors from pheochromocytoma in companion animals. *Vet Clin Pathol*. 2014 Sep;43:453–459.
19. Berzina I, Sharkey LC, Matise I, Kramek B. Correlation between cytologic and histopathologic diagnoses of bone lesions in dogs: a study of the diagnostic accuracy of bone cytology. *Vet Clin Pathol*. 2008;37:332–338.
20. Beserra HEO, Grandi F, Dufloth RM, et al. Metastasis of Mammary Carcinoma in Bitches: Evaluation of the Sentinel Lymph Node Technique. *Adv Breast Cancer Res*. 2016;05:58–65.
21. Biller B, Berg J, Garrett L, et al. 2016 AAHA oncology guidelines for dogs and cats. *J Am Anim Hosp Assoc*. 2016;52:181–204.
22. Blackwood L, Murphy S, Buracco P, et al. European consensus document on mast cell tumours in dogs and cats. *Vet Comp Oncol*. 2012;10:1–29.
23. Blackwood L. Cats with Cancer: Where to start. *J Feline Med Surg*. 2013 Apr 19;15:366–377.

24. Bonfanti U, Bertazzolo W, Gracis M, et al. Diagnostic value of cytological analysis of tumours and tumour-like lesions of the oral cavity in dogs and cats: A prospective study on 114 cases. *Vet J*. 2015;205:322–327.
25. Bookbinder PF, Butt MT, Harvey HJ. Determination of the number of mast cells in lymph node, bone marrow, and buffy coat cytologic specimens from dogs. *J Am Vet Med Assoc*. 1992 Jun;200:1648–1650.
26. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Ann Clin Biochem*. 2003 Feb;40:357–363.
27. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015 Oct 28;351:h5527.
28. Bottero E, Melega M, Dimartino ER, et al. Diagnosis of feline mesenchymal nasal hamartoma by squash preparation cytology. *Vet Clin Pathol*. 2018 Dec 1;47:629–633.
29. Bourgeois P, Nogaret JM, Veys I, et al. How “hot” is the pathologically positive sentinel lymph node in breast cancer patients? *Nucl Med Commun*. 2003 May;24:513–518.
30. Braun AO, Hauser B. Korrelation zwischen zytologischen und histologischen Haut-, Lymphknoten- Und Milzbefunden bei 500 Hunden und Katzen. *Schweiz Arch Tierheilkd*. 2007 Jun;149:249–257.
31. Brissot HN, Edery EG. Use of indirect lymphography to identify sentinel lymph node in dogs: a pilot study in 30 tumours. *Vet Comp Oncol*. 2017;15:740–753.
32. Burkhard MJ, Bienzle D. Making Sense of Lymphoma Diagnostics in Small Animal Patients. *Clin Lab Med*. 2015;35:591–607.
33. Cahalane AK, Payne S, Barber LG, et al. Prognostic factors for survival of dogs with inguinal and perineal mast cell tumors treated surgically with or without adjunctive treatment: 68 cases (1994-2002). *J Am Vet Med Assoc*. 2004 Aug;225:401–408.
34. Camus MS, Priest HL, Koehler JW, et al. Cytologic Criteria for Mast Cell Tumor Grading in Dogs With Evaluation of Clinical Outcome. *Vet Pathol*. 2016;53:1117–1123.
35. Caniatti M, Roccabianca P, Ghisleni G, Mortellaro CM, Romussi S, Mandelli G. Evaluation of brush cytology in the diagnosis of chronic intranasal disease in cats. *J Small Anim Pract*. 1998 Feb;39:73–77.
36. Caniatti M, Cunha NP da, Avallone G, et al. Diagnostic accuracy of brush cytology in canine chronic intranasal disease. *Vet Clin Pathol*. 2012;41:133–140.
37. Carmon M, Olsha O, Schecter WP, et al. The “Sentinel Chain”: a new concept for prediction of axillary node status in breast cancer patients. *Breast Cancer Res Treat*. 2006 Jun;97:323–328.
38. Chaffin K, Thrall DE. Results of radiation therapy in 19 dogs with cutaneous mast cell tumor and regional lymph node metastasis. *Vet Radiol ultrasound Off J Am Coll Vet Radiol Int Vet Radiol Assoc*. 2002;43:392–395.
39. Chino J, Fujino Y, Kobayashi T, et al. Cytomorphological and immunological classification of feline lymphomas: clinicopathological features of 76 cases. *J Vet Med Sci*. 2013;75:701–707.
40. Christensen NI, Canfield PJ, Martin PA, Krockenberger MB, Spielman DS, Bosward KL. Cytopathological and histopathological diagnosis of canine splenic disorders. *Aust Vet J*. 2009 May;87:175–181.
41. Christopher MM. Cytology of the spleen. *Vet Clin North Am Small Anim Pract*. 2003 Jan;33:135–152.
42. Christopher MM, Hotz CS. Cytologic diagnosis: expression of probability by clinical pathologists. *Vet Clin Pathol*. 2004;33:84–95.
43. Christopher MM, Hotz CS, Shelly SM, Pion PD. Use of cytology as a diagnostic method in veterinary practice and assessment of communication between veterinary practitioners and veterinary clinical pathologists. *J Am Vet Med Assoc*. 2008 Mar;232:747–754.
44. Cleveland MJ, Casale S. Incidence of malignancy and outcomes for dogs undergoing splenectomy for incidentally detected nonruptured splenic nodules or masses: 105 cases (2009–2013). *J Am Vet Med Assoc*. 2016 Jun;248:1267–1273.
45. Clifford CA, Pretorius ES, Weisse C, et al. Magnetic resonance imaging of focal splenic and hepatic lesions in the dog. *J Vet Intern Med*. 2004;18:330–338.
46. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016 Nov;6:e012799.
47. Cohen JF, Korevaar DA, Gatsonis CA, et al. STARD for Abstracts: essential items for reporting diagnostic accuracy studies in journal or conference abstracts. 2017 Aug 17;358:j3751.
48. Cohen M, Bohling MW, Wright JC, Welles EA, Spano JS. Evaluation of sensitivity and

specificity of cytologic examination: 269 cases (1999-2000). *J Am Vet Med Assoc*. 2003;222:964–967.
49. Davies O, Taylor AJ. Refining the “double two-thirds” rule: Genotype-based breed grouping and clinical presentation help predict the diagnosis of canine splenic mass lesions in 288 dogs. *Vet Comp Oncol*. 2020 Dec;18:548–558.
50. Delgado-Bocanegra R, Millen E, Nascimento C, Bruno K. Intraoperative imprint cytology versus histological diagnosis for the detection of sentinel lymph nodes in breast cancer treated with neoadjuvant chemotherapy. *Clinics*. 2018;73:1–6.
51. Dettmeyer R. Lethal Infections, Sepsis, and Shock. In: *Forensic histopathology: fundamentals and perspectives*. Heidelberg: Springer; 2011:303–331.
52. Dey P. Time for evidence-based cytology. *Cytojournal*. 2007 Jan 8;4:1.
53. Dobson JM. Breed-predispositions to cancer in pedigree dogs. *ISRN Vet Sci*. 2013;2013:941275.
54. Dogan NU, Dogan S, Favero G, Köhler C, Dursun P. The Basics of Sentinel Lymph Node Biopsy: Anatomical and Pathophysiological Considerations and Clinical Aspects. Adrian TE, ed. *J Oncol*. 2019;2019:3415630.
55. Dolka I, Czopowicz M, Gruk-Jurka A, Wojtkowska A, Sapierzyński R, Jurka P. Diagnostic efficacy of smear cytology and Robinson’s cytological grading of canine mammary tumors with respect to histopathology, cytomorphometry, metastases and overall survival. *PLoS One*. 2018;13:e0191595.
56. Donnelly L, Mullin C, Balko J, et al. Evaluation of histological grade and histologically tumour-free margins as predictors of local recurrence in completely excised canine mast cell tumours. *Vet Comp Oncol*. 2015;13:70–76.
57. Eberle N, von Babo V, Nolte I, Baumgärtner W, Betz D. Splenic masses in dogs. Part 1: Epidemiologic, clinical characteristics as well as histopathologic diagnosis in 249 cases (2000-2011). *Tierarztl Prax Ausg K Kleintiere Heimtiere*. 2012;40:250–260.
58. Eich C, Whitehair J, Moroff S, Heeb L. The accuracy of intraoperative cytopathological diagnosis compared with conventional histopathological diagnosis. *J Am Anim Hosp Assoc*. 2000 Jan;36:16–18.
59. Elliott JW, Cripps P, Blackwood L, Berlato D, Murphy S, Grant IA. Canine oral mucosal mast cell tumours. *Vet Comp Oncol*. 2016 Mar;14:101–111.
60. Elston LB, Sueiro FAR, Cavalcanti JN, Metze K. The importance of the mitotic index as a prognostic factor for survival of canine cutaneous mast cell tumors: a validation study. Vol. 46, *Veterinary pathology*. United States; 2009:362–364, author reply 364-5.
61. Emanuelli MP, Kommers GD, Antoniazzi AQ, Bernardes FCS, Lopes STA, Fighera RA. Myoepithelial cells and extracellular matrix in the cytologic differentiation of canine mammary tumors. *Vet Clin Pathol*. 2020 Sep;49:451–458.
62. Fabrizio F, Calam AE, Dobson JM, et al. Feline mediastinal lymphoma: a retrospective study of signalment, retroviral status, response to chemotherapy and prognostic indicators. *J Feline Med Surg*. 2014;16:637–644.
63. Favril S, Stock E, Hernot S, et al. Sentinel lymph node mapping by near-infrared fluorescence imaging and contrast-enhanced ultrasound in healthy dogs. *Vet Comp Oncol*. 2019 Mar;17:89–98.
64. Fernandez S, Lang JM, Maritato KC. Evaluation of Nodular Splenic Lesions in 370 Small-Breed Dogs (<15 kg). *J Am Anim Hosp Assoc*. 2019;55:201–209.
65. Ferrari R, Marconato L, Buracco P, et al. The impact of extirpation of non-palpable/normal-sized regional lymph nodes on staging of canine cutaneous mast cell tumours: A multicentric retrospective study. *Vet Comp Oncol*. 2018;16:505–510.
66. Ferrari R, Chiti LE, Manfredi M, et al. Biopsy of sentinel lymph nodes after injection of methylene blue and lymphoscintigraphic guidance in 30 dogs with mast cell tumors. *Vet Surg*. 2020 Aug 6;
67. Finora K, Leibman NF, Fettman MJ, Powers BE, Hackett TA, Withrow SJ. Cytological comparison of fine-needle aspirates of liver and spleen of normal dogs and of dogs with cutaneous mast cell tumours and an ultrasonographically normal appearing liver and spleen. *Vet Comp Oncol*. 2006;4:178–183.
68. Finotello R, Vasconi ME, Sabattini S, et al. Feline large granular lymphocyte lymphoma: An Italian Society of Veterinary Oncology (SIONCOV) retrospective study. *Vet Comp Oncol*. 2018 Mar;16:159–166.

69. Fossum TW, Caplan ER. Surgery of the Hemolymphatic System. In: Fossum TW, ed. *Small animal surgery*. St. Louis, Mo.: Elsevier Mosby; 2013:685–704.
70. Fournel-Fleury C, Magnol JP, Bricaire P, et al. Cytohistological and immunological classification of canine malignant lymphomas: comparison with human non-Hodgkin's lymphomas. *J Comp Pathol*. 1997;117:35–59.
71. Fournier Q, Cazzini P, Bavcar S, Pecceu E, Ballber C, Elders R. Investigation of the utility of lymph node fine-needle aspiration cytology for the staging of malignant solid tumors in dogs. *Vet Clin Pathol*. 2018;47:489–500.
72. Fournier Q, Cazzini P, Elders R, Del Pozo J. Correlation of lymph node cytology with 1-year metastasis-free survival in dogs with mast cell tumours after lymphadenectomy. In: *ESVONC Annual Congress 2018*. 2018:
73. Fournier Q, Thierry F, Longo M, et al. Contrast-enhanced ultrasound for sentinel lymph node mapping in the routine staging of canine mast cell tumours: A feasibility study. *Vet Comp Oncol*. 2020 Aug 24;n/a.
74. Gabor LJ, Malik R, Canfield PJ. Clinical and anatomical features of lymphosarcoma in 118 cats. *Aust Vet J*. 1998 Nov 1;76:725–732.
75. Gelb HR, Freeman LJ, Rohleder JJ, Snyder PW. Feasibility of contrast-enhanced ultrasound-guided biopsy of sentinel lymph nodes in dogs. *Vet Radiol Ultrasound*. 2010;51:628–633.
76. Ghisleni G, Roccabianca P, Ceruti R, et al. Correlation between fine-needle aspiration cytology and histopathology in the evaluation of cutaneous and subcutaneous masses from dogs and cats. *Vet Clin Pathol*. 2006 Mar;35:24–30.
77. Gieger TL, Théon AP, Werner JA, McEntee MC, Rassnick KM, DeCock HE V. Biologic behavior and prognostic factors for mast cell tumors of the canine muzzle: 24 cases (1990-2001). *J Vet Intern Med*. 2003;17:687–692.
78. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. 2014 Jan;383:267–276.
79. Gochhait D, Dey P, Rajwanshi A, et al. Role of Fine needle aspiration cytology of spleen. *APMIS acta Pathol Microbiol Immunol Scand*. 2015 Mar;123:190–193.
80. Goldfarb LR, Alazraki NP, Eshima D, Eshima LA, Herda SC, Halkar RK. Lymphoscintigraphic identification of sentinel lymph nodes: clinical evaluation of 0.22-micron filtration of Tc-99m sulfur colloid. *Radiology*. 1998 Aug;208:505–509.
81. Gomez-Macías GS, Garza-Guajardo R, Segura-Luna J, Barboza-Quintana O. Inadequate fine needle aspiration biopsy samples: pathologists versus other specialists. *Cytojournal*. 2009 Jun 18;6:9.
82. Gómez-Rubio M, López-Cano A, Rendón P, et al. Safety and diagnostic accuracy of percutaneous ultrasound-guided biopsy of the spleen: A multicenter study. *J Clin Ultrasound*. 2009 Oct;37:445–450.
83. Grimes JA, Matz BM, Christopherson PW, et al. Agreement Between Cytology and Histopathology for Regional Lymph Node Metastasis in Dogs With Melanocytic Neoplasms. *Vet Pathol*. 2017 Jul;54:579–587.
84. Grimes JA, Secrest SA, Wallace ML, Laver T, Schmiedt CW. Use of indirect computed tomography lymphangiography to determine metastatic status of sentinel lymph nodes in dogs with a pre-operative diagnosis of melanoma or mast cell tumour. *Vet Comp Oncol*. 2020 Dec;18:818–824.
85. Guzera M, Cian F, Leo C, Winnicka A, Archer J. The use of flow cytometry for immunophenotyping lymphoproliferative disorders in cats: a retrospective study of 19 cases. *Vet Comp Oncol*. 2016 Aug 1;14:40–51.
86. Hammond TN, Pesillo-Crosby SA. Prevalence of hemangiosarcoma in anemic dogs with a splenic mass and hemoperitoneum requiring a transfusion: 71 cases (2003-2005). *J Am Vet Med Assoc*. 2008 Feb;232:553–558.
87. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982 Apr;143:29–36.
88. Hayes A. Feline lymphoma 2. Specific disease presentations. *In Pract*. 2006 Nov 1;28:578 LP – 585.
89. Hendrick MJ. Mesenchymal Tumors of the Skin and Soft Tissues. In: Meuten DJ, ed. *Tumors in domestic animals*. Ames, Iowa: Wiley/Blackwell; 2017:142–175.
90. Hergt F, von Bomhard W, Kent MS, Hirschberger J. Use of a 2-tier histologic grading system for canine cutaneous mast cell tumors on cytology specimens. *Vet Clin Pathol*. 2016;45:477–

483.
91. Herman EJ, Stern AW, Fox RJ, Dark MJ. Understanding the Efficiency of Splenic Hemangiosarcoma Diagnosis Using Monte Carlo Simulations. *Vet Pathol*. 2019 Nov;56:856–859.
92. Herring ES, Smith MM, Robertson JL. Lymph node staging of oral and maxillofacial neoplasms in 31 dogs and cats. *J Vet Dent*. 2002 Sep;19:122–126.
93. Hillman LA, Garrett LD, de Lorimier L-P, Charney SC, Borst LB, Fan TM. Biological behavior of oral and perioral mast cell tumors in dogs: 44 cases (1996-2006). *J Am Vet Med Assoc*. 2010 Oct;237:936–942.
94. Horta RS, Lavalle GE, Monteiro LN, Souza MCC, Cassali GD, Araújo RB. Assessment of Canine Mast Cell Tumor Mortality Risk Based on Clinical, Histologic, Immunohistochemical, and Molecular Features. *Vet Pathol*. 2018;55:212–223.
95. Hume CT, Kiupel M, Rigatti L, Shofer FS, Skorupski KA, Sorenmo KU. Outcomes of dogs with grade 3 mast cell tumors: 43 cases (1997-2007). *J Am Anim Hosp Assoc*. 2011;47:37–44.
96. Jeruss JS, Hunt KK, Xing Y, et al. Is intraoperative touch imprint cytology of sentinel lymph nodes in patients with breast cancer cost effective? *Cancer*. 2006;107:2328–2336.
97. Johnson KA, Powers BE, Withrow SJ, Sheetz MJ, Curtis CR, Wrigley RH. Splenomegaly in Dogs. *J Vet Intern Med*. 1989 Jul 1;3:160–166.
98. Kaltreider HB, Barth E, Pellegrini C. The effect of splenectomy on the appearance of specific antibody-forming cells in lungs of dogs after intravenous immunization with sheep erythrocytes. *Exp Lung Res*. 1981;2:231–238.
99. Karaman S, Detmar M. Mechanisms of lymphatic metastasis. *J Clin Invest*. 2014/03/03. 2014 Mar;124:922–928.
100. Kawamura M, Maede Y, Namioka S. Mitogenic responsibilities of lymphocytes in canine babesiosis and the effects of splenectomy on it. *Jpn J Vet Res*. 1987;35:1–10.
101. Kemming G, Messick JB, Mueller W, et al. Can we continue research in splenectomized dogs? Mycoplasma haemocanis: Old problem - New insight. *Eur Surg Res*. 2004;36:198–205.
102. Kiupel M. Mast Cell Tumors. In: Meuten DJ, ed. *Tumors in domestic animals*. Ames, Iowa: Wiley/Blackwell; 2017:176–202.
103. Kiupel M, Webster JD, Bailey KL, et al. Proposal of a 2-tier histologic grading system for canine cutaneous mast cell tumors to more accurately predict biological behavior. *Vet Pathol*. 2011;48:147–155.
104. Krick EL, Billings AP, Shofer FS, Watanabe S, Sorenmo KU. Cytological lymph node evaluation in dogs with mast cell tumours: Association with grade and survival. *Vet Comp Oncol*. 2009;7:130–138.
105. Krick EL, Kiupel M, Durham AC, Thaiwong T, Brown DC, Sorenmo KU. Investigating Associations Between Proliferation Indices, C-kit, and Lymph Node Stage in Canine Mast Cell Tumors. *J Am Anim Hosp Assoc*. 2017;53:258–264.
106. Ku CK, Kass PH, Christopher MM. Cytologic–histologic concordance in the diagnosis of neoplasia in canine and feline lymph nodes: a retrospective study of 367 cases. *Vet Comp Oncol*. 2017;15:1206–1217.
107. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977 May 27;33:159–174.
108. Langenbach A, McManus PM, Hendrick MJ, Shofer FS, Sorenmo KU. Sensitivity and specificity of methods of assessing the regional lymph nodes for evidence of metastasis in dogs and cats with solid tumors. *J Am Vet Med Assoc*. 2001 May;218:1424–1428.
109. Lee M, Park J, Choi H, Lee H, Jeong SM. Presurgical assessment of splenic tumors in dogs: a retrospective study of 57 cases (2012–2017). *J Vet Sci*. 2018 Nov;19:827–834.
110. Lester SJ, Hume JB, Phipps B. Haemobartonella canis infection following splenectomy and transfusion. *Can Vet J*. 1995;36:444–445.
111. Liu L-C, Lang JE, Jenkins T, et al. Is It Necessary to Harvest Additional Lymph Nodes after Resection of the Most Radioactive Sentinel Lymph Node in Breast Cancer? *J Am Coll Surg*. 2008;207:853–858.
112. London CA, Thamm DH. Mast Cell Tumors. In: Withrow SJ, Vail DM, Page RL, eds. *Small Animal Clinical Oncology*. St. Louis: Elsevier Saunders; 2013:335–355.
113. Louwerens M, London CA, Pedersen NC, Lyons LA. Feline Lymphoma in the Post—Feline Leukemia Virus Era. *J Vet Intern Med*. 2005 May 1;19:329–335.
114. Lurie DM, Seguin B, Schneider PD, Verstraete FJ, Wisner ER. Contrast-assisted ultrasound for

sentinel lymph node detection in spontaneously arising canine head and neck tumors. *Invest Radiol*. 2006;41:415–421.
115. Maki LC, Males KN, Byrnes MJ, El-Saad AA, Coronado GS. Incidence of gastric dilatation-volvulus following a splenectomy in 238 dogs. *Can Vet J*. 2017;58:1275–1280.
116. Marconato L, Bettini G, Giacoboni C, et al. Clinicopathological features and outcome for dogs with mast cell tumors and bone marrow involvement. *J Vet Intern Med*. 2008;22:1001–1007.
117. Marconato L, Marchetti V, Francione D, et al. Morphometrical approach for predicting regional lymph node micrometastatic load in canine mast cell tumours: preliminary results. *Vet Comp Oncol*. 2008 Sep;6:162–170.
118. Marconato L. The staging and treatment of multicentric high-grade lymphoma in dogs: A review of recent developments and future prospects. *Vet J*. 2011;188:34–38.
119. Marconato L, Polton G, Stefanello D, et al. Therapeutic impact of regional lymphadenectomy in canine stage II cutaneous mast cell tumours. *Vet Comp Oncol*. 2018;16:580–589.
120. Marconato L, Stefanello D, Kiupel M, et al. Adjuvant medical therapy provides no therapeutic benefit in the treatment of dogs with low-grade mast cell tumours and early nodal metastasis undergoing surgery. *Vet Comp Oncol*. 2020 Sep 1;18:409–415.
121. Marino DJ, Matthiesen DT, Fox PR, Lesser MB, Stamoulis ME. Ventricular Arrhythmias in Dogs Undergoing Splenectomy : *Vet Surg*. 1994;23:101–106.
122. Martini V, Bernardi S, Marelli P, Cozzi M, Comazzi S. Flow cytometry for feline lymphoma: a retrospective study regarding pre-analytical factors possibly affecting the quality of samples. *J Feline Med Surg*. 2018 Jun;20:494–501.
123. Masserdotti C. Architectural patterns in cytology: correlation with histology. *Vet Clin Pathol*. 2006 Dec;35:388–396.
124. McAloney CA, Sharkey LC, Feeney DA, Seelig DM. Diagnostic utility of renal fine-needle aspirate cytology and ultrasound in the cat. *J Feline Med Surg*. 2017 Jul 18;20:544–553.
125. McAloney CA, Sharkey LC, Feeney DA, Seelig DM, Avery AC, Jessen CR. Evaluation of the diagnostic utility of cytologic examination of renal fine-needle aspirates from dogs and the use of ultrasonographic features to inform cytologic diagnosis. *J Am Vet Med Assoc*. 2018 May;252:1247–1256.
126. McHugh ML. Interrater reliability: the kappa statistic. *Biochem medica*. 2012;22:276–282.
127. Melkane AE, Mamelle G, Wycisk G, et al. Sentinel node biopsy in early oral squamous cell carcinomas: a 10-year experience. *Laryngoscope*. 2012 Aug;122:1782–1788.
128. Ménard M, Fontaine M, Morin M. Fine needle aspiration biopsy of malignant tumors in dogs and cats: a report of 102 cases. *Can Vet J = La Rev Vet Can*. 1986 Dec;27:504–510.
129. Mendez SE, Drobatz KJ, Duda LE, White P, Kubicek L, Sorenmo KU. Treating the locoregional lymph nodes with radiation and/or surgery significantly improves outcome in dogs with high-grade mast cell tumours. *Vet Comp Oncol*. 2020 Jun;18:239–246.
130. Messick JB. The Lymph Nodes. In: Valenciano AC, Cowell R, eds. *Cowell and Tyler's Diagnostic Cytology and Hematology of the Dog and Cat*. St. Louis, Mo.: Elsevier Mosby; 2016:180–194.
131. Michels GM, Knapp DW, DeNicola DB, Glickman N, Bonney P. Canine Cutaneous Mast Cell Tumors With Histopathologically Tumor-Free Versus Nontumor-Free Margins : A Retrospective Study of 31 Cases. *J Am Anim Hosp Assoc*. 2002;38:458–466.
132. Miko I, Nemeth N, Peto K, Furka A, Toth L, Furka I. Changes of red blood cell aggregation parameters in a long-term follow-up of splenectomy, spleen-autotransplantation and partial or subtotal spleen resections in a canine model. *Clin Hemorheol Microcirc*. 2017;67:91–100.
133. Miko I, Nemeth N, Sogor V, et al. Comparative erythrocyte deformability investigations by filtrometry, slit-flow and rotational ektacytometry in a long-term follow-up animal study on splenectomy and different spleen preserving operative techniques: Partial or subtotal spleen resection and. *Clin Hemorheol Microcirc*. 2017;66:83–96.
134. Millanta F, Asproni P, Aquino G, Poli A. Cytologic grading of canine and feline spindle-cell sarcomas of soft tissues and its correlation with histologic grading. *Top Companion Anim Med*. 2020 Nov;41:100458.
135. Miller RL, Van Lelyveld S, Warland J, Dobson JM, Foale RD. A retrospective review of treatment and response of high-risk mast cell tumours in dogs. *Vet Comp Oncol*. 2016 Dec;14:361–370.
136. Moo T-A, Edelweiss M, Hajiyeva S, et al. Is Low-Volume Disease in the Sentinel Node After Neoadjuvant Chemotherapy an Indication for Axillary Dissection? *Ann Surg Oncol*. 2018/03/23.

2018 Jun;25:1488–1494.
137. Moore AS, Frimberger AE, Sullivan N, Moore PF. Histologic and Immunohistochemical Review of Splenic Fibrohistiocytic Nodules in Dogs. *J Vet Intern Med*. 2012;26:1164–1168.
138. Moore A. Extranodal Lymphoma in the Cat: Prognostic factors and treatment options. *J Feline Med Surg*. 2013 Apr 19;15:379–390.
139. Moore PF. Canine and Feline Histiocytic Diseases. In: Meuten DJ, ed. *Tumors in domestic animals*. Ames, Iowa: Wiley/Blackwell; 2017:322–336.
140. Moore PF. A Review of Histiocytic Diseases of Dogs and Cats. *Vet Pathol*. 2014;51:167–184.
141. Moore PF, Rodriguez-Bertos A, Kass PH. Feline gastrointestinal lymphoma: mucosal architecture, immunophenotype, and molecular clonality. *Vet Pathol*. 2012 Jul;49:658–668.
142. Mullins MN, Dernell WS, Withrow SJ, Ehrhart EJ, Thamm DH, Lana SE. Evaluation of prognostic factors associated with outcome in dogs with multiple cutaneous mast cell tumors treated with surgery with and without adjuvant treatment: 54 cases (1998–2004). *J Am Vet Med Assoc*. 2006;228:91–95.
143. Murphy S, Sparkes AH, Blunden AS, Brearley MJ, Smith KC. Effects of stage and number of tumours on prognosis of dogs with cutaneous mast cell tumours. *Vet Rec*. 2006;158:287–291.
144. Murphy S, Sparkes AH, Smith KC, Blunden AS, Brearley MJ. Relationships between the histological grade of cutaneous mast cell tumours in dogs, their survival and the efficacy of surgical resection. *Vet Rec*. 2004;154:743–746.
145. Mutz ML, Boudreaux BB, Royal A, et al. Cytologic comparison of the percentage of mast cells in lymph node aspirate samples from clinically normal dogs versus dogs with allergic dermatologic disease and dogs with cutaneous mast cell tumors. *J Am Vet Med Assoc*. 2017 Aug;251:421–428.
146. Newkirk KM, Brannik EM, Kusewitt DF. Neoplasia and Tumor Biology. In: Zachary JF, ed. *Pathologic Basis of Veterinary Disease*. St. Louis, Missouri: Elsevier Inc.; 2017:286–323.
147. Northrup NC, Howerth EW, Harmon BG, et al. Variation among pathologists in the histologic grading of canine cutaneous. *J Vet Diagnostic Investig*. 2005;17:245–248.
148. O'Keefe DA, Couto CG. Fine-needle aspiration of the spleen as an aid in the diagnosis of splenomegaly. *J Vet Intern Med*. 1987 Sep;1:102–109.
149. Oakes DD, Sherck J. Splenic trauma, splenosis, and death from sepsis. 1982 Mar 12;247:1404–1405.
150. Owen LN. Skin - mastocytoma. In: Owen LN, ed. *TNM Classification of Tumours in Domestic Animals*. Geneva: World Health Organization.; 1980:14–15.
151. Pastarapatee N, Kijtawornrat A, Buranakarl C. Imbalance of autonomic nervous systems involved in ventricular arrhythmia after splenectomy in dogs. *J Vet Med Sci*. 2017;79:2002–2010.
152. Patil P, Shukla S, Bhake A, Hiwale KM. Accuracy of frozen section analysis in correlation with surgical pathology diagnosis. *Int J Res Med Sci*. 2015;3:399–404.
153. Patnaik AK, Ehler WJ, MacEwen EG. Canine Cutaneous Mast Cell Tumor: Morphologic Grading and Survival Time in 83 Dogs. *Vet Pathol*. 1984;21:469–474.
154. Patten SG, Boston SE, Monteith GJ. Outcome and prognostic factors for dogs with a histological diagnosis of splenic hematoma following splenectomy: 35 cases (2001-2013). *Can Vet J*. 2016;57:842–846.
155. Pavel G, Malancus RN, Condrea M. Fine Needle Aspiration Cytology (FNAC) of Neoplasms in Dogs and Cats, with Emphasis on Differential Diagnosis-A Retrospective Study. *Bull Univ Agric Sci Vet Med Cluj-Napoca Vet Med*. 2016;73:116–126.
156. Pecceu E, Serra Varela JC, Handel I, Piccinelli C, Milne E, Lawrence J. Ultrasound is a poor predictor of early or overt liver or spleen metastasis in dogs with high-risk mast cell tumours. *Vet Comp Oncol*. 2020 Sep;18:389–401.
157. La Perle KMD. Comparative Pathologists: Ultimate Control Freaks Seeking Validation! *Vet Pathol*. 2019 Jan;56:19–23.
158. Pierini A, Lubas G, Gori E, Binanti D, Millanta F, Marchetti V. Epidemiology of Breed-Related Mast Cell Tumour Occurrence and Prognostic Significance of Clinical Features in a Defined Population of Dogs in West-Central Italy. *Vet Sci*. 2019 Jun;6.
159. Pizzoni S, Sabattini S, Stefanello D, et al. Features and prognostic impact of distant metastases in 45 dogs with de novo stage IV cutaneous mast cell tumours: A prospective study. *Vet Comp Oncol*. 2018;16:28–36.
160. Ponce F, Marchal T, Magnol JP, et al. A morphological study of 608 cases of canine malignant

lymphoma in France with a focus on comparative similarities between canine and human lymphoma morphology. *Vet Pathol*. 2010;47:414–433.
161. Ponce F, Magnol J-P, Ledieu D, et al. Prognostic significance of morphological subtypes in canine malignant lymphomas during chemotherapy. *Vet J*. 2004 Mar;167:158–166.
162. Pratschke KM, Atherton MJ, Sillito JA, Lamm CG. Evaluation of a modified proportional margins approach for surgical resection of mast cell tumors in dogs: 40 cases (2008-2012). *J Am Vet Med Assoc*. 2013 Nov;243:1436–1441.
163. Przeździecki R, Czopowicz M, Sapierzyński R. Accuracy of routine cytology and immunocytochemistry in preoperative diagnosis of oral amelanotic melanomas in dogs. *Vet Clin Pathol*. 2015 Dec;44:597–604.
164. Raskin RE. General Categories of Cytologic Interpretation. In: Raskin RE, Meyer DJBT, eds. *Canine and Feline Cytology - A Color Atlas and Interpretation Guide*. St. Louis: W.B. Saunders; 2016:16–33.
165. Raskin RE. Hemolymphatic System. In: Raskin RE, Meyer DJBT, eds. *Canine and Feline Cytology - A Color Atlas and Interpretation Guide*. St. Louis: W.B. Saunders; 2016:91–137.
166. Raskin RE. Chapter 4 - Hemolymphatic System. In: Raskin RE, Meyer DJBT, eds. *Canine and Feline Cytology - A Color Atlas and Interpretation Guide*. St. Louis: W.B. Saunders; 2016:91–137.
167. Ressel L, Finotello R. Lymph node histology for the assessment of residual neoplastic disease in canine mast cell tumours: does the presence of metachromatic granules always identify mast cells? *Vet Comp Oncol*. 2017;15:1119–1121.
168. Richardson EF, Brown NO. Hematological and biochemical changes and results of aerobic bacteriological culturing in dogs undergoing splenectomy. *J Am Anim Hosp Assoc*. 1996;32:199–210.
169. Roccabianca P, Vernau W, Caniatti M, Moore PF. Feline large granular lymphocyte (LGL) lymphoma with secondary leukemia: primary intestinal origin with predominance of a CD3/CD8(alpha)(alpha) phenotype. *Vet Pathol*. 2006 Jan;43:15–28.
170. Rodak O, Dzimira S, Podolak A, Płóciennik M, Niżański W. Accuracy of ultrasonography and fine-needle aspiration cytology in the diagnosis of prostate diseases in dogs. *Reprod Domest Anim*. 2018 Nov;53 Suppl 3:79–84.
171. Romansik EM, Reilly CM, Kass PH, Moore PF, London CA. Mitotic index is predictive for survival for K9 cutaneous MCTs. *Vet Pathol*. 2007;44:335–341.
172. Rossi F, Körner M, Suárez J, et al. Computed tomographic-lymphography as a complementary technique for lymph node staging in dogs with malignant tumors of various sites. *Vet Radiol ultrasound Off J Am Coll Vet Radiol Int Vet Radiol Assoc*. 2018 Mar;59:155–162.
173. Roth L. Comparison of liver cytology and biopsy diagnoses in dogs and cats: 56 cases. *Vet Clin Pathol*. 2001;30:35–38.
174. Sabattini S, Renzi A, Marconato L, et al. Comparison between May-Grünwald-Giemsa and rapid cytological stains in fine-needle aspirates of canine mast cell tumour: Diagnostic and prognostic implications. *Vet Comp Oncol*. 2018;16:511–517.
175. Sabattini S, Bottero E, Turba ME, Vicchi F, Bo S, Bettini G. Differentiating feline inflammatory bowel disease from alimentary lymphoma in duodenal endoscopic biopsies. *J Small Anim Pract*. 2016 Aug 1;57:396–401.
176. Sabattini S, Renzi A, Buracco P, et al. Comparative Assessment of the Accuracy of Cytological and Histologic Biopsies in the Diagnosis of Canine Bone Lesions. *J Vet Intern Med*. 2017 May;31:864–871.
177. Sabattini S, Renzi A, Rigillo A, et al. Cytological differentiation between benign and malignant perianal gland proliferative lesions in dogs: a preliminary study. *J Small Anim Pract*. 2019 Oct;60:616–622.
178. Sabattini S, Lopparelli RM, Rigillo A, et al. Canine Splenic Nodular Lymphoid Lesions: Immunophenotyping, Proliferative Activity, and Clonality Assessment. *Vet Pathol*. 2018;55:645–653.
179. Santagostino SF, Mortellaro CM, Boracchi P, et al. Feline upper respiratory tract lymphoma: site, cyto-histology, phenotype, FeLV expression, and prognosis. *Vet Pathol*. 2015 Mar;52:250–259.
180. Sapierzyński R, Czopowicz M, Jagielski D. Metastatic lymphadenomegaly in dogs - cytological study. *Pol J Vet Sci*. 2017 Dec;20:731–736.
181. Sartor AJ, Bentley AM, Brown DC. Association between previous splenectomy and gastric

dilatation-volvulus in dogs: 453 cases (2004–2009). *J Am Vet Med Assoc*. 2013;242:1381–1384.
182. Sato H, Fujino Y, Chino J, et al. Prognostic analyses on anatomical and morphological classification of feline lymphoma. *J Vet Med Sci*. 2014 Jun;76:807–811.
183. Sayag D, Fournel-Fleury C, Ponce F. Prognostic significance of morphotypes in canine lymphomas: A systematic review of literature. *Vet Comp Oncol*. 2018 Mar;16:12–19.
184. Scarpa F, Sabattini S, Bettini G. Cytological grading of canine cutaneous mast cell tumours. *Vet Comp Oncol*. 2016;14:245–251.
185. Scase TJ, Edwards D, Miller J, et al. Canine mast cell tumors: correlation of apoptosis and proliferation markers with prognosis. *J Vet Intern Med*. 2006;20:151–158.
186. Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. *Philos Trans R Soc B Biol Sci*. 2015;370:20140231.
187. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–682.
188. Sever AR, Mills P, Jones SE, et al. Preoperative Sentinel Node Identification With Ultrasound Using Microbubbles in Patients With Breast Cancer. *Am J Roentgenol*. 2011 Feb 1;196:251–256.
189. Sharkey LC, Dial SM, Matz ME. Maximizing the Diagnostic Value of Cytology in Small Animal Practice. *Vet Clin North Am - Small Anim Pract*. 2007;37:351–372.
190. Sharkey LC, Seelig DM, Overmann J. All lesions great and small, part 1: Diagnostic cytology in veterinary medicine. *Diagn Cytopathol*. 2014 Jun 1;42:535–543.
191. Sharkey LC, Seelig DM, Overmann J. All lesions great and small, part 2: Diagnostic cytology in veterinary medicine. *Diagn Cytopathol*. 2014 Jun 1;42:544–552.
192. da Silva EZM, Jamur MC, Oliver C. *Mast Cell Function: A New Vision of an Old Cell*. Vol. 62, *Journal of Histochemistry & Cytochemistry*. 2014.
193. da Silva MBG, Pascoal IC, da Silva ÓP, et al. Nonconformities in Veterinary Cytopathological Examinations: A Retrospective Study of Unsuitable Samples for Analysis. *Acta Vet Brno*. 2019;69:251–261.
194. Sirvan SS, Demir IA, Irmak F, Kafi M, Budak K, Karsidag S. Does Wide Excisional Biopsy in Skin Cancer Prevent Finding The Real Sentinel Lymph Node? *J Investig Surg*. 2020 Sep 13;33:741–747.
195. Skinner OT, Boston SE, Souza CH de M. Patterns of lymph node metastasis identified following bilateral mandibular and medial retropharyngeal lymphadenectomy in 31 dogs with malignancies of the head. *Vet Comp Oncol*. 2017 Sep;15:881–889.
196. Sledge DG, Webster J, Kiupel M. Canine cutaneous mast cell tumors: A combined clinical and pathologic approach to diagnosis, prognosis, and treatment selection. *Vet J*. 2016 Sep;215:43–54.
197. Soultani C, Patsikas MN, Karayannopoulou M, et al. Assessment of Sentinel Lymph Node Metastasis in Canine Mammary Gland Tumors Using Computed Tomographic Indirect Lymphography. *Vet Radiol Ultrasound*. 2017;58:186–196.
198. Spangler WL, Culbertson MR. Prevalence, type, and importance of splenic diseases in dogs: 1,480 cases (1985-1989). *J Am Vet Med Assoc*. 1992 Mar;200:829–834.
199. Spangler WL, Kass PH. Pathologic factors affecting postsplenectomy survival in dogs. *J Vet Intern Med*. 1997;11:166–171.
200. Stefanello D, Valenti P, Faverzani S, et al. Ultrasound-Guided Cytology of Spleen and Liver: A Prognostic Tool in Canine Cutaneous Mast Cell Tumor. *J Vet Intern Med*. 2009 Sep;23:1051–1057.
201. Stefanello D, Buracco P, Sabattini S, et al. Comparison of 2- and 3-category histologic grading systems for predicting the presence of metastasis at the time of initial evaluation in dogs with cutaneous mast cell tumors: 386 cases (2009-2014). *J Am Vet Med Assoc*. 2015 Apr;246:765–769.
202. Stockhaus C, Van Den Ingh T, Rothuizen J, Teske E. A multistep approach in the cytologic evaluation of liver biopsy samples of dogs with hepatic diseases. *Vet Pathol*. 2004 Sep;41:461–470.
203. Suami H, Yamashita S, Soto-Miranda MA, Chang DW. Lymphatic Territories (Lymphosomes) in a Canine: An Animal Model for Investigation of Postoperative Lymphatic Alterations. *PLoS One*. 2013;8:1–9.
204. Sulce M, Marconato L, Martano M, et al. Utility of flow cytometry in canine primary cutaneous

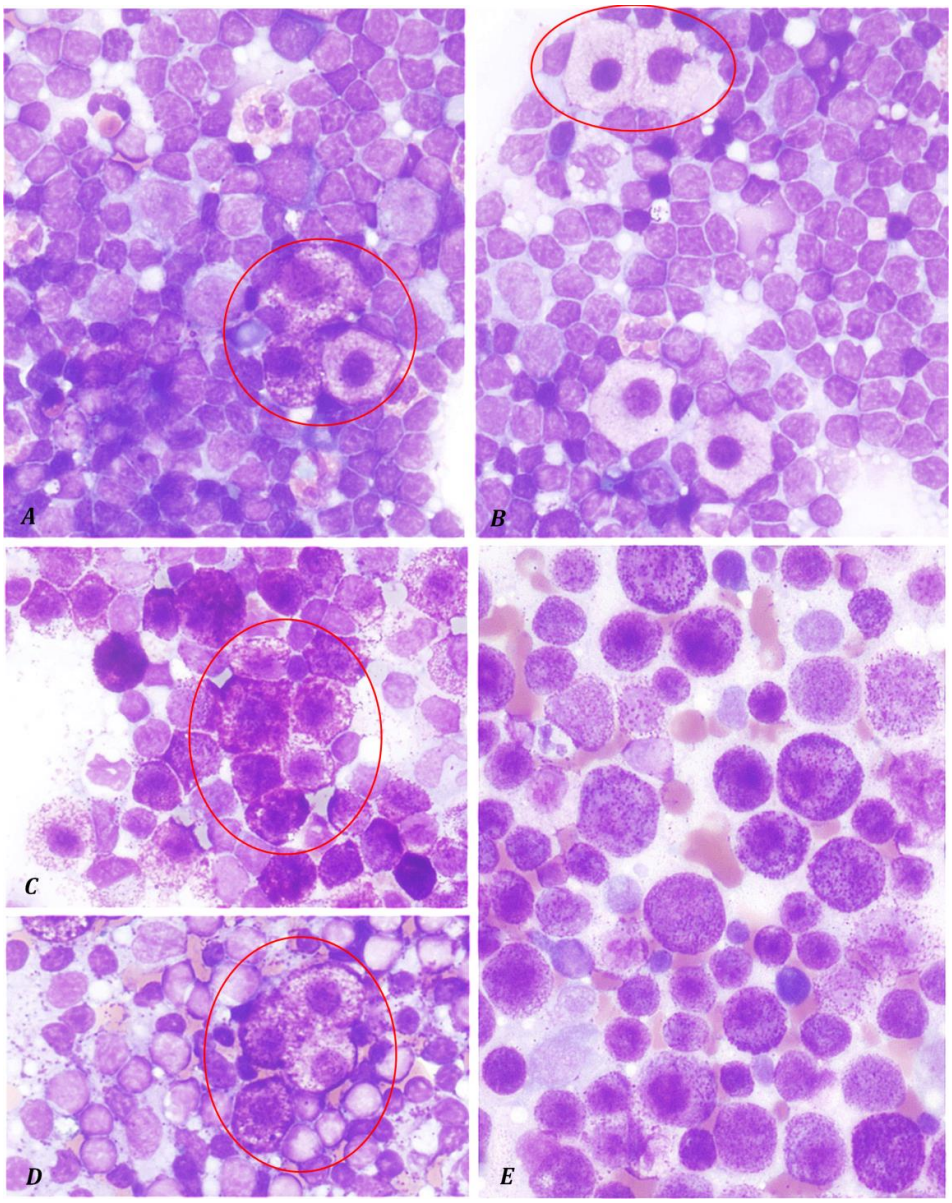
and matched nodal mast cell tumor. *Vet J*. 2018;242:15–23.
205. Takahashi T, Kadosawa T, Nagase M, et al. Visceral mast cell tumors in dogs: 10 cases (1982-1997). *J Am Vet Med Assoc*. 2000 Jan;216:222–226.
206. Tanis PJ, Boom RP, Koops HS, et al. Frozen section investigation of the sentinel node in malignant melanoma and breast cancer. *Ann Surg Oncol*. 2001 Apr;8:222–226.
207. Tecilla M, Gambini M, Forlani A, Caniatti M, Ghisleni G, Roccabianca P. Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. *PLoS One*. 2019;14.
208. Teske E, van Heerde P. Diagnostic value and reproducibility of fine-needle aspiration cytology in canine malignant lymphoma. *Vet Q*. 1996 Sep 1;18:112–115.
209. Thamm DH, Turek MM, Vail DM. Outcome and prognostic factors following adjuvant prednisone/vinblastine chemotherapy for high-risk canine mast cell tumour: 61 cases. *J Vet Med Sci*. 2006 Jun;68:581–587.
210. Thompson JJ, Pearl DL, Yager JA, Best SJ, Coomber BL, Foster RA. Canine subcutaneous mast cell tumor: Characterization and prognostic indices. *Vet Pathol*. 2011;48:156–168.
211. Tiškina V, Capligina V, Must K, Berzina I, Ranka R, Jokelainen P. Fatal Babesia canis canis infection in a splenectomized Estonian dog. *Acta Vet Scand*. 2016;58:1–6.
212. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65:87–108.
213. Tuohy JL, Milgram J, Worley DR, Dernell WS. A review of sentinel lymph node evaluation and the need for its incorporation into veterinary oncology. *Vet Comp Oncol*. 2009;7:81–91.
214. Twomey LN, Alleman AR. Cytodiagnosis of feline lymphoma. 2005;27.
215. U.S. Food and Drug Administration. Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests - Guidance for Industry and FDA Staff. 2007.
216. Vail DM, Pinkerton ME, Young KM. Hematopoietic Tumors. In: Withrow SJ, Vail DM, Page RL, eds. *Small Animal Clinical Oncology*. St. Louis: Elsevier Saunders; 2013:608–678.
217. Valli VE, Kass PH, Myint MS, Scott F. Canine Lymphomas: Association of Classification Type, Disease Stage, Tumor Subtype, Mitotic Rate, and Treatment With Survival. *Vet Pathol*. 2013;50:738–748.
218. Valli VE, Myint M, Barthel A, et al. Classification of canine malignant lymphomas according to the world health organization criteria. *Vet Pathol*. 2011;48:198–211.
219. Valli VE, Vernau W, De Lorimier LP, Graham PS, Moore PF. Canine Indolent Nodular Lymphoma. *Vet Pathol*. 2006 May 26;43:241–256.
220. Valli VE. Normal and Benign Reactive Hematopoietic Tissues - Spleen. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:47–76.
221. Valli VE. Chronic Myeloproliferative Diseases. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:425–460.
222. Valli VE. T-Cell and NK-Cell Neoplasms. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:275–366.
223. Valli VE. B-Cell Neoplasms. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:119–274.
224. Valli VE. Hodgkin's Lymphoma. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:491–504.
225. Valli VE. Myeloproliferative/Myelodysplastic Diseases. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:461–467.
226. Valli VE. Acute Myeloid Leukemias. In: Valli VE, ed. *Veterinary Comparative Hematology*. Ames, Iowa: Blackwell Publishing; 2007:367–424.
227. Valli VE, Bienzle D, Meuten DJ, Linder KE. Tumors of the Hemolymphatic System. In: Meuten DJ, ed. *Tumors in domestic animals*. Ames, Iowa: Wiley/Blackwell; 2017:203–321.
228. Valli VE, Jacobs RM, Norris A, et al. The histologic classification of 602 cases of feline lymphoproliferative disease using the National Cancer Institute working formulation. *J Vet diagnostic Investig Off Publ Am Assoc Vet Lab Diagnosticians, Inc*. 2000 Jul;12:295–306.
229. Vos JH, van den Ingh TS, van Mil FN. Non-exfoliative canine cytology: the value of fine needle aspiration and scraping cytology. *Vet Q*. 1989 Oct;11:222–231.
230. Warland J, Amores-Fuster I, Newbury W, Brearley M, Dobson J. The utility of staging in canine mast cell tumours. *Vet Comp Oncol*. 2014;12:287–298.
231. Warland J, Brioschi V, Owen L, Dobson J. Canine mast cell tumours: decision-making and treatment. *In Pract*. 2015 Jul 1;37:315 LP – 332.

232. Watson AT, Penninck D, Knoll JS, Keating JH, Sutherland-Smith J. Safety and correlation of test results of combined ultrasound-guided fine-needle aspiration and needle core biopsy of the canine spleen. *Vet Radiol Ultrasound*. 2011 Jun;52:317–322.
233. Watson PF, Petrie A. Method agreement analysis: A review of correct methodology. *Theriogenology*. 2010;73:1167–1179.
234. Weishaar KM, Thamm DH, Worley DR, Kamstock DA. Correlation of nodal mast cells with clinical outcome in dogs with mast cell tumour and a proposed classification system for the evaluation of node metastasis. *J Comp Pathol*. 2014;151:329–338.
235. Wells A, Yates C, Shepard CR. E-cadherin as an indicator of mesenchymal to epithelial reverting transitions during the metastatic seeding of disseminated carcinomas. *Clin Exp Metastasis*. 2008;25:621–628.
236. Wendelburg KM, O'Toole TE, McCobb E, Price LL, Lyons JA, Berg J. Risk factors for perioperative death in dogs undergoing splenectomy for splenic masses: 539 cases (2001-2012). *J Am Vet Med Assoc*. 2014 Dec;245:1382–1390.
237. Whitehair JG, Griffey SM, Olander HJ, Vasseur PB, Naydan D. The Accuracy of Intraoperative Diagnoses Based on Examination of Frozen Sections A Prospective Comparison with Paraffin-Embedded Sections. *Vet Surg*. 1993 Jul 1;22:255–259.
238. Wolfesberger B, Fuchs-Baumgartinger A, Greß V, et al. World Health Organisation Classification of Lymphoid Tumours in Veterinary and Human Medicine: a Comparative Evaluation of Gastrointestinal Lymphomas in 61 Cats. *J Comp Pathol*. 2018 Feb;159:1–10.
239. Wolfesberger B, Skor O, Hammer SE, et al. Does categorisation of lymphoma subtypes according to the World Health Organization classification predict clinical outcome in cats? *J Feline Med Surg*. 2016 Aug 30;19:897–906.
240. Worley DR. Incorporation of sentinel lymph node mapping in dogs with mast cell tumours: 20 consecutive procedures. *Vet Comp Oncol*. 2014;12:215–226.
241. Wu S, Wang Y, Zhang N, et al. Intraoperative Touch Imprint Cytology in Targeted Axillary Dissection After Neoadjuvant Chemotherapy for Breast Cancer Patients with Initial Axillary Metastasis. *Ann Surg Oncol*. 2018;25:3150–3157.
242. Zandvliet M. Canine lymphoma: a review. *Vet Q*. 2016 Apr 2;36:76–104.

APPENDICES

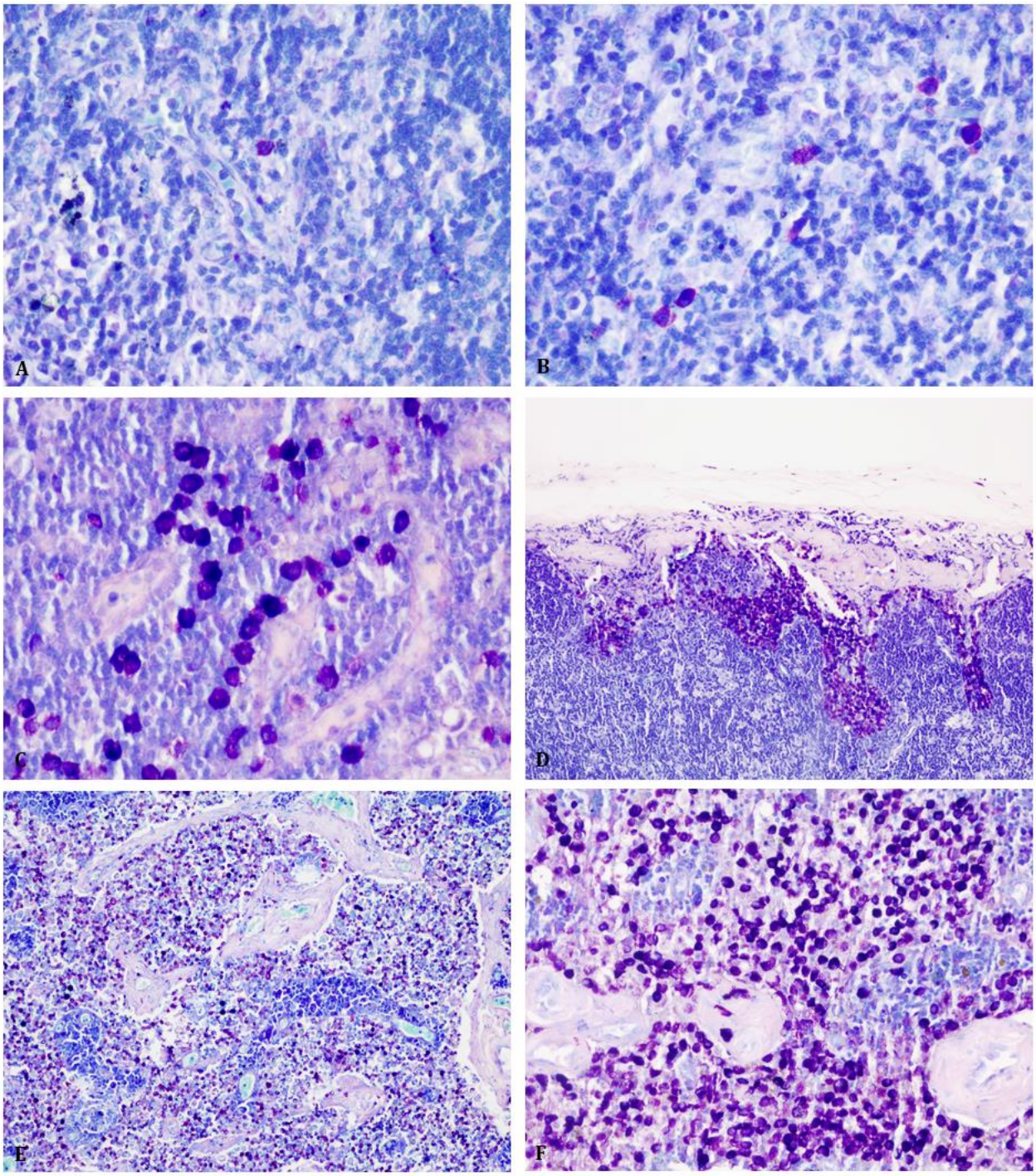
APPENDIX 1 – Supplementary material to Chapter 3 – Section 1 and Section 2

Supplementary Figure 1 – Examples of morphological features evaluated to classify the cytological cases



A) Encircled in red, example of mast cells (MCs) arranged in a triplet. **B)** Encircled in red, example of MCs arranged in a couple. **C) and D)** Encircled in red, examples of aggregates composed by >3 MCs. **E)** Example of replacement of >50% total lymphoid cells by MCs.

Supplementary Figure 2 – Examples of histopathological features evaluated to classify the histological cases according to the interpretative system by Weishaar et al.[234]



All sections were stained with Giemsa histochemical stain. **A)** Lymph node (LN) classified as HN0 ("non-metastatic"). An individualized mast cell (MC) is present in a 400x high power field (HPF). **B)** LN classified as HN1 ("pre-metastatic"). Greater than 3 individualized MCs are present in a 400x field. **C) and D)** LNs classified as HN2 ("early metastasis"). Clusters composed by >3 MCs (figure C) or sinusoidal sheets of MCs (figure D) are needed to assign a LN to the HN2 class. **E) and F)** LNs classified as HN3 ("overt metastasis"). Large clusters of MCs effacing the normal nodal architecture are required to assign a LN to the HN3 class.

APPENDIX 2 – Supplementary material to Chapter 3, Section 1

Supplementary Figure 1 – Examples of microphotographs evaluated in the current study

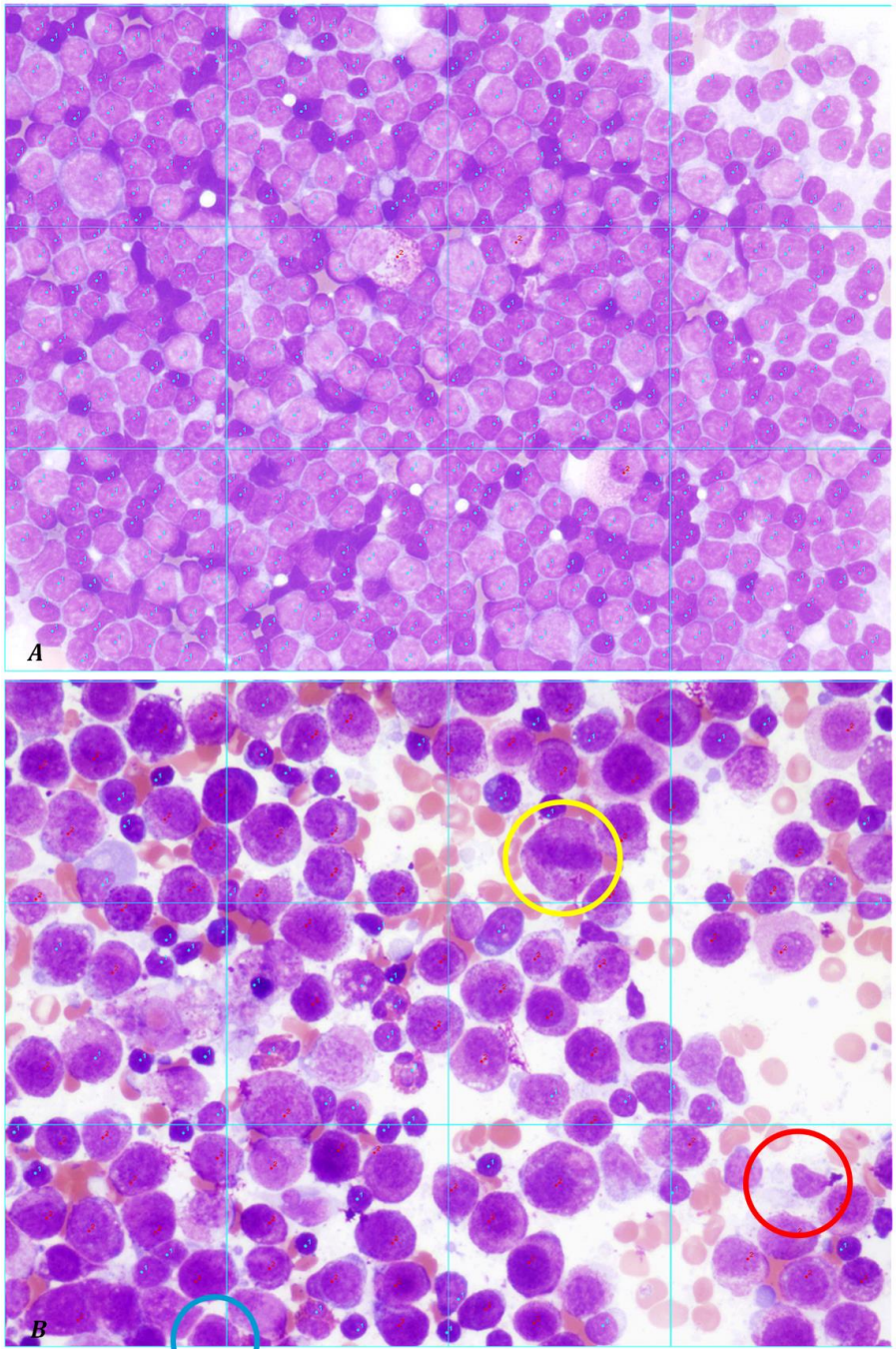
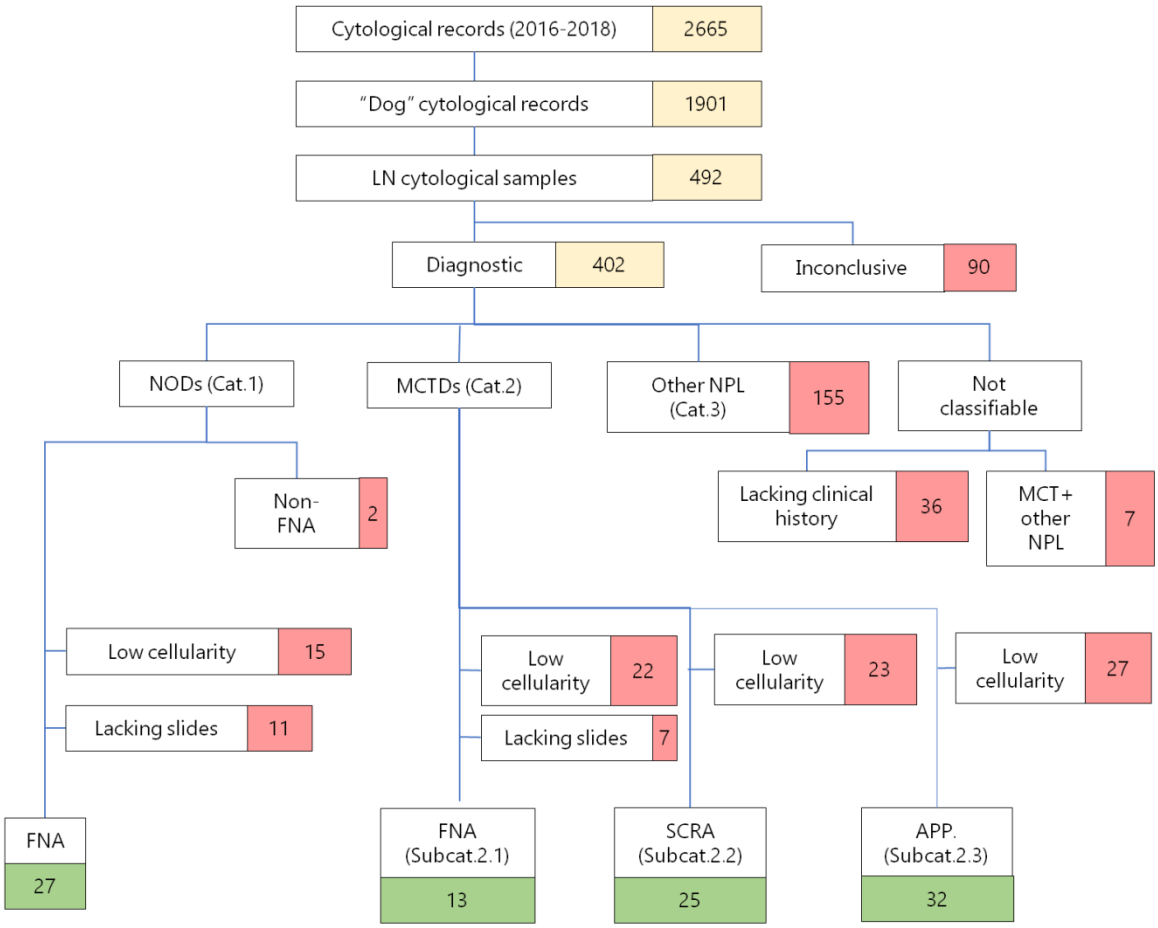


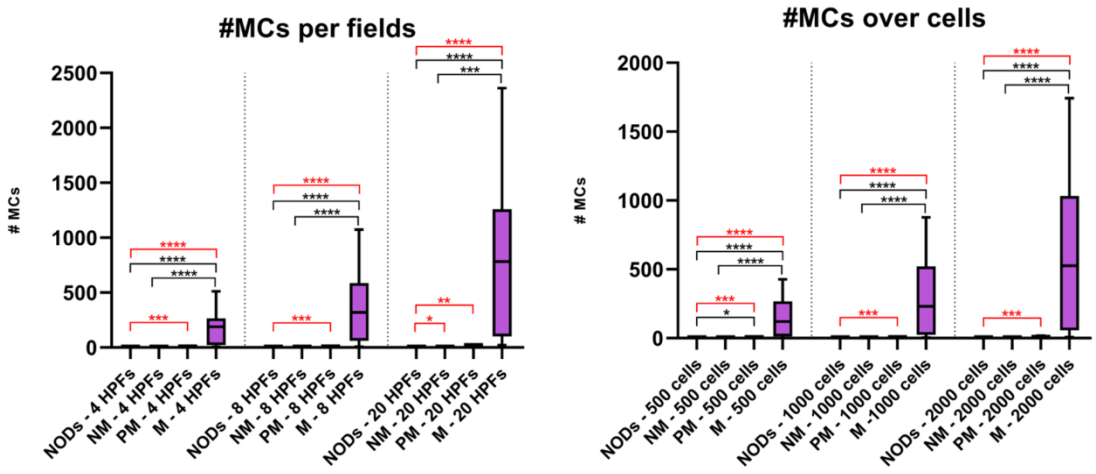
Figure A. Example microphotograph showing the square grid applied to facilitate cell counting. Highlighted by red dots, mast cells (MCs). Highlighted by blue dots, cells other than MCs. **Figure B.** Examples of cells excluded from counting. Encircled in yellow, a mitotic cell. Encircled in red, a neked nucleus. Encircled in blue, a cell only partially depicted..

Supplementary Figure 2 – Selection workflow for cases included in the current study.



Legend: FNA, fine-needle aspiration; LN, lymph node; MCT, mast cell tumor; MCTDs, mast cell tumor-bearing dogs; NODs, non-oncological dogs; NPL, neoplasm; cellularity; SCRA., scraping smears; TIC, touch imprints.

Supplementary Figure 3 – Nodal mast cells in non-oncological dogs (NODs) compared to cases obtained from mast cell tumor-bearing dogs (MCTDs) and classified as "non metastatic" (NM), "possibly metastatic" (PM), and "metastatic" (M)



The box represents the interquartile range, with the internal horizontal line indicating the median value. The tip of the upper whisker indicates the maximum value while the tip of the lower whisker reports the minimum value. Statistically significant differences as determined with Kruskall-Wallis test with Dunn's correction for multiple comparisons, are indicated by black bars and stars, and mirrored those reported in Section 1 – Figure 2 and described in the main text . Statistically significant differences as determined with Mann-Withney test are indicated by red bars and stars. Legend: HPFs, high power fields; MCs, mast cells; #, number of; *, p<0,05; **, p<0,01; ***, p<0,001; ****, p<0.0001.

Supplementary Table 1 - Criteria and corresponding interpretation applied for the cytological evaluation of LN metastatic status in MCT-bearing dogs

Cytological criteria	Interpretation
No MCs observed; OR >50% small lymphocytes with a mixed population of prolymphocytes, lymphoblasts, plasmacells, and/or few to moderate numbers of macrophages, neutrophils, and eosinophils, and/or rare individual MCs	**Non-metastatic** *(former "normal" + "reactive lymphoid hyperplasia")*
2-3 incidences of MCs aggregated in couples or triplets	**Possibly metastatic** *(former "possible metastasis")*
>3 incidences of MCs aggregated in couples or triplets and/or 2-5 aggregates composed by >3 MCs; OR Effacement of lymphoid tissue by MCs, and/or aggregated, poorly differentiated MCs (pleomorphism, anisocytosis, anisokaryosis, and/or decreased or variable granulation), and/or >5 aggregates composed by >3 MCs	**Metastatic** *(former "probable metastasis" + "certain metastasis")*

Legend: MC, mast cell. Table adapted from Krick et al.[104]

Supplementary Table 2 - Signalment data of dogs included in the study and details regarding investigated LNs.

Dog ID	Breed	Sex	Age	Primary process	Case ID	LN location	LN size	RLN or SLN	Category / Subcategory	Histology (agreement)	Retrieved cytological slides	# cytological slides used for micropic	Entirely-counted micropic
1	Mongrel	NF	3yy	Rhinitis, dermatitis, and reactive histiocytosis	#8	R mandibular	Megalic (mild)	RLN	Cat.1	n/a	2	2	5
					#28	L mandibular	Normal	RLN	Cat.1	n/a	2	2	7
2	Mongrel	CM	9yy	Necrotizing gingivitis	#9	Retromandibular	n/d	RLN	Cat.1	n/a	2	1	8
3	Dachshund	F	5yy	Bilateral purulent otitis	#35	R prescapular	Megalic (moderate)	RLN	Cat.1	n/a	2	1	7
					#16	L prescapular	Megalic (moderate)	RLN	Cat.1	n/a	1	1	6
4	Ibizan Hound	Ma	1yy	Leishmaniasis	#18	L submandibular	Megalic (mild)	RLN	Cat.1	n/a	2	2	6
5	Dachshund	NF	9yy	Leishmaniasis	#37	Prescapular	3 cm	RLN	Cat.1	n/a	2	2	8
6	Shar Pei	NF	4yy 8mm	Hepatic amyloidosis	#38	R popliteal	1 cm	RLN	Cat.1	n/a	1	1	8
7	Mongrel	Ma	9yy	Allergic dermatopathy	#21	n/d	n/d	RLN	Cat.1	n/a	2	2	13
8	English Setter	NF	12yy	Stomatitis	#2	Cervical (ventral)	n/d	RLN	Cat.1	n/a	2	2	12
9	Jack Russell Terrier	NF	11yy 11mm	MCT (n/d)	#39	L iliac	Megalic (n/d)	RLN	Subcat.M Subcat.2.1	n/a	3	1	14
					#19	R iliac	Megalic (n/d)	RLN	Subcat.M Subcat.2.1	n/a	1	1	14
10	Coton de Tulear	NF	13yy 11mm	MCT (L thorax)	#20	L axillary	n/d	RLN	Subcat.M Subcat.2.1	n/a	2	2	19
11	French Bulldog	Ma	7yy	MCT (L forelimb)	#40	L cervical (superficial)	n/d	RLN	Subcat.M Subcat.2.1	HN3 (Y)	3	1	13

Dog ID	Breed	Sex	Age	Primary process	Case ID	LN location	LN size	RLN or SLN	Category / Subcategory	Histology (agreement)	Retrieved cytological slides	# cytological slides used for micropic	Entirely-counted micropic
12	Beagle	Ma	11yy	MCT (n/d)	#17	R prescapular	Megalic (severe)	RLN	Subcat.M Subcat.2.1	n/a	4	2	14
13	Tosa Inu	Ma	4yy	MCT (prepuce)	#6	Popliteal	Megalic (moderate)	RLN	Subcat.NM Subcat.2.1	n/a	2	2	8
14	Labrador Retriever	Ma	5yy	MCT (R hindlimb)	#34	R prescapular	Megalic (severe)	RLN	Subcat.NM Subcat.2.1	n/a	2	2	16
15	Mongrel	NF	7yy 11mm	MCT (L flank)	#1	L inguinal	Megalic (mild)	RLN	Subcat.NM Subcat.2.1	HN2 (N)	2	2	10
16	Maltese	CM	5yy	MCT (R thorax)	#27	L popliteal	Megalic (mild)	RLN	Subcat.NM Subcat.2.1	n/a	2	2	10
17	Boxer	F	12yy	MCT (breast; L forelimb – III digit)	#36	L prescapular	Megalic (n/d)	RLN	Subcat.PM Subcat.2.1	HN2 (P)	3	1	14
18	Mongrel	NF	6yy	MCT (breast)	#14	R inguinal (lateral)	Normal	RLN	Subcat.M Subcat.2.2	HN3 (Y)	2	1	11
					#33	R inguinal (medial)	Normal	RLN	Subcat.M Subcat.2.2	HN3 (Y)	2	1	6
19	Labrador Retriever	Ma	7yy	MCT (prepuce)	#15	R inguinal	3x0.8x0.5 cm	SLN	Subcat.PM Subcat.2.2	HN0 (N)	1	1	9
20	Mongrel	CM	12yy	MCT (head)	#23	L prescapular #2	1.2x0.7x0.4 cm	SLN	Subcat.NM Subcat.2.2	HN0 (Y)	2	2	7
					#3	L prescapular #2	1.2x0.7x0.4 cm	SLN	Subcat.NM Subcat.2.3	HN0 (Y)	1	1	8
					#24	L prescapular #1	2.45x1.1x0.95 cm	SLN	Subcat.NM Subcat.2.3	HN2 (N)	1	1	6
					#22	L prescapular #3	2.4x1.3x0.6 cm	SLN	Subcat.NM Subcat.2.3	HN0 (Y)	1	1	10
				MCT (breast)	#4	R inguinal	1.35x0.95x0.45 cm	SLN	Subcat.M Subcat.2.2	HN2 (Y)	1	1	7

Dog ID	Breed	Sex	Age	Primary process	Case ID	LN location	LN size	RLN or SLN	Category / Subcategory	Histology (agreement)	Retrieved cytological slides	# cytological slides used for micropic	Entirely-counted micropic
					#5	R inguinal	1.35x0.95x0.45 cm	SLN	Subcat.PM Subcat.2.3	HN2 (P)	1	1	9
				MCT (R thight)	#25	R inguinal #2	2.5x1x0.7 cm	SLN	Subcat.PM Subcat.2.2	HN2 (P)	1	1	8
					#26	R inguinal #2	2.5x1x0.7 cm	SLN	Subcat.M Subcat.2.3	HN2 (Y)	1	1	9
21	English Setter	M	6yy	MCT (R hindlimb)	#10	R inguinal #1	1.8x1.5 cm	SLN	Subcat.M Subcat.2.2	HN3 (Y)	1	1	12
					#30	R inguinal #2	0.9x1 cm	SLN	Subcat.M Subcat.2.2	HN3 (Y)	1	1	19
					#31	R popliteal	3.1x1.4 cm	SLN	Subcat.PM Subcat.2.3	HN2 (P)	1	1	7
22	Mongrel	NF	7yy	MCT (n/d)	#32	Prescapular	Normal	SLN	Subcat.M Subcat.2.2	HN3 (Y)	2	1	16
					#11	Prescapular	Normal	SLN	Subcat.M Subcat.2.3	HN3 (Y)	2	1	8
23	Boxer	NF	6yy	MCT (L forelimb)	#12	L prescapular	n/d	RLN	Subcat.NM Subcat.2.2	HN3 (N)	4	4	5
					#13	L prescapular	n/d	RLN	Subcat.M Subcat.2.3	HN3 (Y)	5	2	6
24	Mongrel	CM	4yy	MCT (L thorax)	#7	Axillary (accessory)	0.5x0.5x0.3 cm	SLN	Subcat.NM Subcat.2.3	HN0 (Y)	2	2	8
25	Golden Retriever	NF	4yy	MCT (R hock)	#29	R popliteal	n/d	RLN	Subcat.M Subcat.2.3	HN3 (Y)	5	4	12

Legend: Case ID, randomized blinded ID assigned to each cytological specimen considered as an independent case; Cat.1, nodal sample obtained from non-oncological dog; CM, castrated male; F, female; HNO, "non-metastatic" according to Weishaar et al.;[234] HN2, "early metastasis" according to Weishaar et al.; HN3, "overt metastasis" according to Weishaar et al.; L, left; LN, limph node; Ma, male; MCT, mast cell tumor; Megalic (mild), <2cm; Megalic (moderate), 2-5cm; Megalic (severe), i.e. >5cm; micropic, microphotographs; mm, months; NF, neutered female; n/a, not applied; n/d, not determined; R, right; RLN, regional lymph node; SNL, sentinel lymph node; Subcat.M, cytologically metastatic nodal sample obtained from MCT-bearing dog; Subcat.NM, cytologically non metastatic nodal sample obtained from MCT-bearing dog; Subcat.PM, cytologically possibly metastatic nodal sample obtained from MCT-bearing dog; Subcat.2.1, nodal sample obtained from MCT-bearing dog via fine-needle aspiration;

Subcat.2.2, nodal sample obtained from MCT-bearing dog via scraping smearing; Subcat.2.3, nodal sample obtained from MCT-bearing dog via touch imprinting; yy, years; (N), lack of agreement between cytological and histological diagnosis; (P), partial agreement between cytological and histological diagnosis; (Y), agreement between cytological and histological diagnosis.

Supplementary Table 3 – Median and mean absolute and percentage number of nodal mast cells in non-oncological dogs (NODs) and mast cell tumor-bearing dogs (MCTDs).

Group	# cases	Counting method	Median	Min-max range	Mean	Std dev	Median%	Min-max range%	Mean%	Std dev%
NODs (Cat .1)	10	4 HPFs	0	0-1	0.1	0.32	//	//	//	//
		8 HPFs	0	0-2	0.2	0.63	//	//	//	//
		20 HPFs	0.5	0-11	1.7	3.37	//	//	//	//
		500 cells	0	0	0	0	0	0	0	0
		1000 cells	0	0-0.86	0.09	0.27	0	0-0.09	0.01	0.03
		2000 cells	0	0-2	0.2	0.63	0	0-0.1	0.01	0.04
MCTDs (Cat.2)	30	4 HPFs	8	0-511	85.6	130.5	//	//	//	//
		8 HPFs	13.5	0-1073	173.1	268.6	//	//	//	//
		20 HPFs	28	1-2362	428.3	657.2	//	//	//	//
		500 cells	3.90	0-427.9	79.22	130.2	0.78	0-85.58	15.84	26.05
		1000 cells	6.06	0-877.1	153.8	257.5	0.61	0-87.71	15.38	25.75
		2000 cells	12.8	0-1743	307.9	513.6	0.64	0-87.14	15.39	25.68

Legend: HPFs, high power fields; Median%, median of the number of mast cells expressed as the percentage over 500, 1000, and 2000 cells; Min-max range, minimum-mximum range referred to the Median; Min-max range%, minimum-maximum range referred to the Median%; Mean%, mean of the number of mast cells expressed as the percentage over 500, 1000, and 2000 cells; Std dev, standard deviation; Std dev%, standard deviation referred to the Mean%.

Supplementary Table 4 – Median and mean absolute and percentage number of mast cells in nodal specimens obtained from mast cell tumor-bearing dogs (MCTDs) and classified as "non metastatic" (NM), "possibly metastatic" (PM), and "metastatic" (M).

Group	# cases	Counting method	Median	Min-max range	Mean	Std dev	Median%	Min-max range%	Mean%	Std dev%
NM (Subcat.2.NM)	10	4 HPFs	0	0-3	0.6	0.97	//	//	//	//
		8 HPFs	0.50	0-4	1.10	1.45	//	//	//	//
		20 HPFs	2	1-6	2.40	1.58	//	//	//	//
		500 cells	0	0-0.9	0.25	0.40	0	0-0.18	0.05	0.08
		1000 cells	0	0-2.6	0.68	1.02	0	0-0.26	0.07	0.10
		2000 cells	0.50	0-2.91	1.04	1.22	0.03	0-0.15	0.05	0.06
PM (Subcat.2.PM)	5	4 HPFs	5	1-9	5.40	2.97	//	//	//	//
		8 HPFs	9	3-16	8.80	5.12	//	//	//	//
		20 HPFs	28	6-28	19.40	11.78	//	//	//	//
		500 cells	4.39	0.97-6.38	4.05	2.26	0.88	0.19-1.28	0.81	0.45
		1000 cells	5.24	1.99-7.54	5.12	2.24	0.52	0.20-0.75	0.51	0.22
		2000 cells	8.70	4.84-16.82	9.39	5.02	0.43	0.24-0.84	0.47	0.25
M (Subcat.2.M)	15	4 HPFs	188	5-511	169	142.7	//	//	//	//
		8 HPFs	319	7-1073	342.6	296.4	//	//	//	//
		20 HPFs	782	21-2362	848.6	718.3	//	//	//	//
		500 cells	120	0.88-427.9	156.9	149	24	0.18-85.58	31.39	29.79
		1000 cells	230.8	4.34-877.1	305.4	296.7	23.08	0.43-87.71	30.54	29.67
		2000 cells	525.3	6.98-1743	611.9	590.2	26.26	0.35-87.14	30.60	29.51

Legend: HPFs, high power fields; Median%, median of the number of mast cells expressed as the percentage over 500, 1000, and 2000 cells; Min-max range, minimum-mximum range referred to the Median; Min-max range%, minimum-maximum range referred to the Median%; Mean%, mean of the number of mast cells expressed as the percentage over 500, 1000, and 2000 cells; Std dev, standard deviation; Std dev%, standard deviation referred to the Mean%.

Supplementary Table 5 – Median and mean absolute and percentage number of mast cells in nodal specimens obtained from mast cell tumor-bearing dogs (MCTDs) and sampled via fine-needle aspiration (FNA), scraping smearing (SCRA), and touch imprinting (TIC).

Group	# cases	Counting method	Median	Min-max range	Mean	Std dev	Median%	Min-max range%	Mean%	Std dev%
FNA (Subcat.2.1)	10	4 HPFs	94.5	0-511	156.3	182.6	//	//	//	//
		8 HPFs	181.5	0-1073	328.4	385.5	//	//	//	//
		20 HPFs	620	1-2361	833.6	931.7	//	//	//	//
		500 cells	124.4	0-427.9	165.2	181.8	24.89	0-85.58	33.04	36.37
		1000 cells	251.3	0-877.1	332.5	366.1	25.13	0-87.71	33.25	36.61
		2000 cells	476.9	0-1743	652.1	726.1	23.85	87.14	32.61	36.30
SCRA (Subcat.2.2)	5	4 HPFs	26	0-215	73.5	88.93	//	//	//	//
		8 HPFs	55.50	0-426	145.4	168.9	//	//	//	//
		20 HPFs	136.5	2-958	365.8	413.6	//	//	//	//
		500 cells	19.10	0-250	57.83	82.82	3.82	0-50	11.57	16.56
		1000 cells	36.55	0-389.5	102.6	138.5	3.66	0-38.95	10.26	13.86
		2000 cells	66.11	0-886.6	226.2	313.1	3.31	0-44.33	11.31	15.65
TIC (Subcat.2.3)	15	4 HPFs	7	0-208	27	63.97	//	//	//	//
		8 HPFs	10.50	0-319	45.60	97.88	//	//	//	//
		20 HPFs	17.50	1-610	85.60	187.2	//	//	//	//
		500 cells	2.25	0-120	14.66	37.22	0.45	0-24	2.93	7.44
		1000 cells	4.32	0-201.2	26.22	62.01	0.44	0-20.12	2.62	6.20
		2000 cells	9.42	0-309.7	45.29	95.03	0.47	0-15.49	2.27	4.75

Legend: HPFs, high power fields; Median%, median of the number of mast cells expressed as the percentage over 500, 1000, and 2000 cells; Min-max range, minimum-mximum range referred to the Median; Min-max range%, minimum-maximum range referred to the Median%; Mean%, mean of the number of mast cells expressed as the percentage over 500, 1000, and 2000 cells; Std dev, standard deviation; Std dev%, standard deviation referred to the Mean%.

Supplementary Table 6 – Median and mean cellularity per field of each category and subcategory (i.e. "group") investigated in the current study.

Group	# cases	Median	Min-max range	Mean	Std dev
NODs (Cat .1)	10	286.7	144.6-407.2	282.1	78.52
MCTDs (Cat.2)	30	222.2	74.8-448.6	222.5	87.67
NM (Subcat.2.NM)	10	251.8	144.2-448.6	260.9	87.48
PM (Subcat.2.PM)	5	238.6	125.2-301.4	231.4	65.59
M (Subcat.2.M)	15	162.2	74.8-357.8	193.9	88.40
FNA (Subcat.2.1)	10	147.2	74.80-230.4	153.5	53.20
SCRA (Subcat.2.2)	10	245.8	108.8-448.6	249.7	101.5
TIC (Subcat.2.3)	10	281.7	155.6-342.6	264.2	59.86

Legend: FNA, fine-needle aspiration; M, metastatic; MCTDs, mast cell tumor-bearing dogs; Min-max range, minimum-maximum range referred to the Median; NM, non metastatic; NODs, non-oncological dogs; PM, possibly metastatic; SCRA, scraping smearing; Std dev, standard deviation; TIC, touch imprinting.

APPENDIX 3 – Supplementary material to Chapter 3, Section 2

Supplementary Section 1 – Rationale applied by our research group to modify the cytological interpretative system proposed by Krick et al.[104] and to set criteria included in "Amendment 1" to the same system.

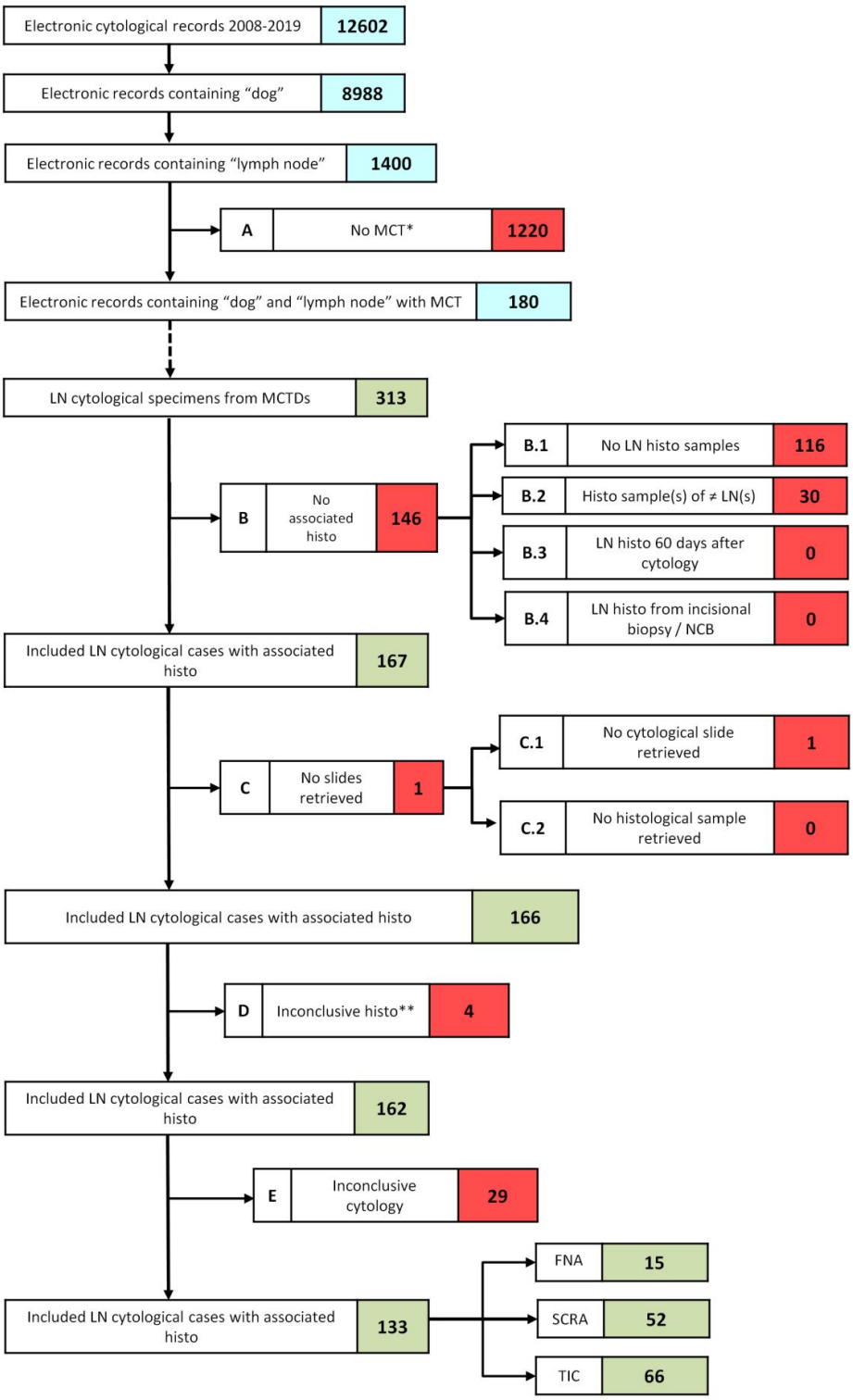
The changes made by our research group to modify the interpretative system proposed by Krick et al.,[104] aiming to standardize and make more reproducible the criteria for each diagnostic class, were the following:

- merge of the interpretative classes "probable metastasis" and "certain metastasis" into a single diagnostic class renamed "metastatic" (M), as already proposed by Mutz et al.,[145] with consequent incorporation of the criteria of the previous class "probable metastasis" into the new class M;
- merge of the classes "normal" and "reactive lymphoid hyperplasia" into a single class renamed "non metastatic" (NM), based on the absence of a statistically significant difference between the median survival times of dogs belonging to the two original classes, as reported by Krick et al.[104] and as already performed by previous investigators;[73]
- change of the name of the interpretative class "possible metastasis" into "possibly metastatic" (PM);
- introduction of cut-offs for the quantification of discrete, morphologically normal MCs. Specifically, the cut-off for shifting from NM class to PM class was set at 4 normal MCs for HPF on 8 HPFs in total (corresponding to 25-56 normal MCs/8 HPFs). This choice was based on the cut-off established by Weishaar et al.(Weishaar et al., 2014) for the transition from HNO class to HN1 class. The doubling of HPFs to be evaluated compared to the ones proposed by Weishaar et al. is in agreement with Scarpa et al.[184] who suggested that, in terms of cellularity, 1 histological HPF corresponds approximately to 2 cytological HPF. For the transition from PM class to M class, the cut-off of 4 MCs was duplicated on an arbitrary basis to 8 well differentiated MCs for HPF on 8 HPFs (corresponding to ≥57 normal MCs/8 HPFs);
- introduction of criteria and cut-offs for the quantification of discrete, atypical MCs. Considering that atypical MCs should not be present in a normal or reactive lymph node,[112,116] the cut-off for the latter was set on an arbitrary basis at 4 atypical MCs per 8

HPF for the transition from NM class to PM class, and at 8 atypical MCs per 8 HPF for the transition from PM class to M class.

The resulting interpretative system is illustrated in Section 2 – Table 1.

Supplementary Figure 1 – Selection workflow for cases included in the current study.



Legend: FNA, fine-needle aspiration; histo, histology/histological; LN, lymph node; MCT, mast cell tumor; MCTDs, mast cell tumor-bearing dogs; NCB, needle-core biopsy; NPL, neoplasm; cellularity; SCRA., scraping smears; TIC, touch imprints; *, no MCT reported in clincial history, or insufficient clincial history; **, histological samples characterized by nodal neoplasms other than MCT (e.g. lymphoma), or by the absence of LN tissue with the exclusive presence of different tissues (e.g. adipose tissue, salivary gland, etc.).

Supplementary Table 1 - Signalment data of dogs included in the study and details regarding investigated LNs.

Dog ID	Breed	Sex	Age (yy)	MCT location	LN ID	LN Location	LN size (cm)	Histological diagnosis (Weishaar et al. 2014)[234]	RLN or SLN	Sampling technique	Final case ID
#1	Setter	CM	12	Multiple MCT (disseminated)	#1.1	Axillary	ND	HN2	RLN	SCRA	#1.1.1
#9	Labrador Retriever	M	7	Ventral prepuce	#9.1	Inguinal L	2.4x0.7x0.4	HN2	SLN	SCRA	#9.1.1
										TIC	#9.1.2
					#9.2	Inguinal R #1	1.6X1X0.7	HN3	SLN	SCRA	#9.2.1
										TIC	#9.2.2
					#9.3	Inguinal R #2	3x0.8x0.5	HN0	SLN	SCRA	#9.3.1
#18	Mongrel	NF	6	Breast	#18.1	Inguinal lateral R	ND	HN3	RLN	SCRA	#18.1.1
										TIC	#18.1.2
					#18.2	Inguinal median R	ND	HN3	RLN	SCRA	#18.2.1
										TIC	#18.2.2
					#18.3	Inguinal medial R	ND	HN3	RLN	SCRA	#18.3.1
										TIC	#18.3.2
#19	Mongrel	NF	11	L forearm	#19.1	Prescapular L	1.8x1.2x0.3	HN0	SLN	TIC	#19.1.1
										SCRA	#19.1.2
#20	Breton	M	10	L stifle	#20.1	Mandibular #1	ND	HN0	RLN	TIC	#20.1.1
					#20.2	Mandibular #2	ND	HN0	RLN	TIC	#20.2.1
					#20.3	Mandibular #3	ND	HN0	RLN	TIC	#20.3.1
#24	English Setter	M	6	R hindlimb	#24.1	Inguinal R #1	1.8x1.5	HN3	SLN	SCRA	#24.1.1
										TIC	#24.1.2
					#24.2	Inguinal R #2	1x0.9	HN3	SLN	SCRA	#24.2.1
										TIC	#24.2.2
					#24.3	Popliteal R #1	2.1x1.4	HN2	SLN	SCRA	#24.3.1

Dog ID	Breed	Sex	Age (yy)	MCT location	LN ID	LN Location	LN size (cm)	Histological diagnosis (Weishaar et al. 2014)[234]	RLN or SLN	Sampling technique	Final case ID
										TIC	#24.3.2
#25	French Bulldog	M	7	L forelimb	#25.2	Cervical caudal L	3 cm	HN3	RLN	FNA	#25.2.1
#30	Boxer	NF	12	Breast (II L)	#30.1	Prescapular L	Normal	HN2	RLN	FNA	#30.1.1
#31	Cocker	NF	10	R stifle	#31.1	Inguinal R	Megalic (moderate)	HN3	RLN	TIC	#31.1.1
					#31.2	Popliteal R	Megalic (mild)	HN2	RLN	TIC	#31.2.1
#34	Golden Retriever	NF	4	R hindlimb	#34.1	Popliteal R	Megalic (moderate)	HN3	RLN	SCRA	#34.1.1
										TIC	#34.1.2
										FNA	#34.1.3
#38	Labrador Retriever	CM	11	R thoax	#38.1	Axillary #1	1.75x1.15	HN2	SLN	SCRA	#38.1.1
										TIC	#38.1.2
#40	English Setter	F	6	Multiple MCT (2x L flank + medial L thigh)	#40.1	Inguinal R	Normal	HN2	RLN	FNA	#40.1.1
#43	Labrador Retriever	F	6	Sternum	#43.1	Axillary #1	1.7x0.7	HN2	SLN	SCRA	#43.1.1
										TIC	#43.1.2
					#43.2	Axillary #2	1.1x0.6	HN2	SLN	SCRA	#43.2.1
										TIC	#43.2.2
					#43.3	Preaxillary	1x1.3	HN2	SLN	SCRA	#43.3.1
										TIC	#43.3.2
#46	Chihuahua	CM	9	L foot	#46.1	Popliteal L	0.4x0.3	HN0	SLN	TIC	#46.1.1
#52	Mongrel	CM	5	R elbow	#52.1	Prescapular	ND	HN2	RLN	TIC	#52.1.1
#53	American Staffordshire	M	4	L auricolar pinna	#53.1	Prescapular L #1	2.3x1.6x0.7	HN1	SLN	TIC	#53.1.1
										SCRA	#53.1.2
					#53.2	Prescapular L #2	1.2x0.7x0.4	HN1	SLN	TIC	#53.2.1
										SCRA	#53.2.2

Dog ID	Breed	Sex	Age (yy)	MCT location	LN ID	LN Location	LN size (cm)	Histological diagnosis (Weishaar et al. 2014)[234]	RLN or SLN	Sampling technique	Final case ID
#55	Mongrel	NF	7yy 11mm	L flank	#55.1	Inguinal L	Megalic (mild)	HN2	RLN	FNA	#55.1.1
#60	Sharpei	M	8	Upper lip	#60.3	Mandibular	Megalic (severe)	HN3	RLN	TIC	#60.3.2
#62	Mongrel (Labrador Retriever)	M	5	R hindlimb	#62.2	Inguinal R	2x0.9	HN1	SLN	TIC	#62.2.1
					#62.3	Popliteal R	0.85x0.4	HN0	SLN	TIC	#62.3.1
#63	Mongrel	NF	13	Breast	#63.1	Inguinal L #1	0.85x0.5	HN3	SLN	TIC	#63.1.1
					#63.2	Inguinal L #2	0.55x0.4	HN3	SLN	TIC	#63.2.2
#67	Boxer	M	9	L neck	#67.1	Prescapular L #1	1.85x1.2x0.5	HN0	SLN	SCRA	#67.1.1
										TIC	#67.1.2
					#67.2	Prescapular L #2	0.9x0.85x0.35	HN0	SLN	SCRA	#67.2.1
										TIC	#67.2.2
#73	Boxer	F	6	L forelimb	#73.1	Prescapular L	ND	HN3	RLN	SCRA	#73.1.1
										TIC	#73.1.2
#74	Golden Retriever	NF	6	L hindlimb	#74.1	Inguinal L	2.5x1.1x0.2	HN1	SLN	TIC	#74.1.1
										SCRA	#74.1.2
					#74.2	Popliteal L #1	0.7x0.6x0.3	HN0	SLN	TIC	#74.2.1
										SCRA	#74.2.2
					#74.3	Popliteal L #2	1.1x0.5x0.3	HN1	SLN	TIC	#74.3.1
										SCRA	#74.3.2
#75	Pug Dog	NF	10	L hock	#75.1	Popliteal L	ND	HN0	RLN	SCRA	#75.1.1
#76	Dachshund	NF	9	Tail (R)	#76.1	Abdominal	0.9x0.7x0.3	HN0	SLN	SCRA	#76.1.1
										TIC	#76.1.2
					#76.2	Lumbar aortic	1.2x0.6x0.2	HN0	SLN	SCRA	#76.2.1
										TIC	#76.2.2
#78	Setter	NF	10	L medial thigh	#78.1	Inguinal L #1	0.7x0.3x0.2	HN1	SLN	TIC	#78.1.1

Dog ID	Breed	Sex	Age (yy)	MCT location	LN ID	LN Location	LN size (cm)	Histological diagnosis (Weishaar et al. 2014)[234]	RLN or SLN	Sampling technique	Final case ID
					#78.2	Inguinal L #2	0.8x0.6x0.2	HN2	SLN	TIC	#78.2.1
#80	Mongrel	M	12	Multiple MCT (neck + abdomen)	#80.2	Prescapular	ND	HN3	RLN	TIC	#80.2.2
#81	Yorkshire	NF	13	R stifle	#81.2	Popliteal R	1.5x1	HN1	RLN	FNA	#81.2.1
										SCRA	#81.2.2
										TIC	#81.2.3
#85	English Setter	NF	9	Multiple MCT (>3 – bilateral hindlimb)	#85.1	Popliteal R	Normal	HN2	RLN	FNA	#85.1.1
#87	Pug Dog	M	11	R shoulder	#87.1	Prescapular L	Normal	HN1	RLN	FNA	#87.1.1
#92	Mongrel	CM	12	Breast	#92.1	Inguinal R #3	1.35x0.95x0.45	HN2	SLN	SCRA	#92.1.1
										TIC	#92.1.2
				R thigh	#92.2	Inguinal R #1	1.1x1x0.45	HN0	SLN	SCRA	#92.2.1
										TIC	#92.2.2
					#92.3	Inguinal R #2	2.5x1x0.7	HN2	SLN	SCRA	#92.3.1
										TIC	#92.3.2
				Head	#92.4	Prescapular L #1	2.45x1.1x0.95	HN2	SLN	SCRA	#92.4.1
										TIC	#92.4.2
					#92.5	Prescapular L #2	1.2x0.7x0.4	HN0	SLN	SCRA	#92.5.1
										TIC	#92.5.2
					#92.6	Prescapular L #3	2.4x1.3x0.6	HN0	SLN	SCRA	#92.6.1
										TIC	#92.6.2
#98	Boxer	NF	11yy 9mm	R thorax	#98.2	Axillary R	Megalic (moderate)	HN3	RLN	FNA	#98.2.1
#99	Labrador Retriever	M	4	R dorsum + R neck	#99.1	Prescapular R	Megalic (moderate)	HN0	RLN	FNA	#99.1.1
#100	Weimaraner	CM	7	R forearm	#100.1	Prescapular R	1.5x1x0.6	HN2	SLN	SCRA	#100.1.1
										TIC	#100.1.2
#102	Golden Retriever	NF	7	Neck	#102.	Prescapular L	1.5x0.95	HN2	SLN	SCRA	#102.1.1

Dog ID	Breed	Sex	Age (yy)	MCT location	LN ID	LN Location	LN size (cm)	Histological diagnosis (Weishaar et al. 2014)[234]	RLN or SLN	Sampling technique	Final case ID
					1					TIC	#102.1.2
#103	Labrador Retriever	M	6yy 6mm	Multiple MCT (R hock + R thorax + L shoulder + L groin + L prepuce + scrotum)	#103.1	Popliteal R	Normal	HN2	RLN	FNA	#103.1.1
					#103.3	Prescapular L	Megalic (mild)	HN3	RLN	FNA	#103.3.1
					#103.4	Prescapular R	ND	HN0	RLN	FNA	#103.4.1
#107	English Setter	F	8	Neck	#107.1	Prescapular L	ND	HN3	RLN	TIC	#107.1.1
#110	Mongrel	CM	4	L thorax	#110.1	Axillary L	0.5x0.5x0.3	HN0	SLN	TIC	#110.1.1
#112	Golden Retriever	NF	11	L thigh	#112.1	Inguinal L #1	1.7x0.8x0.25	HN2	SLN	SCRA	#112.1.1
										TIC	#112.1.2
					#112.2	Inguinal L #2	0.8x0.3x0.2	HN0	SLN	SCRA	#112.2.1
										TIC	#112.2.2
					#112.3	Inguinal L #3	0.9x0.6x0.35	HN0	SLN	SCRA	#112.3.1
										TIC	#112.3.2
#115	Pug Dog	M	3	L auricolar pinna	#115.1	Prescapular L	1.8X0.95X0.52	HN2	SLN	SCRA	#115.1.1
										TIC	#115.1.2
				R shoulder	#115.2	Prescapular R	2X1X0.5	HN0	SLN	SCRA	#115.2.1
										TIC	#115.2.2
#121	Pitbull	NF	5	R foot	#121.1	Middle iliac R	2.6x1.2x0.8	HN1	SLN	TIC	#121.1.1
									SLN	SCRA	#121.1.2
									SLN	FNA	#121.1.3
					#121.2	Popliteal R	1.3x0.7x0.6	HN2	SLN	SCRA	#121.2.2
#130	Labrador Retriever	M	11	Scrotum (x2)	#130.1	Inguinal R #1	1.4x1.3	HN0	SLN	TIC	#130.1.1
					#130.2	Inguinal R #2	0.9x0.6	HN0	SLN	TIC	#130.2.1

Dog ID	Breed	Sex	Age (yy)	MCT location	LN ID	LN Location	LN size (cm)	Histological diagnosis (Weishaar et al. 2014)[234]	RLN or SLN	Sampling technique	Final case ID
#132	Labrador Retriever	M	1	Nasal planum	#132.1	Mandibular L #1	1.9x1.3	HN0	SLN	SCRA	#132.1.1
										TIC	#132.1.2
					#132.2	Mandibular L #2	1.2x1.2	HN0	SLN	SCRA	#132.2.1
										TIC	#132.2.2
					#132.3	Mandibular L #3	1.4x1.1	HN0	SLN	SCRA	#132.3.1
										TIC	#132.3.2
					#132.4	Mandibular R #1	0.6x0.4	HN0	SLN	SCRA	#132.4.1
					#132.5	Mandibular R #2	1x0.7	HN0	SLN	SCRA	#132.5.1
										TIC	#132.5.2
					#132.6	Mandibular R #3	1.2x0.6	HN0	SLN	SCRA	#132.6.1
										TIC	#132.6.2
#134	Mongrel	CM	5	R hock	#134.1	Popliteal R	Megalic (mild)	HN1	RLN	FNA	#134.1.2
#136	Golden Retriever	CM	6	L neck (x2)	#136.3	Popliteal L	ND	HN1	RLN	TIC	#136.3.2
#143	Mongrel	NF	7	Thorax	#143.1	Prescapular	1.4x0.7	HN3	SLN	SCRA	#143.1.1
										TIC	#143.1.2
#148	Tosa Inu	M	4	L prepuce	#148.1	Inguinal #1	1.8x0.9	HN2	SLN	SCRA	#148.1.1
										TIC	#148.1.2
					#148.2	Inguinal #2	1.4x0.5	HN2	SLN	SCRA	#148.2.1

Legend: CM, castrated male; cm; centimeter; F, female; FNA, fine needle aspirate; Final case ID, ID identifying eachs ingel case and taking simultaneously into account the ID of the dog and the ID of the LN from which the case was sampled: HNO, “non-metastatic” according to Weishaar et al.;[234] HN1, “pre-metastatic” according to Weishaar et al.; HN2, “early metastasis” according to Weishaar et al.; HN3, “overt metastasis” according to Weishaar et al.; L, left; LN, limph node; M, male; MCT, mast cell tumor; Megalic (mild), <2cm; Megalic (moderate), 2-5cm; Megalic (severe), i.e. >5cm; mm, months; ND, not determined; NF, neutered female; R, right; RLN, regional lymph node; SCRA, scraping smear; SNL, sentinel lymph node; yy, years; TIC, touch imprint.

Supplementary Table 2 – Tabulation of cytological diagnoses expressed according to each interpretative system investigated in the current study, with the corresponding histological diagnosis.

Final case ID	T.C.	Krick's system (5-classes)					AM1					AM1.1					AM1.2					Histological diagnosis (Weishaar et al. 2014)[234]
		R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	
#1.1.1	SCRA	5	5	5	3	5	M	M	M	M	M	M	M	M	M	M	M	M	M	NM	M	HN2
#9.1.1	SCRA	2	2	2	2	3	PM	M	M	NM	PM	NM	M	M	NM	PM	NM	M	M	NM	NM	HN2
#9.1.2	TIC	2	4	4	3	1	NM	M	M	PM	NM	NM	M	M	M	NM	NM	M	M	NM	NM	HN2
#9.2.1	SCRA	2	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#9.2.2	TIC	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#9.3.1	SCRA	2	5	4	2	3	NM	M	M	M	M	NM	M	M	M	M	NM	M	M	NM	M	HN0
#18.1.1	SCRA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#18.1.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#18.2.1	SCRA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#18.2.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#18.3.1	SCRA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#18.3.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#19.1.1	TIC	2	2	2	2	2	PM	M	NM	NM	NM	PM	M	NM	PM	NM	NM	NM	NM	NM	NM	HN0
#19.1.2	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#20.1.1	TIC	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#20.2.1	TIC	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#20.3.1	TIC	2	1	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#24.1.1	SCRA	5	4	4	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#24.1.2	TIC	5	4	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#24.2.1	SCRA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#24.2.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#24.3.1	SCRA	3	4	2	4	2	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2

#24.3.2	TIC	2	3	2	2	2	M	M	NM	NM	M	M	M	PM	M	M	M	M	NM	NM	NM	HN2
#25.2.1	FNA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#30.1.1	FNA	2	2	2	2	2	M	PM	PM	PM	PM	M	NM	M	NM	NM	M	NM	NM	NM	NM	HN2
#31.1.1	TIC	4	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#31.2.1	TIC	2	2	3	2	2	M	M	NM	NM	NM	M	M	M	M	NM	M	M	NM	NM	NM	HN2
#34.1.1	SCRA	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#34.1.2	TIC	4	4	4	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#34.1.3	FNA	2	2	3	2	2	M	M	M	NM	M	M	M	M	PM	M	NM	M	NM	NM	NM	HN3
#38.1.1	SCRA	2	5	2	2	4	PM	M	PM	NM	M	NM	M	M	M	M	NM	M	NM	NM	M	HN2
#38.1.2	TIC	2	2	2	2	2	PM	NM	NM	NM	NM	M	NM	NM	PM	NM	NM	NM	NM	NM	NM	HN2
#40.1.1	FNA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#43.1.1	SCRA	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#43.1.2	TIC	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#43.2.1	SCRA	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#43.2.2	TIC	4	5	5	4	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#43.3.1	SCRA	5	5	4	2	4	M	M	M	PM	M	M	M	M	M	M	M	M	M	NM	M	HN2
#43.3.2	TIC	4	5	4	5	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#46.1.1	TIC	2	2	2	2	2	NM	PM	NM	NM	NM	NM	M	NM	NM	NM	NM	M	NM	NM	NM	HN0
#52.1.1	TIC	4	4	4	3	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#53.1.1	TIC	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	NM	HN1
#53.1.2	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	M	NM	NM	NM	NM	NM	HN1
#53.2.1	TIC	2	4	2	2	2	NM	M	NM	NM	NM	NM	M	PM	PM	M	NM	M	NM	NM	NM	HN1
#53.2.2	SCRA	2	4	3	3	2	PM	M	PM	PM	NM	PM	M	M	M	PM	NM	M	NM	NM	NM	HN1
#55.1.1	FNA	2	2	2	2	3	PM	PM	NM	NM	PM	NM	M	NM	PM	M	NM	M	NM	NM	M	HN2
#60.3.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#62.2.1	TIC	2	2	2	2	2	PM	NM	NM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	NM	HN1
#62.3.1	TIC	2	1	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	HN0

#63.1.1	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#63.2.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#67.1.1	SCRA	1	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#67.1.2	TIC	1	1	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#67.2.1	SCRA	2	2	2	2	4	M	M	NM	NM	M	M	M	PM	PM	NM	M	M	NM	NM	NM	HN0
#67.2.2	TIC	4	4	5	2	2	M	M	M	PM	M	M	M	M	M	M	M	M	M	NM	NM	HN0
#73.1.1	SCRA	2	2	2	2	2	NM	PM	NM	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN3
#73.1.2	TIC	4	5	4	5	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#74.1.1	TIC	2	2	2	2	2	NM	M	PM	NM	PM	PM	M	PM	M	PM	NM	NM	NM	NM	NM	HN1
#74.1.2	SCRA	2	4	2	2	2	NM	M	NM	PM	NM	PM	M	PM	NM	PM	NM	M	NM	NM	NM	HN1
#74.2.1	TIC	2	1	1	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#74.2.2	SCRA	2	1	2	1	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#74.3.1	TIC	2	2	2	2	2	PM	NM	NM	NM	NM	M	NM	PM	NM	NM	NM	NM	NM	NM	NM	HN1
#74.3.2	SCRA	2	2	3	2	2	NM	NM	PM	NM	PM	NM	NM	M	PM	M	NM	NM	NM	NM	NM	HN1
#75.1.1	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN0
#76.1.1	SCRA	2	3	2	2	2	NM	M	NM	NM	NM	NM	M	M	PM	PM	NM	M	NM	NM	NM	HN0
#76.1.2	TIC	2	2	2	1	2	PM	PM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#76.2.1	SCRA	2	3	2	2	2	NM	M	PM	NM	NM	NM	M	M	M	PM	NM	M	NM	NM	NM	HN0
#76.2.2	TIC	2	2	2	2	2	PM	M	M	NM	NM	NM	M	M	NM	NM	NM	M	NM	NM	NM	HN0
#78.1.1	TIC	4	4	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN1
#78.2.1	TIC	4	4	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#80.2.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#81.2.1	FNA	2	1	2	1	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN1
#81.2.2	SCRA	2	3	2	2	2	NM	M	PM	NM	PM	NM	M	NM	M	NM	NM	M	NM	NM	NM	HN1
#81.2.3	TIC	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN1
#85.1.1	FNA	4	5	4	3	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#87.1.1	FNA	1	1	2	1	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN1

#92.1.1	SCRA	4	4	4	4	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#92.1.2	TIC	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#92.2.1	SCRA	2	4	2	2	2	NM	M	PM	PM	PM	PM	M	PM	M	M	NM	M	NM	NM	NM	HN0
#92.2.2	TIC	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN0
#92.3.1	SCRA	4	4	4	4	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#92.3.2	TIC	4	4	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#92.4.1	SCRA	2	3	2	2	2	NM	PM	NM	NM	NM	NM	M	PM	PM	PM	NM	NM	NM	NM	NM	HN2
#92.4.2	TIC	2	2	2	2	2	PM	NM	NM	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN2
#92.5.1	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN0
#92.5.2	TIC	2	2	2	2	2	PM	NM	NM	PM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	NM	HN0
#92.6.1	SCRA	2	2	1	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#92.6.2	TIC	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#98.2.1	FNA	2	5	5	5	5	NM	M	M	M	M	NM	M	M	M	M	NM	M	M	M	M	HN3
#99.1.1	FNA	1	1	1	1	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#100.1.1	SCRA	2	1	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	HN2
#100.1.2	TIC	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN2
#102.1.1	SCRA	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN2
#102.1.2	TIC	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	NM	HN2
#103.1.1	FNA	2	2	2	2	3	NM	PM	PM	PM	PM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	HN2
#103.3.1	FNA	4	5	4	4	3	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#103.4.1	FNA	2	2	2	2	2	NM	PM	NM	NM	NM	NM	NM	NM	M	NM	NM	NM	NM	NM	NM	HN0
#107.1.1	TIC	5	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#110.1.1	TIC	2	2	2	2	1	PM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#112.1.1	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN2
#112.1.2	TIC	1	4	2	1	2	NM	M	NM	NM	NM	NM	M	NM	NM	NM	NM	M	NM	NM	NM	HN2
#112.2.1	SCRA	1	2	2	1	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#112.2.2	TIC	2	3	1	1	1	NM	PM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	NM	NM	HN0

#112.3.1	SCRA	1	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#112.3.2	TIC	2	3	2	2	1	NM	PM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#115.1.1	SCRA	3	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#115.1.2	TIC	4	5	4	5	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#115.2.1	SCRA	2	2	2	1	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#115.2.2	TIC	2	2	2	1	1	NM	PM	NM	NM	NM	NM	M	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#121.1.1	TIC	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN2
#121.1.2	SCRA	2	2	2	2	1	NM	PM	NM	NM	NM	NM	NM	NM	PM	NM	NM	NM	NM	NM	NM	HN2
#121.1.3	FNA	2	1	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN2
#121.2.2	SCRA	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN2
#130.1.1	TIC	2	2	2	2	3	NM	NM	NM	PM	PM	NM	NM	PM	PM	PM	NM	NM	NM	NM	NM	HN0
#130.2.1	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN0
#132.1.1	SCRA	2	2	2	2	2	M	M	NM	PM	NM	M	M	PM	M	NM	NM	M	NM	NM	NM	HN0
#132.1.2	TIC	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#132.2.1	SCRA	2	2	2	2	3	PM	M	NM	PM	PM	NM	M	M	M	PM	NM	M	NM	NM	NM	HN0
#132.2.2	TIC	2	1	2	1	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#132.3.1	SCRA	2	2	2	2	2	PM	PM	M	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN0
#132.3.2	TIC	2	2	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#132.4.1	SCRA	2	2	2	2	2	NM	M	NM	PM	NM	NM	M	M	M	NM	NM	M	NM	NM	NM	HN0
#132.5.1	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	PM	PM	NM	NM	NM	NM	NM	NM	HN0
#132.5.2	TIC	1	1	2	1	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#132.6.1	SCRA	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#132.6.2	TIC	2	1	2	2	1	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN0
#134.1.2	FNA	2	2	2	2	2	NM	PM	M	NM	PM	NM	NM	M	M	NM	NM	NM	NM	NM	NM	HN1
#136.3.2	TIC	2	2	2	2	2	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	NM	HN1
#143.1.1	SCRA	5	5	5	4	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3
#143.1.2	TIC	5	5	5	5	5	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN3

#148.1.1	SCRA	4	4	4	4	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#148.1.2	TIC	4	5	4	5	4	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	HN2
#148.2.1	SCRA	4	5	4	3	4	M	M	M	M	M	M	M	M	M	M	M	M	M	NM	M	HN2

Highlighted in yellow, those cases for which the majority of the Readers expressed a cytological diagnosis of metastasis disregarding the interpretative system applied, in spite of a corresponding HN0 histological diagnosis. Legend: AM1, amended system 1; AM1.1, amended system 1.1; AM1.2, amended system 1.2; FNA, fine needle aspirate; HNO, “non-metastatic” according to Weishaar et al.;[234] HN1, “pre-metastatic” according to Weishaar et al.; HN2, “early metastasis” according to Weishaar et al.; HN3, “overt metastasis” according to Weishaar et al.; Krick’s system, cytological interpretative system proposed by Krick et al.[104]; M. metastatic; NM, non metastatic; PM, possibly metastatic; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; SCRA, scraping smear; TIC, touch imprint; 1, “normal”; 2, “reactive lymphoid hyperplasia”; 3, “possible metastas”; 4, “probable metastasis”; 5, “certain metastasis”.

Supplementary Table 3 – Weighted and Cohen's kappa coefficients for pairwise comparisons of interrater agreement among the 5 Readers.

	Weighted K				Cohen's K (95% C.I.)			
Couple investigated	**Krick's system - 5 classes**	**Krick's system - 3 classes**	**AM1**	**AM1.2**	**Krick's system - 5 classes**	**Krick's system - 3 classes**	**AM1**	**AM1.2**
R1 vs R2	0.706	0.778	0.646	n/a	0.527 (0.420-0.633)	0.728 (0.623-0.832)	0.558 (0.447-0.670)	0.690 (0.573-0.807)
R1 vs R3	0.809	0.889	0.726	n/a	0.645 (0.539-0.752)	0.848 (0.761-0.935)	0.635 (0.525-0.746)	0.875 (0.792-0.959)
R1 vs R4	0.782	0.870	0.740	n/a	0.608 (0.504-0.713)	0.816 (0.723-0.909)	0.640 (0.531-0.749)	0.857 (0.767-0.947)
R1 vs R5	0.745	0.849	0.787	n/a	0.521 (0.417-0.625)	0.792 (0.694-0.889)	0.690 (0.586-0.794)	0.859 (0.771-0.948)
R2 vs R3	0.736	0.825	0.693	n/a	0.528 (0.423-0.634)	0.760 (0.0661-0.860)	0.612 (0.504-0.721)	0.719 (0.607-0.832)
R2 vs R4	0.715	0.778	0.624	n/a	0.558 (0.455-0.662)	0.706 (0.602-0.810)	0.537 (0.428-0.646)	0.620 (0.499-0.741)
R2 vs R5	0.681	0.756	0.694	n/a	0.452 (0.345-0.559)	0.681 (0.572-0.790)	0.630 (0.522-0.737)	0.705 (0.591-0.819)
R3 vs R4	0.795	0.873	0.817	n/a	0.624 (0.520-0.729)	0.823 (0.733-0.913)	0.749 (0.650-0.847)	0.889 (0.809-0.968)
R3 vs R5	0.757	0.851	0.863	n/a	0.546 (0.442-0.649)	0.783 (0.686-0.880)	0.827 (0.740-0.914)	0.922 (0.855-0.989)
R4 vs R5	0.742	0.833	0.831	n/a	0.516 (0.410-0.621)	0.753 (0.652-0.854)	0.764 (0.667-0.860)	0.904 (0.829-0.979)

According to previous literature,[126] results of kappa coefficient calculation were interpreted as follows: 0-.20, no agreement; .21-.39, minimal agreeement; .40-.59, weak agreement; .60-.79, moderate agreement; .80-.90, strong agreement; >.90, almost perfect agreement. Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system – 5-classes, original version of the system by Krick et al;[104] Krick's system – 3-classes, 3-classes simplified version of the system by Krick et al; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; vs, versus.

Supplementary Table 4 – Results of chi-squared test applied to investigate significant differences of percentage agreement between groups of cases stratifyied according to the corresponding histological diagnosis.

Reader	Comparison and interpretative system	CA (%)	#cases	CA (%)	#cases	p value
R1	HN0 vs HN1 - Krick's system	93,62	47	0,00	16	<0,0001
	HN0 vs HN2/3 – Krick's system	93,62	47	61,43	70	0,0001
	HN1 vs HN2/3 – Krick's system	0,00	16	61,43	70	<0,0001
	HN0 vs HN1 - AM1	74,47	47	18,75	16	0,0001
	HN0 vs HN2/3 – AM1	74,47	47	71,43	70	0,7188
	HN1 vs HN2/3 – AM1	18,75	16	71,43	70	0,0001
	HN0 vs HN1 - AM1.2	91,49	47	93,75	16	0.7744
	HN0 vs HN2/3 – AM1.2	91,49	47	70,00	70	0.0056
	HN1 vs HN2/3 – AM1.2	93,75	16	70,00	70	0.0508
R2	HN0 vs HN1 - Krick's system	80,85	47	6,25	16	<0,0001
	HN0 vs HN2/3 – Krick's system	80,85	47	71,43	70	0,2493
	HN1 vs HN2/3 – Krick's system	6,25	16	71,43	70	<0,0001
	HN0 vs HN1 - AM1	57,45	47	6,25	16	0,0004
	HN0 vs HN2/3 – AM1	57,45	47	77,14	70	0,0243
	HN1 vs HN2/3 – AM1	6,25	16	77,14	70	<0,0001
	HN0 vs HN1 - AM1.2	68,09	47	75,00	16	0.6058
	HN0 vs HN2/3 – AM1.2	68,09	47	80,00	70	0.1453
	HN1 vs HN2/3 – AM1.2	75,00	16	80,00	70	0.6592
R3	HN0 vs HN1 - Krick's system	91,49	47	12,50	16	<0,0001
	HN0 vs HN2/3 – Krick's system	91,49	47	67,14	70	0,0023
	HN1 vs HN2/3 – Krick's system	12,50	16	67,14	70	0,0001
	HN0 vs HN1 - AM1	82,98	47	25,00	16	<0,0001
	HN0 vs HN2/3 – AM1	82,98	47	71,43	70	0,1529
	HN1 vs HN2/3 – AM1	25,00	16	71,43	70	0,0006

	HN0 vs HN1 - AM1.2	91,49	47	93,75	16	0.7744
	HN0 vs HN2/3 – AM1.2	91,49	47	70,00	70	0.0056
	HN1 vs HN2/3 – AM1.2	93,75	16	70,00	70	0.0508
R4	HN0 vs HN1 - Krick's system	95,74	47	6,25	16	< 0.0001
	HN0 vs HN2/3 – Krick's system	95,74	47	60,00	70	< 0.0001
	HN1 vs HN2/3 – Krick's system	6,25	16	60,00	70	0.0001
	HN0 vs HN1 - AM1	78,72	47	12,5	16	< 0.0001
	HN0 vs HN2/3 – AM1	78,72	47	65,71	70	0.1303
	HN1 vs HN2/3 – AM1	12,5	16	65,71	70	0.0001
	HN0 vs HN1 - AM1.2	95,74	47	93,75	16	0.7488
	HN0 vs HN2/3 – AM1.2	95,74	47	62,86	70	< 0.0001
	HN1 vs HN2/3 – AM1.2	93,75	16	62,86	70	0.0169
R5	HN0 vs HN1 - Krick's system	87,23	47	0,00	16	< 0.0001
	HN0 vs HN2/3 – Krick's system	87,23	47	65,71	70	0.0093
	HN1 vs HN2/3 – Krick's system	0,00	16	65,71	70	< 0.0001
	HN0 vs HN1 - AM1	82,98	47	25,00	16	< 0.0001
	HN0 vs HN2/3 – AM1	82,98	47	71,43	70	0.1529
	HN1 vs HN2/3 – AM1	25,00	16	71,43	70	0.0006
	HN0 vs HN1 - AM1.2	93,62	47	93,75	16	0.9854
	HN0 vs HN2/3 – AM1.2	93,62	47	70,00	70	0.0020
	HN1 vs HN2/3 – AM1.2	93,75	16	70,00	70	0.0508
Median	HN0 vs HN1 - Krick's system	91,49	47	6,25	16	< 0.0001
	HN0 vs HN2/3 – Krick's system	91,49	47	65,71	70	0.0014
	HN1 vs HN2/3 – Krick's system	6,25	16	65,71	70	< 0.0001
	HN0 vs HN1 - AM1	78,72	47	18,75	16	< 0.0001
	HN0 vs HN2/3 – AM1	78,72	47	71,43	70	0.3780
	HN1 vs HN2/3 – AM1	18,75	16	71,43	70	0.0001

	HN0 vs HN1 - AM1.2	91,49	47	93,75	16	0.7744
	HN0 vs HN2/3 – AM1.2	91,49	47	70,00	70	0.0056
	HN1 vs HN2/3 – AM1.2	93,75	16	70,00	70	0.0508

Values of CA are expressed in percentages. Highlighted in grey, p values <.05. Legend: AM1, amended system 1; AM1.2, amended system 1.2; CA, complete agreement (i.e. proprotion of “non metastatic”, “possibly metastatic”, or “metastatic” diagnoses respectively corresponding to HN0, HN1, or HN2/3 diagnosis); HNO, “non-metastatic” according to Weishaar et al.;[234] HN1, “pre-metastatic” according to Weishaar et al.; HN2, “early metastasis” according to Weishaar et al.; HN3, “overt metastasis” according to Weishaar et al.; Krick’s system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1;R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

Supplementary Table 5 – Results of chi-squared test applied to investigate significant differences of percentage agreement between groups of cases stratified according to the corresponding cytological sampling technique.

Reader	Comparison and interpretative system	CA (%)	#cases	CA (%)	#cases	p value
R1	FNA vs TIC – Krick's system	40,00	15	69,7	66	0.0314
	FNA vs SCRA – Krick's system	40,00	15	67,31	52	0.0577
	TIC vs SCRA – Krick's system	69,7	66	67,31	52	0.7820
	FNA vs TIC – AM1	53,33	15	68,18	66	0.2780
	FNA vs SCRA – AM1	53,33	15	67,31	52	0.3235
	TIC vs SCRA – AM1	68,18	66	67,31	52	0.9204
	FNA vs TIC – AM1.2	66,67	15	83,33	66	0.1460
	FNA vs SCRA – AM1.2	66,67	15	80,77	52	0.2520
	TIC vs SCRA – AM1.2	83,33	66	80,77	52	0.7193
R2	FNA vs TIC – Krick's system	46,67	15	69,7	66	0.0925
	FNA vs SCRA – Krick's system	46,67	15	69,23	52	0.1111
	TIC vs SCRA – Krick's system	69,7	66	69,23	52	0.9563
	FNA vs TIC – AM1	53,33	15	65,15	66	0.3951
	FNA vs SCRA – AM1	53,33	15	59,62	52	0.6658
	TIC vs SCRA – AM1	65,15	66	59,62	52	0.5391
	FNA vs TIC – AM1.2	80,00	15	81,82	66	0.8707
	FNA vs SCRA – AM1.2	80,00	15	65,38	52	0.2859
	TIC vs SCRA – AM1.2	81,82	66	65,38	52	0.0426
R3	FNA vs TIC – Krick's system	46,67	15	71,21	66	0.0705
	FNA vs SCRA – Krick's system	46,67	15	73,08	52	0.0568
	TIC vs SCRA – Krick's system	71,21	66	73,08	52	0.8230
	FNA vs TIC – AM1	53,33	15	71,21	66	0.1833
	FNA vs SCRA – AM1	53,33	15	73,08	52	0,1494
	TIC vs SCRA – AM1	71,21	66	73,08	52	0,8230

	FNA vs TIC – AM1.2	66,67	15	81,82	66	0.1961
	FNA vs SCRA – AM1.2	66,67	15	82,69	52	0.1821
	TIC vs SCRA – AM1.2	81,82	66	82,69	52	0.9028
R4	FNA vs TIC – Krick's system	40,00	15	69,7	66	0.0314
	FNA vs SCRA – Krick's system	40,00	15	69,23	52	0.0407
	TIC vs SCRA – Krick's system	69,7	66	69,23	52	0.9563
	FNA vs TIC – AM1	46,67	15	66,67	66	0.1502
	FNA vs SCRA – AM1	46,67	15	65,38	52	0.1935
	TIC vs SCRA – AM1	66,67	66	65,38	52	0.8836
	FNA vs TIC – AM1.2	66,67	15	81,82	66	0.1961
	FNA vs SCRA – AM1.2	66,67	15	76,92	52	0.4251
	TIC vs SCRA – AM1.2	81,82	66	76,92	52	0.5133
R5	FNA vs TIC – Krick's system	40,00	15	69,7	66	0.0314
	FNA vs SCRA – Krick's system	40,00	15	67,31	52	0.0577
	TIC vs SCRA – Krick's system	69,7	66	67,31	52	0.7820
	FNA vs TIC – AM1	60,00	15	71,21	66	0.3992
	FNA vs SCRA – AM1	60,00	15	71,15	52	0.4157
	TIC vs SCRA – AM1	71,21	66	71,15	52	0.9943
	FNA vs TIC – AM1.2	73,33	15	81,82	66	0.4588
	FNA vs SCRA – AM1.2	73,33	15	82,69	52	0.4228
	TIC vs SCRA – AM1.2	81,82	66	82,69	52	0.9028
Median	FNA vs TIC – Krick's system	40,00	15	69,70	66	0.0314
	FNA vs SCRA – Krick's system	40,00	15	69,23	52	0.0407
	TIC vs SCRA – Krick's system	69,70	66	69,23	52	0.9563
	FNA vs TIC – AM1	53,33	15	68,18	66	0.2780
	FNA vs SCRA – AM1	53,33	15	67,31	52	0.3235
	TIC vs SCRA – AM1	68,18	66	67,31	52	0.9204

	FNA vs TIC – AM1.2	66,67	15	81,82	66	0.1961
	FNA vs SCRA – AM1.2	66,67	15	80,77	52	0.2520
	TIC vs SCRA – AM1.2	81,82	66	80,77	52	0.8849

Values of CA are expressed in percentages. Highlighted in grey, p values <.05. Legend: AM1, amended system 1; AM1.2, amended system 1.2; CA, complete agreement (i.e. proprotion of “non metastatic”, “possibly metastatic”, or “metastatic” diagnoses respectively corresponding to HN0, HN1, or HN2/3 diagnosis); FNA, fine needel aspirates; Krick’s system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1;R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; SCRA, scraping smears; TIC, touch imprints.

Supplementary Table 6 - Concordance between the cytological diagnoses of the cases referring to the same LN but sampled with different techniques.

			R1		R2		R3		R4		R5		Median	
Krick's system	Complete intersampling coherence	Right diagnoses	86,96	58,70	76,09	56,52	84,78	63,04	86,96	63,04	86,96	60,87	86,67	62,22
		Wrong diagnoses		28,26		19,57		21,74		23,91		26,09		24,44
	Partial intersampling coherence	2 right diagnoses	2,17	2,17	4,35	2,17	2,17	2,17	2,17	2,17	2,17	2,17	2,22	2,22
		1 right diagnoses		0,00		2,17		0,00		0,00		0,00		0,00
	Intersampling incoherence		10,87		19,57		13,04		10,87		10,87		11,11	
AM1	Complete intersampling coherence	Right diagnoses	76,09	52,17	78,26	50,00	82,61	63,04	78,26	54,35	86,96	65,22	80,00	55,56
		Wrong diagnoses		23,91		28,26		19,57		23,91		21,74		24,44
	Partial intersampling coherence	2 right diagnoses	0,00	0,00	0,00	0,00	2,17	0,00	2,17	2,17	2,17	0,00	0,00	0,00
		1 right diagnoses		0,00		0,00		2,17		0,00		2,17		0,00
	Intersampling incoherence		23,91		21,74		15,22		19,57		10,87		20,00	
AM1.2	Complete intersampling coherence	Right diagnoses	93,48	76,09	82,61	65,22	89,13	76,09	89,13	73,91	89,13	76,09	89,36	74,47
		Wrong diagnoses		17,39		17,39		13,04		15,22		13,04		14,89
	Partial intersampling coherence	2 right diagnoses	2,17	2,17	2,17	2,17	2,17	2,17	2,17	2,17	2,17	2,17	2,13	2,13
		1 right diagnoses		0,00		0,00		0,00		0,00		0,00		0,00
	Intersampling incoherence		4,35		15,22		8,70		8,70		8,70		8,51	

Values are expressed in percentages. Complete intersampling coherence was defined when the diagnoses for all the 2 or 3 sampling techniques were in agreement among each other, and was further distinguished on the basis of the correctness of the diagnoses compared to the histological reference standard. Partial intersampling coherence was defined exclusively for those cases sampled with all the 3 different techniques, and was further distinguished on the basis of the number of techniques which were in

agreement with histology (i.e. 2 or 1 out of 3). Intersampling inchoerence could have been observed exclusively for those cases sampled with 2 different techniques, among which one diagnosis was right and the other was wrong. Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

Supplementary Table 7 - Weighted and Cohen's kappa coefficients for the evaluation of the agreement between each cytological interpretative system investigated and histopathology.

Evaluator	Interpretative system	Weighted K	Cohen's K (95% C.I.)
R1	Krick's system	0.468	0.413 (0.299-0.526)
	AM1	0.508	0.436 (0.316-0.556)
	AM1.2	n/a	0.613 (0.483-0.743)
R2	Krick's system	0.498	0.432 (0.310-0.553)
	AM1	0.416	0.335 (0.209-0.460)
	AM1.2	n/a	0.501 (0.354-0.648)
R3	Krick's system	0.518	0.474 (0.356-0.592)
	AM1	0.527	0.488 (0.365-0.610)
	AM1.2	n/a	0.613 (0.483-0.743)
R4	Krick's system	0.495	0.435 (0.323-0.547)
	AM1	0.474	0.403 (0.285-0.521)
	AM1.2	n/a	0.571 (0.440-0.701)
R5	Krick's system	0.495	0.418 (0.303-0.532)
	AM1	0.543	0.493 (0.373-0.613)
	AM1.2	n/a	0.628 (0.501-0.755)
Median	Krick's system	0.520	0.467 (0.351-0.583)
	AM1	0.514	0.448 (0.328-0.569)

	AM1.2	n/a	0.613 (0.483-0.743)

According to previous literature,[126] results of kappa coefficient calculation were interpreted as follows: 0-.20, no agreement; .21-.39, minimal agreement; .40-.59, weak agreement; .60-.79, moderate agreement; .80-.90, strong agreement; >.90, almost perfect agreement. Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system, 3-classes simplified version of the system by Krick et al;[104] R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; vs, versus.

Supplementary Table 8 – Number of cases classified as true positive, false positive, false negative, and true negative applying each of the cytological interpretative system investigated in the current study.

Evaluator	Interpretative system	Best case scenario (95% C.I.)		Worst case scenario (95% C.I.)	
		TP	FP	TP	FP
		FN	TN	FN	TN
R1	Krick's system	43	4	45	4
		27	59	25	59
	AM1	50	6	55	16
		20	57	15	47
	AM1.1	51	7	51	12
		19	56	19	51
	AM1.2	49		5	
		21		58	
R2	Krick's system	50	9	52	14
		23	51	21	46
	AM1	54	19	60	27
		16	44	10	36
	AM1.1	56	21	56	23
		14	42	14	40
	AM1.2	55		18	
		15		45	
R3	Krick's system	47	5	49	7
		23	58	21	56
	AM1	50	8	53	14
		20	55	17	49
	AM1.1	53	13	59	28
		17	50	11	35
	AM1.2	49		5	
		21		58	
R4	Krick's system	42	3	47	4
		28	60	23	59
	AM1	46	4	50	13
		24	59	20	50
	AM1.1	51	15	60	26
		19	48	10	37
	AM1.2	44		3	
		26		60	

R5	**Krick's system**	46	4	50	7
		24	59	20	56
	AM1	50	6	54	13
		20	57	16	50
	AM1.1	51	9	54	16
		19	54	16	47
	AM1.2	49		4	
		21		59	
Median	**Krick's system**	46	4	49	7
		24	59	21	56
	AM1	50	6	54	14
		20	57	16	49
	AM1.1	51	13	56	23
		19	50	14	40
	AM1.2	49		5	
		21		58	

Legend: AM1, amended system 1; AM1.2, amended system 1.2; best case scenario, “possibly metastatic” cases considered as negative for metastasis; FN, false negative; FP, false positive; Krick’s system, 3-classes simplified version of the system by Krick et al.;[104] R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; TN, true negative; TP, true positive; worst case scenario, “possibly metastatic” cases considered as positive for metastasis.

Supplementary Table 9 - Results of McNemar's test applied to verify the presence of a test bias for each of the cytological interpretative system investigated in the current study.

Evaluator	Interpretative system	Best case scenario	Worst case scenario
R1	Krick's system	<0.0001	0.0002
	AM1	0.0108	1.000
	AM1.1	0.0310	0.2812
	AM1.2	0.0033	
R2	Krick's system	0.0216	0.3105
	AM1	0.7353	0.0085
	AM1.1	0.3105	0.1885
	AM1.2	0.7277	
R3	Krick's system	0.0013	0.0140
	AM1	0.0376	0.7194
	AM1.1	0.5839	0.0104
	AM1.2	0.0033	
R4	Krick's system	<0.0001	0.0005
	AM1	0.0003	0.2963
	AM1.1	0.6069	0.0124
	AM1.2	<0.0001	
R5	Krick's system	0.0003	0.0209
	AM1	0.0108	0.7103
	AM1.1	0.0890	0.8597
	AM1.2	0.0014	
Median	Krick's system	0.0003	0.0140
	AM1	0.0108	0.8551
	AM1.1	0.3768	0.1885
	AM1.2	0.0033	

Highlighted in grey, p values >.05. Legend: AM1, amended system 1; AM1.1, amended system 1.1; AM1.2, amended system 1.2; best case scenario, "possibly metastatic" cases considered as negative for metastasis; Krick's system, 3-classes simplified version of the system by Krick et al;[104] R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5; worst case scenario, "possibly metastatic" cases considered as positive for metastasis.

Supplementary Table 10 - . Diagnostic accuracy indexes and 95% confidence intervals for all the cytological interpretative systems investigated in the current study.

Evaluator	Interpretative system	ACCURACY		SENSITIVITY		SPECIFICITY		POSITIVE PREDICTIVE VALUE		NEGATIVE PREDICTIVE VALUE	
		Best case scenario (95% C.I.)	Worst case scenario (95% C.I.)	Best case scenario (95% C.I.)	Worst case scenario (95% C.I.)	Best case scenario (95% C.I.)	Worst case scenario (95% C.I.)	Best case scenario (95% C.I.)	Worst case scenario (95% C.I.)	Best case scenario (95% C.I.)	Worst case scenario (95% C.I.)
R1	Krick's system	76,69 (+) (68,58-83,58)	78,2 (+) (70,21-84,88)	61,43 (-) (49,03-72,83)	64,29 (-) (51,93-75,39)	**93,65 (+++) (84,53-98,24)**	**93,65 (+++) (84,53-98,24)**	**91,49 (+++) (80,35-96,58)**	**91,84 (+++) (81,09-96,72)**	68,6 (-) (61,75-74,73)	70,24 (+) (63,13-76,49)
	AM1	**80,45 (++) (72,68-86,81)**	76,69 (+) (68,58-83,58)	71,43 (+) (59,38-81,60)	**78,57 (+) (67,13-87,48)**	90,48 (+++) (80,41-96,42)	74,6 (+) (62,06-84,73)	89,29 (++) (79,33-94,76)	77,46 (+) (68,87-84,23)	74,03 (+) (66,11-80,63)	**75,81 (+) (66,17-83,39)**
	AM1.1	**80,45 (++) (72,68-86,81)**	76,69 (+) (68,58-83,58)	**72,86 (+) (60,90-82,80)**	72,86 (+) (60,90-82,80)	88,89 (++) (78,44-95,41)	80,95 (++) (69,09-89,75)	87,93 (++) (78,13-93,70)	80,95 (++) (71,47-87,82)	**74,67 (+) (66,54-81,37)**	72,86 (+) (64,23-80,05)
	AM1.2	**80,45 (++) (72,68-86,81)**		70 (+) (57,87-80,38)		92,06 (+++) (82,44-97,37)		90,74 (+++) (80,65-95,84)		73,42 (+) (65,72-79,92)	
R2	Krick's system	**75,94 (+) (67,77-82,92)**	73,68 (+) (65,35-80,94)	68,49 (-) (56,56-78,87)	71,23 (+) (59,45-81,23)	**85 (++) (73,43-92,90)**	**76,67 (+) (63,96-86,62)**	**84,75 (++) (74,89-91,19)**	**78,79 (+) (69,65-85,74)**	68,92 (-) (60,87-75,97)	68,66 (-) (59,80-76,34)
	AM1	73,68 (+) (65,35-80,94)	72,18 (+) (63,75-79,60)	77,14 (+) (65,55-86,33)	**85,71 (++) (75,29-92,93)**	69,84 (-) (56,98-80,77)	57,14 (-) (44,05-69,54)	73,97 (+) (65,65-80,87)	68,97 (-) (62,19-75,01)	73,33 (+) (63,45-81,33)	**78,26 (+) (66,12-86,91)**
	AM1.1	73,68 (+) (65,35-80,94)	72,18 (+) (63,75-79,60)	**80 (++) (68,73-88,61)**	80 (++) (68,73-88,61)	66,67 (-) (53,66-78,05)	63,49 (-) (50,40-75,27)	72,73 (+) (64,85-79,40)	70,89 (+) (63,27-77,49)	**75 (+) (64,53-83,18)**	74,07 (+) (63,30-82,55)
	AM1.2	**75,19 (+) (66,96-82,26)**		78,57 (+) (67,13-87,48)		71,43 (+) (58,65-82,11)		75,34 (+) (66,99-82,14)		**75 (+) (65,10-82,83)**	
R3	Krick's system	78,95 (+) (71,03-85,53)	78,95 (+) (71,03-85,53)	67,14 (-) (54,88-77,91)	70 (+) (57,87-80,38)	**92,06 (+++) (82,44-97,37)**	88,89 (++) (78,44-95,41)	90,38 (+++) (79,96-95,68)	87,5 (++) (77,40-93,47)	71,6 (+) (64,16-78,03)	72,73 (+) (64,85-79,40)

Evaluator	Interpretative system	ACCURACY		SENSITIVITY		SPECIFICITY		POSITIVE PREDICTIVE VALUE		NEGATIVE PREDICTIVE VALUE	
		Best case scenario (95% C.I.)	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**
	AM1	78,95 (+) (71,03-85,53)	76,69 (+) (68,58-83,58)	71,43 (+) (59,38-81,60)	75,71 (+) (63,99-85,17)	87,3 (++) (76,50-94,35)	77,78 (+) (65,54-87,28)	86,21 (++) (76,29-92,39)	79,1 (+) (70,07-85,96)	73,33 (+) (65,24-80,12)	74,24 (+) (65,12-81,65)
	AM1.1	77,44 (+) (69,39-84,23)	70,68 (+) (62,16-78,25)	**75,71 (+) (63,99-85,17)**	**84,29 (++) (73,62-91,89)**	79,37 (+) (67,30-88,53)	55,56 (-) (42,49-68,08)	80,3 (++) (71,16-87,07)	67,82 (-) (61,09-73,87)	**74,63 (+) (65,62-81,92)**	**76,09 (+) (63,91-85,11)**
	AM1.2	**80,45 (++) (72,68-86,81)**		70 (+) (57,87-80,38)		**92,06 (+++) (82,44-97,37)**		**90,74 (+++) (80,65-95,84)**		73,42 (+) (65,72-79,92)	
R4	**Krick's system**	76,69 (+) (68,58-83,58)	**79,7 (+) (71,86-86,17)**	60 (-) (47,59-71,53)	67,14 (-) (54,88-77,91)	**95,24 (+++) (86,71-99,01)**	93,65 (+++) (84,53-98,24)	93,33 (+++) (82,03-97,72)	92,16 (+++) (81,78-96,85)	68,18 (-) (61,54-74,16)	71,95 (+) (64,59-78,30)
	AM1	**78,95 (+) (71,03-85,53)**	75,19 (+) (66,96-82,26)	65,71 (-) (53,40-76,65)	71,43 (+) (59,38-81,60)	93,65 (+++) (84,53-98,24)	79,37 (+) (67,30-88,53)	92 (+++) (81,44-96,79)	79,37 (+) (69,86-86,45)	71,08 (+) (63,85-77,38)	71,43 (+) (62,83-78,71)
	AM1.1	74,44 (+) (66,15-81,60)	72,93 (+) (64,55-80,27)	**72,86 (+) (60,90-82,80)**	**85,71 (++) (75,29-92,93)**	76,19 (+) (63,79-86,02)	58,73 (-) (45,62-70,99)	77,27 (+) (68,12-84,40)	69,77 (-) (62,87-75,88)	**71,64 (+) (62,69-79,16)**	**78,72 (+) (66,78-87,20)**
	AM1.2	78,2 (+) (70,21-84,88)		62,86 (-) (50,48-74,11)		**95,24 (+++) (86,71-99,01)**		**93,62 (+++) (82,73-97,82)**		69,77 (-) (62,87-75,88)	
R5	**Krick's system**	78,95 (+) (71,03-85,53)	79,7 (+) (71,86-86,17)	65,71 (-) (53,40-76,65)	71,43 (+) (59,38-81,60)	**93,65 (+++) (84,53-98,24)**	88,89 (++) (78,44-95,41)	92 (+++) (81,44-96,79)	87,72 (++) (77,77-93,58)	71,08 (+) (63,85-77,38)	73,68 (+) (65,68-80,38)
	AM1	80,45 (++) (72,68-86,81)	78,2 (+) (70,21-84,88)	71,43 (+) (59,38-81,60)	**77,14 (+) (65,55-86,33)**	90,48 (+++) (80,41-96,42)	79,37 (+) (67,30-88,53)	89,29 (++) (79,33-94,76)	80,6 (++) (71,57-87,27)	**74,03 (+) (66,11-80,63)**	**75,76 (+) (66,62-83,03)**
	AM1.1	78,95 (+) (71,03-85,53)	75,94 (+) (67,77-82,92)	**72,86 (+) (60,90-82,80)**	**77,14 (+) (65,55-86,33)**	85,71 (++) (74,61-93,25)	74,6 (+) (62,06-84,73)	85 (++) (75,27-91,34)	77,14 (+) (68,45-84,00)	73,97 (+) (65,65-80,87)	74,6 (+) (65,11-82,22)
	AM1.2	**81,2 (++) (73,52-87,45)**		70 (+) (57,87-80,38)		**93,65 (+++) (84,53-98,24)**		**92,45 (+++) (82,42-96,97)**		73,75 (+) (66,14-80,16)	
Median	**Krick's system**	78,95 (+)	78,95 (+)	65,71 (-)	70,00 (+)	**93,65 (+++)**	88,89 (++)	**92,00 (+++)**	87,50 (++)	71,08 (+)	72,73 (+)

Evaluator	Interpretative system	ACCURACY		SENSITIVITY		SPECIFICITY		POSITIVE PREDICTIVE VALUE		NEGATIVE PREDICTIVE VALUE	
		Best case scenario (95% C.I.)	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**	**Best case scenario (95% C.I.)**	**Worst case scenario (95% C.I.)**
		(71,03-85,53)	(71,03-85,53)	(53,40-76,65)	(57,87-80,38)	**(84,53-98,24)**	(78,44-95,41)	**(81,44-96,79)**	(77,40-93,47)	(63,85-77,38)	(64,85-79,40)
	AM1	**80,45 (++)** **(72,68-86,81)**	77,44 (+) (69,39-84,23)	71,43 (+) (59,38-81,60)	77,14 (+) (65,55-86,33)	90,48 (+++) (80,41-96,42)	77,78 (+) (65,54-87,28)	89,29 (++) (79,33-94,76)	79,41 (+) (70,49-86,17)	**74,03 (+)** **(66,11-80,63)**	**75,38 (+)** **(66,13-82,77)**
	AM1.1	75,94 (+) (67,77-82,92)	72,18 (+) (63,75-79,60)	**72,86 (+)** **(60,90-82,80)**	**80 (++)** **(68,73-88,61)**	79,37 (+) (67,30-88,53)	63,49 (-) (50,40-75,27)	79,69 (+) (70,31-86,67)	70,89 (+) (63,27-77,49)	72,46 (+) (63,73-79,76)	74,07 (+) (63,30-82,55)
	AM1.2	**80,45 (++)** **(72,68-86,81)**		70 (+) (57,87-80,38)		**92,06 (+++)** **(82,44-97,37)**		**90,74 (+++)** **(80,65-95,84)**		73,42 (+) (65,72-79,92)	

Values are expressed in percentages. Highlighted in red, values judged as low compared to the Krick's system, or which interpretation class was reduced twice compared to the Krick's system. Interpretation classes: (-), low; (+), moderate; (++), high; (+++), very high. Legend: AM1, amended system 1; AM1.1, amended system 1.1; AM1.2, amended system 1.2; Krick's system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

Supplementary Table 11 - Concordance between the cytological diagnoses expressed according to the 3-classes simplified version of the system by Krick et al.,[104] the AM1 system, and the AM1.2 system.

	R1	R2	R3	R4	R5	Median
Intersystem agreement - right diagnoses	58,65	57,89	66,17	60,15	63,91	61,07
Intersystem agreement - wrong diagnoses	18,05	19,55	18,80	20,30	17,29	19,08
Intersystem disagreement	23,31	22,56	15,04	19,55	18,80	19,85
Krick's system - wrong diagnosis	29,03	16,67	20,00	11,54	24,00	20,83
AM1 - wrong diagnosis	25,81	16,67	0,20	30,77	8,00	20,83
AM1.2 - wrong diagnosis	0,00	0,00	0,00	0,00	0,00	0,00
Krick's system & AM1 - wrong diagnoses	38,71	43,33	55,00	50,00	60,00	54,17
Krick's system & AM1.2 - wrong diagnoses	3,26	0,00	5,00	7,69	8,00	4,17
AM1 & AM1.2 - wrong diagnoses	3,23	23,33	0,00	0,00	0,00	0,00

Values are expressed in percentages. Intersystem agreement was defined when the diagnoses for all the investigated cytological interpretative systems were in agreement among each other, and was further distinguished on the basis of the correctness of the diagnoses compared to the histological reference standard. Intersystem disagreement was defined when the diagnoses of 1 or 2 of the investigated cytological interpretative systems were in disagreement with the other(s) diagnosis/es, and was further distinguished on the basis of which interpretative system(s) expressed the wrong diagnosis/es compared to the histological reference standard. Legend: AM1, amended system 1; AM1.2, amended system 1.2; Krick's system, 3-classes simplified version of the system by Krick et al.[104] as illustrated in Appendix 2 – Supplementary Table 1; R1, Reader1 ; R2, Reader 2; R3, Reader 3; R4, Reader 4; R5; Reader 5.

ADDENDUM - VIRAL ONCOLYSIS IN A CELL CULTURE MODEL OF CANINE HISTIOCYTIC SARCOMA

SCIENTIFIC PRODUCTION OF THE PH.D. CANDIDATE IN THE FIELD

The 2 studies described in this Section were performed during the externship of the Ph.D. candidate at the University of Veterinary Medicine of Hannover (TiHo Hannover, Germany). Both studies have been published in peer-reviewed journals, with the Ph.D. candidate as one of the 2 co-first Authors:

- Armando F*, Gambini M*, Corradi A, Giudice C, Pfankuche VM, Brogden G, Attig F, von Köckritz-Blickwede M, Baumgärtner W, Puff C. Oxidative Stress in Canine Histiocytic Sarcoma Cells Induced by an Infection with Canine Distemper Virus Led to a Dysregulation of HIF-1α Downstream Pathway Resulting in a Reduced Expression of VEGF-B in vitro. Viruses. 2020 Feb 11;12(2):200. doi: 10.3390/v12020200. PMID: 32054075; PMCID: PMC7077254. (available in Open Access at the following link: https://www.mdpi.com/1999-4915/12/2/200/htm)
- Armando F*, Gambini M*, Corradi A, Becker K, Marek K, Pfankuche VM, Mergani AE, Brogden G, de Buhr N, von Köckritz-Blickwede M, Naim HY, Baumgärtner W, Puff C. Mesenchymal to epithelial transition driven by canine distemper virus infection of canine histiocytic sarcoma cells contributes to a reduced cell motility in vitro. J Cell Mol Med. 2020 Aug;24(16):9332-9348. doi: 10.1111/jcmm.15585. Epub 2020 Jul 6. PMID: 32627957; PMCID: PMC7417708. (available in Open Access at the following link: https://onlinelibrary.wiley.com/doi/full/10.1111/jcmm.15585)

Considered that the 2 studies are the main focus of investigation of the Ph.D. Thesis of Dr. Federico Armando from the Department of Veterinary Medicine Sciences of the University of Parma (Parma, Italy), only a brief summary of the main findings is reported in the current Chapter.

INTRODUCTION

Viral oncolysis represents an interesting potential new therapeutic option in both human and veterinary medicine.[8,41,62] Indeed, viruses from different families, including members of the Paramyxoviridae (canine distemper virus, measles virus, and Newcastle disease virus), have been reported as possessing oncolytic properties.[8,28,38,48] Canine distemper virus (CDV) is a Morbillivirus closely related to human measles virus,[67] with the latter already described as a promising oncolytic virus in human medicine.[10,18] Similarly, the attenuated Onderstepoort vaccine strain of canine distemper virus (CDV-Ond) represents a potential oncolytic virus for the treatment of canine histiocytic sarcomas.[26,50]

Histiocytic sarcomas (HS) are malignant tumors with poor prognosis and limited therapeutic options in both humans and dogs,[1,3,11,58,70] which originate from interstitial dendritic cells or from macrophages.[21,29,46,59] According to the higher prevalence of HS in dogs than in human beings, the canine species may represent an interesting translational model for this neoplastic disease.[11] Since its establishment,[73] a canine histiocytic sarcoma cell line (DH82 cells) has been commercially available. DH82 cells can be infected by CDV-Ond,[26] and have been reported as a promising model for the investigation of viral oncolysis.[38,49,52] Specifically, acute infection of DH82 cells with CDV-Ond in vitro resulted in a prominent cell death at 12 days post infection,[26] followed by the establishment of persistent infection in tumor cells surviving the acute lytic phase.[50] In this context, subcutaneous xenotransplantion of persistently CDV-Ond infected DH82 cells resulted in a total regression of the neoplasms in a mouse model.[50] This promising observation was assumed to be related to a decreased vascularization of the transplants,[50] with the underlying mechanisms not fully understood so far. Therefore, additional investigations using persistently CDV-infected DH82 cells might represent a promising model to study virus-induced alterations of cancer hallmarks [27] and of the tumor microenvironment [75] avoiding the confounding effects correlated with ongoing virus-induced cytopathogenic tumor cell death associated with the acute infection.[26] Indeed, as reviewed by Lapp et al.,[38] viral oncolysis mechanisms can be distinguished between primary (i.e. direct virus-induced cytolysis and/or apoptosis) and secondary ones (also defined as “indirect oncolysis”). The latter include a wide range of events leading to tumor cell death,

such as modulation of the antiviral and antitumoral immune response, changes in the organization of the tumor-associated extracellular matrix, and alterations of the tumor-associated vasculature and angiogenesis.[8,24,33,38,41,48,50,56,75]

In addition to the inderect oncolytic effects hypothesized above, persistent infection with CDV might induce direct alterations in DH82 cells mRNA and protein expression, further reducing malignancy potential of neoplastic cells and favouring xenotransplants regression. In this context, it has beeen reported that persistently CDV-infected DH82 cells are characterized by an altered expression of reversion-inducing cysteine-rich protein with Kazal motifs (RECK), matrix metalloproteinases (MMP) −2 and −9, and tissue inhibitors of matrix metalloproteinases (TIMP) −1 and −2,[52] and by an altered distribution of cortactin within the cytoskeleton.[49] Therefore, further investigations are warranted in this direction, such as studies focusing on mesenchymal to epithelial transition (MET).[76]

STUDY 1 - Oxidative stress in canine histiocytic sarcoma cells induced by an infection with Canine Distemper Virus led to a dysregulation of HIF-1A downstream pathway resulting in a reduced expression of VEGF-B in vitro

Considered that persistently CDV-infected DH82 cells (DH82Ond pi) did not show any difference in growth and apoptotic rate compared to non-infected controls in vitro and during the initial phase after transplantation into Scid mice,[49,50,52] it was assumed that tumor regression of DH82Ond pi xenotransplants was not caused primarily by direct virus-induced cell death alone. Indeed, it seems more likely that secondary effects of the viral infection on the tumor microenvironment,[38,75] as similarly reported for Reoviruses,[33] account for the complete regression. Specifically, it was estimated that regression of DH82Ond pi xenotransplants might be related to alterations in cancer-associated angiogenesis,[50] which cover a key role among indirect onclytic effects induced by viral infections.[8,38,41,48,50,75] A reduced vascularization of neoplasms often leads to intratumoral hypoxia [16] which in turn is associated with modifications of intracellular pathways connected with reactive oxygen species (ROS) production and scavenging. CDV infection can increase ROS production and ROS-induced damage in vitro and in vivo as shown for spontaneous CDV infection in dogs.[5,14,25,32,63] Furthermore, CDV can induce an accumulation of viral glycoproteins in the endoplasmic reticulum (ER) of Vero cells and primary rat neurons, resulting in increased endoplasmic reticulum stress,[12] which has been reported as associated with an increased ROS productio.[9] Besides these considerations, it should be remembered that ROS are physiologically involved in a plethora of different intracellular signaling pathways,[45,57] and play a key role in multiple hallmarks of cancer.[20]

Hypoxia-inducible factor 1-alpha (HIF-1α) is a transcription factor involved in the regulation of a wide plethora of cancer features such as invasion, metastasis, and angiogenesis.[20,27,47,61,71,78] After translocation from the cytoplasm to the nucleus, HIF-1α forms a heterodimer with hypoxia-inducible factor 1-beta (HIF-1α), which binds to specific DNA sequences known as hypoxia response elements (HREs).[61,78] This event induces the expression of numerous genes involved in different cellular

responses such as angiogenesis,[71,78] which is driven by several growth factors, including members of the vascular endothelial growth factor (VEGF) family. Together with hypoxia, ROS represent the most important stimuli for HIF-1α stabilization and nuclear translocation.[23,35,47,61,78] During normoxia and redox homeostatic state,[51] HIF-1α is localized within the cytoplasm and is rapidly degraded by the proteasome after hydroxylation by prolyl hydroxylases (PHDs) and subsequent ubiquitination by the von Hippel-Lindau protein (VHL).[35,61,78] In this context, hypoxia and ROS directly down-regulate the activity of PHDs and VHL, playing therefore a key role in the inhibition of the overall HIF-1α degradation.[35,61,78]

Considered the above, the aim of the current study was to investigate the hypothesis that a persistent CDV-Ond infection of DH82 cells induces oxidative stress followed by a massive inhibition of HIF-1α degrading pathways. This in turn leads to cytoplasmic, non-functional accumulation of HIF-1α, which is associated with a reduced expression of HIF-1α downstream targets, such as VEGF-B.

The study was conducted *in vitro* on persistently CDV-infected DH82 cells (DH82Ond pi) by means of microarray data analysis, single and double immunofluorescence further analysed with laser scanning confocal microscopy, flow cytometry, immunoelectron microscopy, and immunoblotting. Further details are provided in the Open Access version of the published paper at the following link: https://www.mdpi.com/1999-4915/12/2/200/htm.

Microarray data analysis, immunofluorescence for 8-hydroxyguanosine, superoxide dismutase 2 and catalase, and flow cytometry for oxidative burst displayed an increased oxidative stress in persistently CDV-infected DH82 cells (DH82Ond pi) compared to controls. Specifically, the upregulation of TXNIP and NCF4 genes might correlate with an increased intracellular oxidative stress. Indeed, NCF4 encodes for p40phox, a protein that is involved in NADPH oxidase 2 activation.[17,45,71] Additionally, thioredoxin-binding protein 2, encoded by the TXNIP gene, is an important inhibitor of the thioredoxin ROS scavenging system.[45,79] On the other hand, ROS-induced nucleic acid damage did not differ in DH82Ond pi cells compared to non-infected controls. This observation might be interpreted as indicative of an increased oxidative stress associated with the neoplastic nature of DH82 cells rather than an effect of the

viral infection. Indeed, increased intracellular ROS levels are described in the literature as a common feature of cancer cells.[20,23,47] In addition, DH82Ond pi cells displayed an increased expression of SOD2 and CAT compared to non-infected controls. The overexpression of these scavenging enzymes involved in ROS detoxification have been correlated with an increased oxidative stress in neoplastic [20,23,47] as well as in inflammatory conditions.[30] The results obtained by microarray analysis of genes correlated with ER stress [5,9,12,23,45] are consistent with a reduced transcription of genes correlated with this process. Our data might be interpreted as suggestive of an acquired ability of DH82 cells to adapt to the persistent infection with CDV-Ond. This hypothesis is further supported by the finding of an increased expression of ROS-scavenging enzymes in DH82Ond pi cells at both a molecular and protein level, highlighting the plasticity of cancer cells in actively contrasting excessively severe alterations in their redox potential.[23,47]

The HIF-1α expression in DH82Ond pi increased, as demonstrated by Western blot, and showed an unexpected, often sub-membranous distribution, as shown by immunofluorescence and immunoelectron microscopy. Hypoxia could be excluded as the cause of the increased HIF-1α protein expression observed in our in vitro model, due to the fact that cells were cultivated under normoxic conditions in the current study. Consequently, it seems more plausible that the increased expression of HIF-1α in DH82Ond pi cells was induced by the increased oxidative stress level compared to non-infected controls. The down-regulation of 2 PHDs as well as of VHL on a molecular level, in association with a lacking regulation of HIF-1α opposed to an increased expression of the corresponding protein, could imply that the increased protein expression of HIF-1α in DH82Ond pi cells does not refer to an increased synthesis, but rather to an inhibition of the degradation pathway. In addition to the overall increased expression of HIF-1α, the present study revealed an unusual localization of this transcription factor in the sub-membranous compartment and, to a lesser extent, within cytosolic vesicles. Further investigations aiming to better characterize these vesicles, revealed a co-localization of HIF-1α expression with CD63, a marker for the tetraspanin-30 expressed by exosomal membranes.[68] Interestingly, the presence of HIF-1α within CD63+ exosomes has previously been reported in Epstein-Barr virus-infected NP69 cells.[4] On the other hand, in the current

study HIF-1α only occasionally co-localized with CD63+ exosomes, while it frequently overlapped with the localization of CDV-NP. The measles virus N-protein, which is closely related to CDV-NP,[64] is transported within the cell through the endolysosomal system,[43] also rendering this a possible mechanism for the canine counterpart. Furthermore, this observation displays an interesting basis for future investigations on the exact sub-cellular localization of HIF-1α within DH82Ond pi cells.

Finally, microarray data analysis and immunofluorescence confirmed a reduced expression of VEGF-B in DH82Ond pi compared to controls. Specifically, the microarray data analysis revealed a significant down-regulation of different genes involved in the HIF-1α angiogenic downstream pathway, which was further substantiated by a significantly reduced expression of VEGF-B on a molecular and protein level. Though VEGF-B is nowadays recognized as not being directly involved in angiogenesis, this growth factor has been reported as an indirect enhancer of VEGF-A (a well-known inducer of angiogenesis), as well as a key promoter of survival of different cell types (including endothelial cells, pericytes and smooth muscle cells) in several pathological conditions.[13,36,40] As already reported in the literature,[52] the markedly reduced expression of VEGF-B in DH82Ond pi cells did not affect cellular growth nor the apoptotic rate.[49] Interestingly, DH82Ond pi cell xenotransplants displayed a significantly reduced microvessel density compared to non-infected controls.[50] According to our results, it can be assumed that HIF-1α might represent an important mediator of the oncolytic effects described for the in vivo model of DH82Ond pi xenotransplants as reported previously in another viral oncolysis model.[2]

In conclusion, these results of the current study confirmed our hypothesis, suggesting a reduced activation of the HIF-1α angiogenic downstream pathway in DH82Ond pi cells in vitro, most likely due to an excessive, unusually localized, and non-functional expression of HIF-1α which might be the consequence of a decreased cytosolic degradation of this transcriptional factor triggered by CDV-induced increased oxidative stress.

STUDY 2 - Mesenchymal to epithelial transition driven by canine distemper virus infection of canine histiocytic sarcoma cells contributes to a reduced cell motility in vitro

In the context of persistent infection with CDV inducing direct alterations in DH82 cells mRNA and protein expression,[49,52] investigations on the mesenchymal to epithelial transition (MET) might provide promising results.[76] The transition of cells from an epithelial to a mesenchymal state (EMT process) has been extensively studied and validated as one of the major features correlated to invasiveness and metastatic rate of carcinomas.[31,34] In contrast, the reverse process, known as mesenchymal to epithelial transition (MET process), came into the research focus only recently.[76] MET is characterized by the expression of markers typical of epithelial cells (such as E-cadherin, β-catenin, cytokeratin, CD44, and CD34) in sarcoma cells, a process which is often linked to a favourable clinical outcome and a better prognosis due to a decreased invasion and metastatic rate, as reported for human synovial sarcoma and in human leiomyosarcomas.[19,34,37,53,65,66,69,72,76] In the MET process in sarcomas, the classical mesenchymal markers (such as N-cadherin, vimentin, desmin, and alpha-smooth muscle actin - α-SMA) still predominate in the tumor cells in spite of the increased expression of classical epithelial markers, thus determining the so-called “metastable phenotype”. [55,74,76]

Considered the above, the aim of the current study was to investigate the hypothesis that a persistent CDV-Ond infection of DH82 cells triggers the MET process by increasing the expression of epithelial markers, resulting in a less invasive phenotype with decreased motility of the neoplastic cells.[49]

The study was conducted *in vitro* on persistently CDV-infected DH82 cells (DH82Ond pi) by means of microarray data analysis, single and double immunofluorescence further analysed with laser scanning confocal microscopy and 3D reconstruction, immunoblotting, cell morphological analysis with phase contrast microscopy, and scratch and invasion assays as functional confirmation. Further details are provided in the Open Access version of the published paper at the following link: https://onlinelibrary.wiley.com/doi/full/10.1111/jcmm.15585.

CDV-infected cells exhibited an increased expression of epithelial markers such as E-cadherin and cytokeratin 8 compared to controls, indicating a MET process. Specifically, as detected with immunofluorescence and confirmed by immunoblotting, E-cadherin and cytokeratin 8 were signifcantly over-expressed in DH82Ond pi cells compared to non-infected controls. Additionally, the expression of β-catenin was also increased in persistently CDV-infected DH82 cells, although no significant differences were noted. Taken together, these results are indicative of the occurrence of MET in DH82 cells that might be the direct consequence of the infection with CDV-Ond. The findings of the current study confirmed that CDV-Ond might represent a promising oncolytic virus, in addition to the already reported antitumoral effects associated with the viral infection such as the alteration of MMP expression, cortactin distribution, and tumour-associated vascularization and angiogenesis.[7,49,50]

When microarray data were analysed for genes associated with the MET process, a significant down-regulation of the myoferlin gene (MYOF) in DH82Ond pi cells compared to non-infected controls was observed. In breast cancer, a depletion of myoferlin has been associated with increased expression of E-cadherin and reduced levels of fibronectin and vimentin, thus reverting the EMT process and affecting tumour invasiveness.[19,39] Similarly, in the current study the down-regulation of MYOF was associated simultaneously with a down-regulation of the fibronectin gene (FN1) and with an increased protein expression of E-cadherin. TWIST represents another important regulator of the MET process,[42,76] which directly interacts with the expression of genes associated with the epithelial and mesenchymal phenotype. TWIST can down-regulate the E-cadherin and activate the transcription of genes such as N-cadherin and vimentin, which are associated with a mesenchymal phenotype.[42] Interestingly, the mRNA data from the current study displayed an increased expression of TWIST together with an up-regulation of N-cadherin. However, at the protein level, DH82Ond pi cells exclsuively displayed an increased E-cadherin protein expression while a high expression of N-cadherin was present only at the mRNA level. This might be attributed to an incomplete MET status, in which DH82Ond pi cells might be still in a hybrid transient phase with a so-called “metastable phenotype”.[55] Similarly, the expression of both E-cadherin and N-cadherin at the protein level has been reported for circulating breast cancer cells,

expressing mixed epithelial and mesenchymal characteristics in a hybrid state.[77] Additionally, the discrepancies observed in the current study between the mRNA amount and the expression of the corresponding protein might be related to the intervention of miRNAs, which play a well-known role in the EMT and MET process,[31] being able to directly influence the expression of E-cadherin and N-cadherin. Therefore, the role of miRNAs in the MET process in DH82Ond pi cells should be taken in consideration for future investigations.

When the intracellular distribution of the aforementioned markers was investigated in further details together with the data on cellular morphology and the results of scratch and invasion assays, DH82Ond pi cells over-expressing E-cadherin showed more frequently an unexpected localization of this protein in a focal area of the cytoplasm around the nucleus and the Golgi apparatus compared to non-infected controls, despite not reaching statistical significance. Interestingly, E-cadherin cytoplasmic internalization and other post-translational modifications of the EMT/MET effectors might be involved in the uncommon expression of this epithelial marker at the perinuclear level.[42] In the current study, cytokeratin 8 also displayed a different intracellular localization depending on the cellular infection status. Notably, non-infected DH82 cells were characterized by a significantly more frequent cytoplasmic expression of cytokeratin 8, which was focally arranged around the nucleus and the Golgi apparatus, whereas DH82Ond pi cells showed a pronounced expression of variably sized aggregates of this protein within the cytoplasm and intermingled with the cell membrane. Interestingly, the MET process seems to be associated with a decreased cell motility in DH82Ond pi cells. Indeed, as already reported, a knockdown of cytokeratin 8 and 18 in neoplastic epithelial cells was associated with a significantly increased cancer cell motility and invasiveness.[22] In addition, cytoplasmic expression of cytokeratin 8 is linked to a general inhibition of the migratory potential, while a perinuclear localization is related to an increased tumour cell motility.[15] Similar results were obtained in the present study, which revealed an up-regulation of cytokeratin 8 within DH82Ond pi cells compared to non-infected controls. The observation of cytokeratin 8 mainly in a “membranous to cytoplasmic” localization in DH82Ond pi cells might be one of the factors leading to the reduced cell motility observed in the scratch and in the invasion assay. In contrast, an

increased perinuclear expression of cytokeratin 8 in non-infected DH82 cells was associated with an increased cell motility in both functional assays applied. Furthermore, intermediate filaments and specifically cytokeratins are involved in cell adhesion, localization of the organelles, and changing of cellular shape.[54] This might be correlated with the more homogeneous round cellular shape of DH82Ond pi cultures compared to non-infected controls, with the latter characterized by a more pleomorphic phenotype. The hypothesis of a reduced cell motility due to a rearrangement of intermediate filaments during the MET process might be further supported by the fact that a reduced expression of vimentin is also associated with a reduced cell motility during MET.[44] In the present study, a similar number of cells expressed vimentin regardless of the infection status. However, DH82Ond pi and non-infected DH82 cells showed a higher number of positive round and spindle cells, respectively. This finding was mirrored by the fact that DH82Ond pi and non-infected DH82 cells showed a more frequent focal and diffuse cytoplasmic distribution of vimentin, respectively, which was confirmed by 3D reconstructions. Taken together, these results suggest that the different intracellular distribution of vimentin between DH82Ond pi and non-infected controls might be correlated to the predominant cellular phenotype among each cell population, rather than to the infection status. Nevertheless, the predominant cellular phenotype among each cell population seems to be dependent on the infection status, thus suggesting an indirect role of the virus in the intracellular redistribution of vimentin. Interestingly, a spontaneous CDV infection of canine brain cells has also been reported to modify cytoskeletal proteins such as vimentin and glial fibrillary acid protein (GFAP) in vivo.[60] Considered that the literature highlighted the fundamental influence of the cell shape on motility,[6,54] virus-induced morphological and structural (i.e. intermediate filament rearrangement) modifications might be the cause of the observed alterations in cell motility and invasiveness.

In conclusion, these results of the current study confirmed our hypothesis, suggesting that CDV infection of DH82 cells triggers the MET process through an increased expression of epithelial markers, resulting in decreased cell motility and invasiveness which were most likely caused by a rearrangement of cytoskeletal intermediate filaments.

ADDENDUM REFERENCES

1. Abadie J, Hédan B, Cadieu E, et al. Epidemiology, pathology, and genetics of histiocytic sarcoma in the Bernese mountain dog breed. *J Hered*. 2009;100:S19–S27.
2. Abd-Aziz N, Stanbridge EJ, Shafee N. Newcastle disease virus degrades HIF-1α through proteasomal pathways independent of VHL and p53. *J Gen Virol*. 2016;97:3174–3182.
3. Adams VJ, Evans KM, Sampson J, Wood JLN. Methods and mortality results of a health survey of purebred dogs in the UK. *J Small Anim Pract*. 2010;51:512–524.
4. Aga M, Bentz GL, Raffa S, et al. Exosomal HIF1α supports invasive potential of nasopharyngeal carcinoma-associated LMP1-positive exosomes. *Oncogene*. 2014;33:4613–4622.
5. Attig F, Spitzbarth I, Kalkuhl A, et al. Reactive oxygen species are key mediators of demyelination in canine distemper leukoencephalitis but not in theiler's murine encephalomyelitis. *Int J Mol Sci*. 2019;20:1–29.
6. Battaglia RA, Delic S, Herrmann H, Snider NT. Vimentin on the move: New developments in cell migration [version 1; referees: 2 approved]. 2018;7:1–10.
7. Beineke A, Puff C, Seehusen F, Baumgärtner W. Pathogenesis and immunopathology of systemic and nervous canine distemper. *Vet Immunol Immunopathol*. 2009;127:1–18.
8. Bell J, McFadden G. Viruses for tumor therapy. *Cell Host Microbe*. 2014;15:260–265.
9. Bhandary B, Marahatta A, Kim HR, Chae HJ. An involvement of oxidative stress in endoplasmic reticulum stress and its associated diseases. *Int J Mol Sci*. 2013;14:434–456.
10. Bluming A, Ziegler J. Regression of Burkitt's lymphoma in association with measles infection. *Regres Burkitt's lymphoma Assoc with measles Infect*. 1971;105–106.
11. Boerkamp KM, van der Kooij M, van Steenbeek FG, et al. Gene expression profiling of histiocytic sarcomas in a canine model: the predisposed flatcoated retriever dog. *PLoS One*. 2013;8:e71094.
12. Brunner JM, Plattet P, Doucey MA, et al. Morbillivirus glycoprotein expression induces ER stress, alters Ca 2+ homeostasis and results in the release of vasostatin. *PLoS One*. 2012;7:1–12.
13. Bry M, Kivelä R, Leppänen VM, Alitalo K. Vascular endothelial growth factor-B in physiology and disease. *Physiol Rev*. 2014;94:779–794.
14. Bürge T, Griot C, Vandevelde M, Peterhans E. Antiviral antibodies stimulate production of reactive oxygen species in cultured canine brain cells infected with canine distemper virus. *J Virol*. 1989;63:2790–2797.
15. Busch T, Armacki M, Eiseler T, et al. Keratin 8 phosphorylation regulates keratin reorganization and migration of epithelial tumor cells. *J Cell Sci*. 2012 May;125:2148–2159.
16. Challapalli A, Carroll L, Aboagye EO. Molecular mechanisms of hypoxia in cancer. *Clin Transl Imaging*. 2017;5:225–253.
17. Dahlgren C, Karlsson A, Bylund J. Intracellular Neutrophil Oxidants: From Laboratory Curiosity to Clinical Reality. *J Immunol*. 2019;202:3127–3134.
18. Dispenzieri A, Tong C, Laplant B, et al. Phase I trial of systemic administration of edmonston strain of measles virus genetically engineered to express the sodium iodide symporter in patients with recurrent or refractory multiple myeloma. *Leukemia*. 2017;31:2791–2798.
19. Dong Y, Kang H, Liu H, et al. Myoferlin, a Membrane Protein with Emerging Oncogenic Roles. *Biomed Res Int*. 2019;2019.
20. Fiaschi T, Chiarugi P. Oxidative stress, tumor microenvironment, and metabolic reprogramming: A diabolic liaison. *Int J Cell Biol*. 2012;2012.
21. Fidel J, Schiller I, Hauser B, et al. Histiocytic sarcomas in flat-coated retrievers: a summary of 37 cases (November 1998-March 2005). *Vet Comp Oncol*. 2006;4:63–74.
22. Fortier A-M, Asselin E, Cadrin M. Keratin 8 and 18 loss in epithelial cancer cells increases collective cell migration and cisplatin sensitivity through claudin1 up-regulation. *J Biol Chem*. 2013 Apr;288:11555–11571.
23. Galadari S, Rahman A, Pallichankandy S, Thayyullathil F. Reactive oxygen species and cancer paradox: To promote or to suppress? *Free Radic Biol Med*. 2017;104:144–164.
24. Gopas J, Itzhaky D, Segev Y, et al. Persistent measles virus infection enhances major histocompatibility complex class I expression and immunogenicity of murine neuroblastoma cells. *Cancer Immunol Immunother*. 1992;34:313–320.

25. Griot C, Bürge T, Vandevelde M, Peterhans E. Antibody-induced generation of reactive oxygen radicals by brain macrophages in canine distemper encephalitis: a mechanism for bystander demyelination. *Acta Neuropathol*. 1989;78:396–403.
26. Grőne A, Fonfara S, Baumgärtner W. Cell type-dependent cytokine expression after canine distemper virus infection. *Viral Immunol*. 2002;15:493–505.
27. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 2011;144:646–674.
28. Heinzerling L, Künzi V, Oberholzer PA, Kündig T, Naim H, Dummer R. Oncolytic measles virus in cutaneous T-cell lymphomas mounts antitumor immune responses in vivo and targets interferon-resistant tumor cells. *Blood*. 2005;106:2287–2294.
29. Hornick JL, Jaffe ES, Fletcher CDM. Extranodal histiocytic sarcoma: Clinicopathologic analysis of 14 cases of a rare epithelioid malignancy. *Am J Surg Pathol*. 2004;28:1133–1144.
30. van Horssen J, Schreibelt G, Drexhage J, et al. Severe oxidative damage in multiple sclerosis lesions coincides with enhanced antioxidant enzyme expression. *Free Radic Biol Med*. 2008;45:1729–1737.
31. Kalluri R, Neilson EG. Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest*. 2003;112:1776–1784.
32. Karadeniz A, Hanedan B, Cemek M, Börkü MK. Relationship between canine distemper and oxidative stress in dogs. *Rev Med Vet (Toulouse)*. 2008;159:462–467.
33. Kim M, Chung YH, Johnston RN. Reovirus and tumor oncolysis. *J Microbiol*. 2007;45:187–192.
34. Korpal M, Ell BJ, Buffa FM, et al. Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. *Nat Med*. 2011;17:1101.
35. Krock BL, Skuli N, Simon MC. Hypoxia-Induced Angiogenesis: Good and Evil. 2011 Dec;2:1117–1133.
36. Lal N, Puri K, Rodrigues B. Vascular Endothelial Growth Factor B and Its Signaling. *Front Cardiovasc Med*. 2018;5:1–9.
37. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat Rev Mol cell Biol*. 2014;15:178–196.
38. Lapp S, Pfankuche VM, Baumgärtner W, Puff C. Viral oncolysis - Can insights from measles be transferred to canine distemper virus? *Viruses*. 2014;6:2340–2375.
39. Li R, Ackerman IV WE, Mihai C, Volakis LI, Ghadiali S, Kniss DA. Myoferlin depletion in breast cancer cells promotes mesenchymal to epithelial shape change and stalls invasion. *PLoS One*. 2012;7:e39766.
40. Li X, Lee C, Tang Z, et al. VEGF-B. A survival, or an angiogenic factor? 2009;3:322–327.
41. Liu T-CC, Galanis E, Kirn D. Clinical trial results with oncolytic virotherapy: a century of promise, a decade of progress. *Nat Clin Pract Oncol*. 2007;4:101–117.
42. Lu W, Kang Y. Epithelial-mesenchymal plasticity in cancer progression and metastasis. *Dev Cell*. 2019;49:361–374.
43. Marie JC, Saltel F, Escola J-M, Jurdic P, Wild TF, Horvat B. Cell Surface Delivery of the Measles Virus Nucleoprotein: a Viral Strategy To Induce Immunosuppression. *J Virol*. 2004;78:11952–11961.
44. Mendez MG, Kojima S-I, Goldman RD. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. *FASEB J Off Publ Fed Am Soc Exp Biol*. 2010 Jun;24:1838–1851.
45. Mittal M, Siddiqui MR, Tran K, Reddy SP, Malik AB. Reactive oxygen species in inflammation and tissue injury. *Antioxidants Redox Signal*. 2014;20:1126–1167.
46. Moore PF. A Review of Histiocytic Diseases of Dogs and Cats. *Vet Pathol*. 2014;51:167–184.
47. Pani G, Galeotti T, Chiarugi P. Metastasis: Cancer cell's escape from oxidative stress. *Cancer Metastasis Rev*. 2010;29:351–378.
48. Parato KA, Senger D, Forsyth PAJ, Bell JC. Recent progress in the battle between oncolytic viruses and tumours. *Nat Rev Cancer*. 2005;5:965–976.
49. Pfankuche VM, Sayed-Ahmed M, Contioso VB, et al. Persistent Morbillivirus Infection Leads to Altered Cortactin Distribution in Histiocytic Sarcoma Cells with Decreased Cellular Migration Capacity. *PLoS One*. 2016;11:e0167517.
50. Pfankuche VM, Spitzbarth I, Lapp S, et al. Reduced angiogenic gene expression in morbillivirus-triggered oncolysis in a translational model for histiocytic sarcoma. *J Cell Mol Med*. 2017;21:816–830.
51. Pisoschi AM, Pop A. The role of antioxidants in the chemistry of oxidative stress: A review. *Eur J Med Chem*. 2015;97:55–74.

52. Puff C, Krudewig C, Imbschweiler I, Baumgärtner W, Alldinger S. Influence of persistent canine distemper virus infection on expression of RECK, matrix-metalloproteinases and their inhibitors in a canine macrophage/monocytic tumour cell line (DH82). *Vet J*. 2009;182:100–107.
53. Saito T, Oda Y, Yamamoto H, et al. Nuclear β-catenin correlates with cyclin D1 expression in spindle and pleomorphic sarcomas but not in synovial sarcoma. *Hum Pathol*. 2006;37:689–697.
54. Sanghvi-Shah R, Weber GF. Intermediate Filaments at the Junction of Mechanotransduction, Migration, and Development. *Front cell Dev Biol*. 2017;5:81.
55. Sannino G, Marchetto A, Kirchner T, Grünewald TGP. Epithelial-to-mesenchymal and mesenchymal-to-epithelial transition in mesenchymal tumors: a paradox in sarcomas? *Cancer Res*. 2017;77:4556–4561.
56. Schattner A, Rager-Zisman B, Bloom BR. Persistent viral infection affects tumorigenicity of a neuroblastoma cell line. *Cell Immunol*. 1985;90:103–114.
57. Schieber M, Chandel NS. ROS Function in Redox Signaling and Oxidative Stress Michael. *Curr Biol*. 2014;24:R453–R462.
58. Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. *Philos Trans R Soc B Biol Sci*. 2015;370:20140231.
59. Schlick K, Aigelsreiter A, Pichler M, et al. Histiocytic sarcoma - Targeted therapy: Novel therapeutic options? A series of 4 cases. *Onkologie*. 2012;35:447–450.
60. Seehusen F, Orlando EA, Wewetzer K, Baumgärtner W. Vimentin-positive astrocytes in canine distemper: a target for canine distemper virus especially in chronic demyelinating lesions? *Acta Neuropathol*. 2007;114:597–608.
61. Semenza GL. Oxygen Sensing, Hypoxia-Inducible Factors, and Disease Pathophysiology. *Annu Rev Pathol Mech Dis*. 2014;9:47–71.
62. Sinkovics JG. Viral oncolysates as human tumor vaccines. *Int Rev Immunol*. 1991;7:259–287.
63. Stein VM, Czub M, Schreiner N, et al. Microglial cell activation in demyelinating canine distemper lesions. *J Neuroimmunol*. 2004;153:122–131.
64. Stephenson JR, Ter Meulen V. Antigenic relationships between measles and canine distemper viruses: Comparison of immune response in animals and humans to individual virus-specific polypeptides. *Proc Natl Acad Sci U S A*. 1979;76:6601–6605.
65. Sun BC, Sun Y, Zhao XL, Liu YX, Zhang SW. Expressions and significance of E-cadherin and beta-catenin in synovial sarcoma. *Zhonghua Zhong Liu Za Zhi*. 2005;27:727–730.
66. Suster S. Epithelioid leiomyosarcoma of the skin and subcutaneous tissue: clinicopathologic, immunohistochemical, and ultrastructural study of five cases. *Am J Surg Pathol*. 1994;18:232–240.
67. Suter SE, Chein MB, Von Messling V, et al. In vitro canine distemper virus infection of canine lymphoid cells: a prelude to oncolytic therapy for lymphoma. *Clin Cancer Res*. 2005;11:1579–1587.
68. Théry C, Witwer KW, Aikawa E, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles*. 2018;7.
69. Tian W, Wang G, Yang J, Pan Y, Ma Y. Prognostic role of E-cadherin and Vimentin expression in various subtypes of soft tissue leiomyosarcomas. *Med Oncol*. 2013;30:401.
70. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65:87–108.
71. Ushio-Fukai M, Nakamura Y. Reactive Oxygen Species and Angiogenesis NADPH Oxidase as. *Cancer Lett*. 2008;266:37–52.
72. Wei CH, Wan CY, Chen A, Tseng HH. Epithelioid leiomyosarcoma of the breast: report of a case. *J Formos Med Assoc Taiwan yi zhi*. 1993;92:379–381.
73. Wellman ML, Krakowka S, Jacobs RM, Kociba GJ. A Macrophage-Monocyte Cell Line from a Dog with Malignant Histiocytosis. *Vitr Cell Dev Biol*. 1988;24:223–229.
74. Wells A, Yates C, Shepard CR. E-cadherin as an indicator of mesenchymal to epithelial reverting transitions during the metastatic seeding of disseminated carcinomas. *Clin Exp Metastasis*. 2008;25:621–628.
75. Wojton J, Kaur B. Impact of tumor microenvironment on oncolytic viral therapy. *Cytokine Growth Factor Rev*. 2010;21:127–134.
76. Yang J, Du X, Wang G, et al. Mesenchymal to epithelial transition in sarcomas. *Eur J Cancer*. 2014;50:593–601.

77. Yu M, Bardia A, Wittner BS, et al. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science*. 2013 Feb;339:580–584.
78. Zepeda AB, Pessoa A, Castillo RL, Figueroa CA, Pulgar VM, Farías JG. Cellular and molecular mechanisms in the hypoxic tissue: Role of HIF-1 and ROS. *Cell Biochem Funct*. 2013;31:451–459.
79. Zhou J, Yu Q, Chng WJ. TXNIP (VDUP-1, TBP-2): A major redox regulator commonly suppressed in cancer by epigenetic mechanisms. Vol. 43, *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd; 2011:1668–1673.

REPORT OF THE ACTIVITES OF THE PH.D. CANDIDATE

SCIENTIFIC PRODUCTION

List of publications in peer-reviewed journals:

- Gambini M, Martini V, Bernardi S, Caniatti M, Gelain ME, Roccabianca P, Comazzi S (2021) Cytology of feline nodal lymphoma: low inter-observer agreement and variable accuracy in lymphoma immunophenotype prediction. *Journal of Comparative Pathology*, **184**, 1-6.
- Armando F*, Gambini M*, Corradi A, Becker K, Marek K et al. (2020) Mesenchymal to epithelial transition driven by canine distemper virus infection of canine histiocytic sarcoma cells contributes to a reduced cell motility in vitro. *Journal of Cellular and Molecular Medicine*, **24**, 9332–9348. *these Authors contributed equally to this work
- Armando F*, Gambini M*, Corradi A, Giudice C, Pfankuche VM et al. (2020) Oxidative stress in canine histiocytic sarcoma cells induced by an infection with canine distemper virus led to a dysregulation of HIF-1α downstream pathway resulting in a reduced expression of VEGF-B in vitro. *Viruses*, **12**. *these Authors contributed equally to this work
- Ferrari R, Chiti LE, Manfredi M, Ravasio G, De Zani D, Zani D D, Giudice C, Gambini M, Stefanello D. (2020) Biopsy of sentinel lymph nodes after injection of methylene blue and lymphoscintigraphic guidance in 30 dogs with mast cell tumors. *Veterinary Surgery*, **49**, 1099–1108.
- Tecilla M, Gambini M, Pigoli C, Grieco V, Caniatti M. (2020) What is your diagnosis? Prepubic mass in a mare. *Veterinary Clinical Pathology*.
- Tecilla M*, Gambini M*, Forlani A, Caniatti M, Ghisleni G et al. (2019) Evaluation of cytological diagnostic accuracy for canine splenic neoplasms: An investigation in 78 cases using STARD guidelines. *PLoS ONE*, **14**. *these Authors contributed equally to this work

List of manuscripts currently under review in peer-reviewed journals:

- Armando F*, Pigoli C*, Gambini M*, Ghidelli A, Ghisleni G, Corradi A, Passeri B, Caniatti M, Grieco V, Baumgärtner W, Puff C (2020). Peripheral nerve sheath tumors resembling human atypical neurofibroma in goldfish (*Carassius auratus*). *Veterinary Pathology* – currently under the 3rd round of revision. *these Authors contributed equally to this work
- Sabattini S, Kiupel M, Finotello R, Stefanello D, Faroni E, Bertazzolo W, Bonfanti U, Rigillo A, Del Magno S, Foglia A, Aresu L, Gambini M, Caniatti M, Marconato L (2021). A retrospective study on prophylactic regional lymphadenectomy versus nodal observation only in the management of dogs with stage I, completely resected, low-grade cutaneous mast cell tumors. *BMC Veterinary Research* – currently under the 1st round of revision.
- Armando F, Fayyad A, Arms S, Barthel Y, Schaudien D, Rohn K, Gambini M, Lombardo M S, Beineke A, Baumgärtner W, Puff C (2021). Intratumoral canine distemper virus infection inhibits tumor growth by modulation of the tumor microenvironment in a murine xenograft model of canine histiocytic sarcoma. *International Journal of Molecular Science* – accepted for publication.
- Chiti L E, Boracchi P, Gambini M, Amati M, Pigoli C, Faverzani S, Di Giancamillo M, Stefanello D (2021). The minimum sub-set of variables for oncological prognostication in canine cutaneous mast cell tumour. *Topics in Companion Animal Medicine* - currently under the 1st round of revision.

Oral presentations

- Gambini M, Forlani A, Tecilla M, Caniatti M, Roccabianca P. Diagnostic accuracy of cytology in canine splenic lesions. *Proceedings of Veterinary and Animal Science Days 2018, 6 th - 8 th June, Milan, Italy*.
- Gambini M, Buzzi G, Recordati C, Caniatti M, Giudice C, Pigoli C, Tecilla M. “Should I spend my time counting mast cells?” - preliminary results concerning mast cell quantification in cytological specimens for the evaluation of nodal

metastasis in mast cell tumor-bearing dogs. *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019.*

Scientific posters

- Pigoli C, Caniatti M, Giudice C, Grieco V, Ferrari R, Manfredi M, Gambini M. Agreement between cytology and histology in the diagnosis of sentinel lymph node metastasis in canine cutaneous mast cell tumors: preliminary results. *Proceedings of Joint ESVP-ECVP Congress Cluj-Napoca, Romania September 5th -8th, 2018*.
- Tecilla M, Gambini M, Bardi E, Lubian E, Roccabianca P. Survey of upper alimentary tract neoplasia in prairie dogs (Cynomis ludivicianus). *Proceedings of Joint ESVP-ECVP Congress Cluj-Napoca, Romania September 5th -8th, 2018*.
- Ghisleni G, Attini M, Gambini M, Tecilla M, Cardinelli SS, Melzi S, Caniatti M. Canine keloidal fibrosarcoma: two case reports and review of the literature. *20th ESVCP-ECVCP Meeting – Athens, Greece – 17-20/10/2018*.
- Gambini M, Bardi E, Romussi S, Pigoli C, Roccabianca P, Dell'Aere S, Tecilla M. Ovotesticular disorder of sex determination in a veiled chameleon (*Chamaeleo calyptratus*). *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019*.
- Tecilla M, Gambini M, Pigoli C, Bardi E, Romussi S, Ghisleni G, Origgi F. Oral mycosis in a veiled chameleon (*Chamaeleo calyptratus*). *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019*.
- Pigoli C, Gambini M, Bardi E, Bassi J, Manfredi M, Ghisleni G, Tecilla M. Congenital skull malformation in a captive python (*Pythus regius*). *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019*.
- Pigoli C, Gambini M, Tecilla M, Bielli M, Roccabianca P. Benign ovarian teratoma in a greek tortoise (*Testudo graeca*). *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019*.

- Ghisleni G, Gambini M, Crosta L, Capitelli R, Pigoli C, Paciletti V, Tecilla M. Fine-needle aspiration cytology (FNAC) of palpable masses in Psittaciformes. *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019*.
- Tecilla M, Caniatti M, Pazzini L, Pigoli C, Gambini M, Turchetto S, Ghisleni G, Roccabianca P. Vascular melanosis with muscular necrosis in the western Mediterranean wild elasmobranch fishes, *Raja clavata* and *Raja oxyrhynchus*. *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019*.
- Armando F, Gambini M, Corradi A, Giudice C, Pfankuche V M, Brogden G, Attig F, von Köckritz-Blickwede M, Baumgärtner W, Puff C. Oxidative stress in canine histiocytic sarcoma cells (DH82 cells) induced by a persistent canine distemper virus infection leads to impairment of the hif-1α downstream pathway in vitro. *63th German Pathologists National Congress – Fulda, Germany - 6-8/03/2020*.
- Chiti L E, Ferrari R, Manfredi M, Zani D D, De Zani D, Gambini M, Giudice C, Grieco V, Stefanello D. Gamma-probe guided sentinel lymph node extirpation in spontaneous head and neck malignancies of the dog: impact on tumor staging and surgical anatomy findings. *29th ECVS Annual Scientific Meeting – Valencia, Spain – 2-4/07/2020*. (canceled due to COVID-19 emergency)

Other scientific productions

- Ferrari R, Manfredi M, Chiti LE, Zani DD, De Zani D, Rabbogliatti V, Gambini M, Giudice C, Grieco V, Stefanello D. Sentinel lymph node biopsy guided by combination technique (lymphoscintigraphy and blue dye) for mast cell tumor in dogs: results in 31 consecutive cases. *28th ECVS Annual Scientific Meeting – Budapest, Hungary – 4-6/07/2019* – Oral presentation.
- Stranieri A, Scavone D, Ferrari R, Gambini M, Martinelli L, Martini V, Giordano A, Paltrineri S. Another hint from the sysmex XT-2000IV scattergram. *Joint Congress ECVP-ESVP-ECVCP-ESVCP – Burgers' Zoom, Arnhem, The Netherlands - 25-28/09/2019* – Oral presentation.

- Armando F, Gambini M, Corradi A, Pfankuche V M, Mergani A E, Brogden G, von Köckritz-Blickwede M, Baumgärtner W, Puff C. Canine distemper virus triggers mesenchymal to epithelial transition in canine histiocytic sarcoma cells. *63rd German Pathologists National Congress – Fulda, Germany - 6-8/03/2020* – Oral presentation.
- Armando F, Pigoli C, Gambini M, Ghidelli A, Ghisleni G, Corradi A, Passeri B, Caniatti M, Grieco V, Baumgärtner W, Puff C (2020) Peripheral nerve sheath tumors in goldfish (*Carassius auratus*) similar to human atypical neurofibroma. *64th German Pathologists National Congress – Fulda, Germany - 5-7/03/2020* – Oral presentation.

OTHER ACTIVITIES

Attendance to courses, seminars, workshops and scientific meetings

- Weekly training (each Friday) for European College of Veterinary Pathologists (ECVP) examination by Professor Paola Roccabianca. Departement of Veterinary Medicine, University of Milan, Milan (Italy) – 17/10/2017 – present.
- 16th ECVP (European College of Veterinary Pathologists) Summer school. Faculty of Veterinary Medicine, University CEU Cardenal Herrera, Valencia (Spain) – 16-27/07/2018.
- 17th ECVP (European College of Veterinary Pathologists) Summer school. *Faculty of Veterinary Medicine, University CEU Cardenal Herrera,* Valencia (Spain) – 15-26/07/2019 (88 hours).
- Joint Congress of Veterinary Pathology and Veterinary Clinical Pathology (ECVP-ESVP-ECVCP-ESVCP) - *Burgers' Zoo, Arnhem, The Netherlands* - 25-28/09/2019.
- 2020 Pathology Of Laboratory Animals course (POLA) by Davis-Thompson Foundation – online course – 3-7/08/2020.

Any other relevant activity

- Externship of 6 months (November 2019 – December 2019 and January 2020 – May 2020) at the University of Veterinary Medicine of Hannover (TiHo Hannover – Germany).
- Blinded evaluator of cytological samples in a research study investigating the diagnostic accuracy of cytology in predicting the immunophenotype of canine lymphomas (managed by Dr. Valeria Martini, DVM, Ph.D., and Professor Stefano Comazzi, DVM, Ph.D., DECVCP).
- Peer-reviewer for the journals "*Acta Veterinaria Hungarica*" (1 manuscript) and "*Animals*" (1 manuscript).
- Co-Supervisor for the degree thesis in Veterinary Medicine entitled "Valutazione del numero di mastociti in campioni citologici linfonodali di canni affetti da mastocitoma e non" (Candidate: Giulia Buzzi; Supervisor: Professor Mario Caniatti). *Università degli Studi di Milano – 15/05/2019*.
- Co-Supervisor for the degree thesis in Veterinary Medicine with the provisional title "Accuratezza diagnostica della citologia linfonodale nel mastocitoma canino" (Candidate: Melinda Diotti; Supervisor: Professor Mario Caniatti). *Università degli Studi di Milano – planned for half of October 2019*.
- Lecturer for the microscope practical sessions (5 hours) of the course "Morphological and molecular basis of the central nervous system and its pathologies" (held by Professor Tiziana Brevini and Professor Fabrizio Ceciliani) for the Magistral Degree in "Scienze Biotecnologiche Veterinarie". *Università degli Studi di Milano – 04/04/2019 and 11/04/2019*.
- Daily involved as an ECVP Trainee in the Necropsy, Histopathology and Diagnostic Cytology Services of the Diagnostic Pathology Service of the Veterinary Teaching Hospital of the University of Milan (https://www.ospedaleveterinario.unimi.it/staff/matteo-gambini-anatomia-patologica/)

ACKNOWLEDGMENT

The Ph.D. candidate and his Supervisor want to thank Professor Wolfgang Baumgärtner (DVM, Ph.D., DACVP), Professor Christina Puff (DVM, DECVP), and all the Staff from the Institute of Pathology of the University of Veterinary Medicine of Hannover (TiHo Hannover, Germany), as well as Doctor Federico Armando (DVM) from the Department of Veterinary Medicine Sciences of the University of Parma (Parma, Italy), for the possibility of taking part into the 2 research projects described in the Addendum of the current Thesis.

The Ph.D. candidate and his Supervisor want to thank also Doctor Valeria Martini (DVM, Ph.D.) and Professor Stefano Comazzi (DVM, Ph.D., DECVCP) for the possibility of taking part into the research project described in Chapter 2 of the current Thesis.

The Ph.D. candidate further wish to express his infinite gratitude to Doctor Letizia Pozzi (DVM, Animal Health Specialist) for the accurate proof-reading of the last draft of the current Thesis.