

Scotland's Rural College

Nuclear pleomorphism: role in grading and prognosis of canine mammary carcinomas

Santos, M; Correia-Gomes, C; Santos, A; de Matos, A; Rocha, E; Lopes, C; Pereira, PD

Published in:
The Veterinary Journal

DOI:
[10.1016/j.tvjl.2014.03.019](https://doi.org/10.1016/j.tvjl.2014.03.019)

Print publication: 01/01/2014

Document Version
Peer reviewed version

[Link to publication](#)

Citation for published version (APA):
Santos, M., Correia-Gomes, C., Santos, A., de Matos, A., Rocha, E., Lopes, C., & Pereira, PD. (2014). Nuclear pleomorphism: role in grading and prognosis of canine mammary carcinomas. *The Veterinary Journal*, 200(3), 426 - 433. <https://doi.org/10.1016/j.tvjl.2014.03.019>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Nuclear pleomorphism: Role in grading and prognosis of canine mammary carcinomas

Marta Santos ^{a,*}, Carla Correia-Gomes ^b, Andreia Santos ^c, Augusto de Matos ^{d,e},
Eduardo Rocha ^{a,f}, Carlos Lopes ^g, Patrícia Dias Pereira ^g

^a *Laboratory of Histology and Embryology, Department of Microscopy, Instituto Ciências Biomédicas Abel Salazar, ICBAS – UPorto, University of Porto, Portugal.*

^b *Epidemiology Research Unit, Future Farming Systems, Scotland's Rural College (SRUC), UK*

^c *Faculty of Veterinary Medicine – Lusófona University of Humanities and Technologies (FMV-ULHT)*

^d *Department of Veterinary Clinics, ICBAS – UPorto, Portugal*

^e *Animal Science and Study Central (CECA), Food and Agrarian Sciences and Technologies Institute (ICETA)*

^f *CIIMAR/CIMAR LA, Laboratory of Cellular, Molecular and Analytical studies, Interdisciplinary Center for Marine and Environmental Research, UPorto, Portugal*

^g *Department of Pathology and Molecular Immunology, ICBAS – UPorto, Portugal*

* Corresponding author. Tel.: +351220428243

E-mail address: mssantos@icbas.up.pt (Marta Santos)

Abstract

Canine mammary tumours are highly heterogeneous in morphology and behaviour and successful clinical management requires robust prognostic factors. Histological grade, determined by the Nottingham method, has been considered one of these factors. Despite the adoption of this method, it is unknown if inter-observer agreement exists regarding the assessment of its parameters in canine mammary carcinomas (CMC). In this study, the agreement between two observers using the Nottingham method was evaluated in 89 CMC. Histological evidence of vascular invasion and/or lymph node metastases (both early signs of tumour aggressiveness) was recorded. For 48 animals, 2-years follow-up data were available. Nuclear pleomorphism was quantitatively assessed using a stereological method, that allowed for an unbiased estimation of nuclear size and its variability, by determining the volume-weighted mean nuclear volume (\bar{v}_v). Differences between the \bar{v}_v estimations and the nuclear pleomorphism scores were evaluated. Additionally, the prognostic significance of clinicopathological features including nuclear score and \bar{v}_v was evaluated.

A poor agreement between the two observers was obtained (Kappa value 0.46). Tumours scored as 1 and 2 presented similar \bar{v}_v and only tumours scored 3 presented significantly higher estimates. The \bar{v}_v was not associated with vascular invasion and/or lymph node metastases, but was higher in tumours that progressed during follow-up. In multivariable analysis only tumour size was an independent factor regarding evidence of aggressiveness and an optimal cut-off of 2.9 cm was defined.

Keywords: Canine; Mammary tumours; Histological grade; Nuclear volume; Prognosis; Stereology.

Introduction

Mammary gland tumours are the most frequent neoplasia in female dogs, particularly in countries where spaying is not performed routinely early in life (Sorenmo, 2003). Approximately half of the affected dogs have malignant disease based on the histopathological examination (Lana et al., 2007). Metastases to distant organs are the most common cause of morbidity and mortality associated with canine mammary tumours (Lana et al., 2007).

Mammary neoplasia is a complex disease characterized by a heterogeneous clinical and biological behaviour. This probably contributes to the limited therapeutic options that are currently available (Goldschmidt et al., 2011; Klopffleisch et al., 2011, Sorenmo et al., 2011). Searching for new prognostic parameters is of utmost importance (Philibert et al., 2003; Chang et al., 2005; Santos et al., 2013). Histological grade is well established as an important prognostic factor in human breast cancer (Rakha et al., 2010). In veterinary pathology, different grading systems have been used and claimed to have prognostic value, but such a variety of methods is itself a drawback (Goldschmidt et al., 2011; Matos et al., 2012).

One of the most frequently used methods for histological grading of canine mammary tumours is the Nottingham method, originally developed for human breast cancer (Elston and Ellis, 1998). Karayannopoulou et al. (2005) provided some evidence that this method had advantages when compared to previous methods (that only use nuclear and cellular pleomorphism as parameters) for prognostic purposes in mammary tumours of dogs. The Nottingham grade method is based on the semi-quantitative assessment of three morphological features, namely, tubule formation, mitotic counts

and nuclear pleomorphism. The latter should mainly be evaluated by comparing the nuclear size and variation with normal mammary epithelial cells or lymphocytes (Elston and Ellis, 1998). Nuclear pleomorphism is considered a hallmark of malignant transformation and increases in the absence of cell differentiation.

Irrespective of the method, grading is usually based on a subjective, experience-dependent judgment by a pathologist, and on qualitative or semi-quantitative evaluations of morphologic and cytological features (Sørensen, 1992; Artacho-Pérula and Roldán-Villalobos, 1997). Not only the observer, but also the selection of the tumour areas to assess grade can significantly influence the scoring of the parameters. This is especially relevant for canine mammary tumours, which are intrinsically very heterogeneous (Klopffleisch et al., 2011).

The subjective nature of morphological evaluation is associated with a high risk of inter- and intra-observer variations that can affect reproducibility and accuracy and limit comparative analysis between studies (Artacho-Pérula and Roldán-Villalobos, 1997). A simple way to overcome such subjectivity is to use unbiased quantitative parameters for scoring neoplasms by applying appropriate stereological methods. This has been used in various tumour types, including human breast cancer (Sørensen, 1992; Ladekarl and Sørensen, 1993; Steiner et al., 1994; Artacho-Pérula and Roldán-Villalobos, 1997; Ladekarl, 1998; Yörükoglu et al., 1998; Soda et al., 1999; Fujikawa et al., 2000). The design-based stereological estimates are precise, shape-independent, and allow for an objective evaluation of several cytological features, including nuclear pleomorphism (Ladekarl, 1998). However, to the best of our knowledge, stereological methods were never applied for studying canine mammary carcinomas (CMC).

The point-sampled intercepts (PSI) method is considered the easiest way to objectively quantify nuclear pleomorphism (Gundersen and Jensen, 1985; Ladekarl, 1998). The PSI generates estimations of the volume-weighted mean nuclear volume, \bar{v}_v (Sørensen, 1992). The \bar{v}_v is not an intuitive parameter: it involves sampling the nuclei in proportion to their volumes, therefore reflecting nuclear size variation and pleomorphism (Sørensen, 1992). The \bar{v}_v increases as the nucleus enlarges, and is further augmented when a substantial variation in nuclear size exists. This parameter is mostly used in histopathology, and the \bar{v}_v estimate has been shown to provide relevant information in diagnosis and grading of breast tumours (Ladekarl and Sørensen, 1993; Artacho-Pérula and Roldán-Villalobos, 1997).

The aims of this study were (1) to estimate the inter-observer variability in scoring nuclear pleomorphism-, using the criteria of the Nottingham grading system; (2) to assess nuclear pleomorphism of CMC using stereological tools, namely the PSI method; (3) to compare the stereological estimation with the histological score of nuclear pleomorphism; and (4) to verify the prognostic significance of the nuclear pleomorphism estimations by using univariable and multivariable analyses.

Material and methods

Clinical cases and histological analysis

In this study 89 spontaneous CMC (47 simple carcinomas and 42 complex carcinomas) from 56 female dogs that underwent surgical removal were analysed retrospectively. Owners gave informed consent for the surgery (with curative intents) and follow-up after declining postoperative adjuvant therapy. The protocol was

performed in compliance with the European Union Directives for the protection of animals used for scientific purposes (1999/575/CE and 2010/63/UE) and approved by the Ethics Committee of the Institute of Biomedical Sciences Abel Salazar, University of Porto. Tumours were accurately measured in the largest diameter with a calliper before surgery by the clinician (AdM or AS). Surgical techniques were decided on an individual basis and resulted in 47% radical, 26% regional and 17% local mastectomies.

Histological and stereological evaluations were performed in all the slides resulting from one slab of the tumour, sectioned following the largest cross diameter of the tumour. The histological diagnosis was performed by two observers (MS and PDP) using the criteria of the World Health Organization classification (Misdorp et al., 1999). The presence of peritumoural vascular invasion and regional lymph node metastases was recorded. The lymph node status was evaluated in routine slides and after immunolabelling with pancytokeratin AE1/AE3 and cytokeratin 14, as previously described by Matos et al. (2006).

The histological grading of nuclear pleomorphism was performed according to the Nottingham method criteria, considering luminal epithelial cells present in all the tumoural area. Briefly, nuclei were scored as: 1 if no visible increase in size and shape variability was detected compared to normal surrounding mammary epithelial cells; 2 when moderate variation in size and shape existed (nuclei were generally larger than the normal ones); 3 when a marked increased variation in size and shape was present, with very large and bizarre forms observed in at least one quarter of the tumour area (Fig. 1). Two observers (MS and PDP) performed the nuclear grading independently in order to determine the inter-observer agreement. For cases with scoring discrepancy, a

consensus was subsequently reached by reviewing the slides using a multi-head microscope.

Stereological analysis

For the stereological analysis, all the tumour area of the slides was considered. As tumour size varied from 0.3 cm to 15 cm, an entire slab across the largest cross diameter resulted in one to nine slides. A systematic random sampling approach for the selection of the fields was used in each slide, meaning that the first field of sampling was randomly selected and thereafter fields were sampled systematically by adjusting the distance between individual fields of vision roughly proportional to the overall area of the tumoural tissue. The stereological analysis was performed with a workstation comprising a microscope (BX-50 Olympus) equipped with a 100x oil immersion objective (Olympus Uplan), a CCD video camera (Sony) connected to a PC monitor, and a motorized stage (Prior) for stepwise displacements in x-y directions; the workstation was controlled by the software CAST-Grid (v1.5, Olympus).

The \bar{v}_V was estimated by the PSI method (Gundersen and Jensen, 1985). This parameter quantifies the nuclear size and pleomorphism, estimated with a test-system made of parallel lines bearing a systematic pattern of points (Fig. 2). Only the nuclear profiles of epithelial cells hit by one of these points were sampled. On these profiles, the line segments overlying the nucleus were measured (from boundary to boundary) (Fig. 2); this resulted in a length (l) that was used to estimate the \bar{v}_V as follows (Gundersen and Jensen, 1985):

$$\bar{v}_V(\text{nucleus}) = (\pi/3) \cdot \bar{l}_0^3$$

For estimating the \bar{v}_v , the average of total counted intercepts in each slide was 166 (\pm 69 standard deviation or SD).

Follow-up and survival study

The follow-up study was performed as described by Santos et al. (2011). Briefly, each female dog was evaluated prior to surgery, 3 weeks after the procedure and every 3 months, thereafter for a 2-year period. Each evaluation included a complete physical examination, thoracic radiographs (3 views) and an abdominal ultrasound. Any clinical signs or lesions that could be related to tumour progression, either in the scheduled evaluations or in between them was thoroughly investigated (e.g. fine-needle aspirate, ultrasound-guided biopsy, and skeletal radiography). A complete necropsy was performed when the animals died spontaneously or were euthanised in search for evidence of subclinical local recurrence or metastatic disease. All suspected lesions were confirmed by histopathology.

The survival analysis was performed following the most recent recommendations for prognostic studies in CMC (Matos et al., 2012), using only the cases presented as one malignant tumour per animal (n of animals = 40) and cases of multiple malignant tumours per animal with no evidence of disease progression (recurrence and/or metastases) during the follow-up period (n of animals = 8; n of tumours = 18). Overall survival (OS) was calculated from the date of surgery to the date of animal death/euthanasia due to tumour metastasis. Disease-free interval (DFI) was calculated from the date of surgery to the date of detection of tumour progression, i.e., confirmed local recurrence and/or metastases. In the survival study, cases were censored

if (1) animals died due to causes unrelated to tumour; (2) were lost to follow-up; and (3) were alive and free of distant metastases 24 months after surgery.

Statistical analysis

Cohen's Kappa statistic was used to assess the inter-observer agreement for the histological scoring of nuclear pleomorphism. An almost perfect and substantial agreement exists when the kappa value is comprised between 0.81 to 1 and 0.61 to 0.8, respectively. In contrast, when the kappa value is < 0.6 , the agreement is considered moderate to poor (Vieira and Garrett, 2005). Additionally, an asymptotic marginal homogeneity test was used, to evaluate if statistical differences between the observers existed. ANOVA test and Tukey's post-hoc multicomparison were used to assess for differences between the stereological values of \bar{v}_v in each nuclear pleomorphism consensual score. Receiver operating characteristic (ROC) curves analyses were used to determine if tumour size and \bar{v}_v values could discriminate tumours with or without evidence of vascular invasion and/or lymph node metastases.

The association between clinicopathological parameters (tumour size, histological type, nuclear pleomorphism consensual score, and \bar{v}_v) and the evidence of vascular invasion and/or lymph node metastases was assessed by logistic regression models. Pearson's Chi-square or Fisher test (when the assumptions for the Pearson Chi-square were not fulfilled) was used to assess which categorical variables should enter in the multivariable logistic model, while independent t test was used for continuous variables. In addition to the histological type, tumours were grouped for statistical purposes into two categories: (1) solid and anaplastic types and (2) tubulopapillary and complex types (one case of squamous cell carcinoma was excluded). This was

performed according to previous evidence that solid and anaplastic tumours show increased metastatic activity (Misdorp et al., 1999; Peña et al., 2013; Rasotto et al., 2012). As different categories of tumour size have been described in the literature, this parameter was analysed in three different ways: (1) as a continuous variable; (2) grouped according to the WHO criteria (< 3 cm; 3-5 cm; > 5cm); or (3) categorized as suggested by Santos et al. (2011) (< 3 cm; \geq 3 cm).

A survival analysis of DFI and ODS was performed to determine if differences existed between the group with and without evidence of vascular invasion/lymph node metastases. The Kaplan-Meier plots were used to show differences between these groups. Finally, a log-rank test (Mantel-Cox) was applied to analyse the significance of the differences between groups. The same approach was used for the results of the ROC analysis if significant. The association of nuclear pleomorphism consensual score and \bar{v}_v with survival data was also evaluated by univariable analyses. The counts of tumours with progression (recurrence and/or metastases and death) and tumours without events during the entire follow-up period were used for those univariable analyses. In all the tests, $P < 0.05$ was considered significant. Statistical analyses were performed with IBM SPSS Statistics 21.

Results

The 89 tumours analysed included 42 complex carcinomas, 23 solid carcinomas, 21 tubulopapillary carcinomas and 3 other carcinoma types (1 squamous cell and 2 anaplastic carcinomas). In 26% of the tumours there was histological evidence of vascular invasion and/or lymph node metastases. Survival data were available for 58 cases corresponding to 48 animals. From these, 10 animals progressed (recurred and/or

metastasized), 26 were disease-free at the end of the follow-up period, and the remaining were censored (5 died due to tumour unrelated causes, 5 were lost to follow-up and 2 are still being followed).

When scoring the nuclear pleomorphism, the two observers disagreed in 30/89 cases (33.7%), with higher agreement in the score 3 comparing to scores 1 and 2 (Table 1). The Kappa value was 0.46 ($P < 0.001$), thus indicating poor agreement. An asymptotic marginal homogeneity test was performed to evaluate whether there was a significant difference between the observers and none gave systematically lower or higher values than the other ($P = 0.16$).

In each subtype of carcinomas (except for those named 'other carcinomas'), score 2 was given to 50-60% of the cases. In solid and tubulopapillary carcinomas, score 3 represented 30-40% of the cases, while in complex carcinomas only 19% of cases were score 3. In the latter, 29% had score 1 in nuclear grade (Table 2). It is noteworthy that no association was detected between nuclear pleomorphism scores and histological types of CMC ($P = 0.26$).

Regarding the histological type, no significant association with evidence of vascular invasion and/or lymph node metastases was detected. However, when the tumours were grouped into two histological type categories (solid plus anaplastic and tubulopapillary plus complex), it was noticed that the latter presented a low risk of vascular invasion and/or lymph node metastases comparing to the other category ($P = 0.01$; Table 3). In general, the stereological analysis was simple and the PSI method

lasted 10 to 15 min per slide. The mean \bar{v}_v was $291 \pm 75 \mu\text{m}^3$ and was $267 \pm 50 \mu\text{m}^3$, $265 \pm 50 \mu\text{m}^3$ and $357 \pm 92 \mu\text{m}^3$ in tumours with nuclear scores 1, 2 and 3, respectively.

The ANOVA comparison between nuclear pleomorphism scores and \bar{v}_v revealed significant differences between score 1 vs. 3 ($P < 0.001$) and 2 vs. 3 ($P < 0.001$). Tumours scored as 1 and 2 did not differ significantly regarding their \bar{v}_v . The histogram of \bar{v}_v values showed a larger proportion of values around $300 \mu\text{m}^3$, mainly including tumours with nuclear scores 1 and 2; a much smaller peak appeared to exist after $400 \mu\text{m}^3$, corresponding to tumours that scored 3 in nuclear pleomorphism (Fig. 3).

In univariable analysis, the \bar{v}_v was not associated with histological evidence of vascular invasion and/or lymph node metastases. Therefore, in descriptive ROC analyses the discriminating power of the \bar{v}_v estimations between tumours with and without histological evidence of vascular invasion and/or lymph node metastases was very low (AUC statistics 0.55; 95% confidence interval or CI, 0.41–0.70). According to the ROC findings, no optimal cut-off value of nuclear \bar{v}_v , both with high sensitivity and specificity, could be defined for differentiating tumours with and without vascular invasion and/or lymph node metastases. The \bar{v}_v cut-off of $420 \mu\text{m}^3$ provided 96% specificity, but only 22% sensitivity. However, when the \bar{v}_v values of the tumours that progressed (recurred and/or metastasized) were compared to those that did not progress during the follow-up, it was noticed that the mean \bar{v}_v in the former was significantly higher ($393 \pm 74 \mu\text{m}^3$ versus $270 \pm 68 \mu\text{m}^3$; $P < 0.001$).

The nuclear pleomorphism score categories (score 1 plus score 2 versus score 3) were associated with evidence of vascular invasion and/or lymph node metastases ($P =$

0.02) and with tumour progression during the follow up ($P < 0.001$). In fact, only tumours scored as 2 or 3 showed histological evidence of vascular invasion and/or lymph node metastases and progression (in the latter, 9 out of 10 were scored 3, with only one case being nuclear score 2).

In multivariable analysis, histological type categories and nuclear pleomorphism score failed to retain their association with evidence of vascular invasion and/or lymph node metastases. Tumour size was the only factor with statistical significance as independent predictor of vascular invasion and/or lymph node metastases (Table 3).

In the descriptive ROC analysis, the discriminative power of tumour size was high (AUC statistics = 0.80; range, 0.70-0.91) (Fig. 4). According to the ROC findings, the optimal cut-off tumour size value that would distinguish tumours with and without vascular invasion and/or lymph node metastases was 2.9 cm (sensitivity of 74% and specificity of 77%).

Regarding the survival study, the mean DFI (14.2 ± 2.7 months) and OS (15.2 ± 2.6 months) of the dogs bearing tumours with histological evidence of vascular invasion and/or lymph node metastization were significantly shorter than the DFI (22.4 ± 0.9 months) and OS (23.2 ± 0.5 months) of the cases without such evidence (Fig. 5). The tumours were also grouped by the tumour size cut-off value previously determined by the ROC curve in order to perform a survival analysis (Fig.4). Tumours ≥ 2.9 cm in their maximum diameter were associated with poorer survival times ($P = 0.04$ for DFI and $P = 0.02$ for OS) (Fig.6).

Discussion

In this study, a comprehensive evaluation of the nuclear pleomorphism of CMC and its prognostic value was performed. The tumours were divided according to the presence or absence of vascular invasion and/or lymph node metastases at the time of histological diagnosis, as previously described by Rasotto et al. (2012). This could be viewed as one of the earliest morphological signs of tumour aggressive behaviour and we propose that its recognition is essential for the development of effective post-operative metastatic disease prevention strategies, like chemo- or radiotherapy. This assumption was supported by the survival analysis of this series, where tumours with histological evidence of vascular invasion and/or lymph node metastases had poorer survival after surgery, corroborating results from previous studies (Hellmén et al., 1993; Philibert et al., 2003; Chang et al., 2005).

The inter-observer variability in grading nuclear pleomorphism using the Nottingham method criteria had not, to the best of our knowledge, been previously studied in CMC. In some of the most recent reports focusing on the use of the Nottingham method for mammary tumours of dogs, the number of observers and their concordance were not reported (Karayannopoulou et al., 2005; Clemente et al., 2010, Peña et al., 2013). In this study, there was a low agreement between the two observers regarding nuclear pleomorphism scoring. At a first glance, this fact could be a reason for concern since it anticipates a probable low reproducibility of histological grading in CMC. As already mentioned, CMC are highly heterogeneous tumours and an inherent subjectivity associated with the selection of the highly malignant areas is likely to occur (Ladekarl, 1998; Misdorp, 2002; Kamaki et al., 2006; Rakha et al., 2010). Moreover, according to our stereological findings, tumours with nuclear scores 1 and 2 had

objectively comparable \bar{v}_v . Therefore, a lack of concordance between observers in these nuclear scores, as observed in this study, is more prone to occur. It is noteworthy that similar studies in human breast cancer concluded that the \bar{v}_v estimations in score 1 and score 2 were also similar (Artacho-Pérula and Roldán-Villalobos, 1997).

The stereological assessment of nuclear pleomorphism by the PSI method uses a systematic random sampling, which is not only highly efficient but also obviates the subjectivity associated to the selection of the apparently more aggressive tumour areas (Ladekarl, 1998). To the best of our knowledge, nuclear pleomorphism in CMC has never been quantified using unbiased stereological techniques. According to our statistical analysis, the unbiased assessment of nuclear pleomorphism seems not to be related to the histological evidence of vascular invasion and/or lymph node metastases in CMC. However, \bar{v}_v was significantly higher in tumours that recurred and/or metastasized during the 2-year follow-up period. This opens the possibility that this parameter may be regarded as a survival predictor like in humans, where it has been correlated with survival of patients affected by breast carcinoma (Artacho-Pérula and Roldán-Villalobos, 1997; Ladekarl, 1998).

Canine and human mammary neoplasms are morphologically distinct. Yet, the Nottingham grading method was adapted to the dog, taking for granted that the evaluation of the parameters should be the same as used in human pathology (Matos et al., 2012; Peña et al., 2013), and the three scores of nuclear pleomorphism were considered without any validation studies. The present study showed that tumours scored 1 or 2 in nuclear pleomorphism presented similar mean nuclear volumes and that there was a marked low inter-observer concordance in nuclear score 1. Moreover,

nuclear score 3 was associated to a higher probability of vascular invasion and/or lymph node metastases, and poor survival. Therefore, only two nuclear grading scores appear to exist in CMCs: low grade and high grade, corresponding mostly to Nottingham nuclear scores 1 and 2 and to nuclear score 3, respectively.

There is a general agreement that the tumour size of canine mammary tumours has a prognostic significance (Philibert et al., 2003; Chang et al., 2005; Sorenmo et al., 2011). Our multivariable analysis results confirmed that tumour size represents an independent prognostic factor regarding the capacity of tumour vascular invasion and metastasis. Keeping in mind that different tumour size categories have been proposed and used (Misdorp, 2002; Santos et al., 2011; Peña et al., 2013), a ROC curve analysis was performed in order to define the tumour size threshold value with the highest sensitivity and specificity to predict aggressive tumour behaviour. Our results highlighted that tumours > 2.9 cm in their largest diameter were more associated to vascular system invasion and/or metastasis to the regional lymph nodes. The established cut-off of 2.9 cm supports that the tumour size categories (3 cm < and > 3 cm), defined by Sorenmo (2003) and used by Santos et al. (2011), are suitable for CMC and should be routinely applied for better reproducibility of future studies. Moreover, it indirectly supports the model proposed by Sorenmo et al. (2009) that correlates the development of a malignant phenotype with increased tumour size. According to our findings, tumour size does matter and clinicians should consider more radical surgical procedures and meticulous periodical follow-up evaluations of animals affected by malignant tumours > 3 cm with a special regard to the surgical removal of regional lymph nodes.

Unlike tumour size, nuclear pleomorphism was not an independent prognostic value when included in the multivariable model. This analysis obviates possible confounding factors when assessing prognostic factors and shows that some parameters, although related to prognosis in univariable analysis, lose their value in multivariable models (Peña et al., 2013; Santos et al., 2013). We think that the role of nuclear pleomorphism as a prognostic factor deserves to be studied in larger series by multivariable survival analysis.

A stereological tool was used to evaluate nuclear pleomorphism in CMCs. The estimates of \bar{v}_v did not clearly discriminate tumours of the three different Nottingham nuclear scores. Some evidence that higher tumoural \bar{v}_v could predict a poor animal outcome was provided by the present study. However, the analysed sample with follow-up data, particularly the number of cases with tumour-related events, was relatively small, hampering further statistical analysis (e.g., a ROC curve analysis of the \bar{v}_v values in tumours that progressed and tumours that did not progressed during the follow-up period). In addition, the role of this stereological evaluation of nuclear volume in other subtypes of canine mammary malignant tumours, like carcinosarcomas, remains to be determined.

Conclusions

This study provide evidence that the nuclear pleomorphism evaluation in CMC using the Nottingham method criteria is highly prone to inter-observer variability. Prospective studies of larger series and rigorous follow-up data are needed to validate the possibility that only two nuclear pleomorphism scores should be considered in

CMC. The \bar{v}_v could provide an objective threshold allowing the differentiation between low and high nuclear grade, based on reliable survival information.

Conflict of interest statement

None of the authors has any financial or personal relationships that could inappropriately influence or bias the content of the paper.

References

Artacho-Pérula E., Roldán-Villalobos R., 1997. Unbiased stereological estimation of the number and volume of nuclei and nuclear size variability in invasive ductal breast carcinomas. *Journal of Microscopy* 186, 133-142.

Chang S.C., Chang C.C., Chang T.J., Wong M.L., 2005. Prognostic factors associated with survival two years after surgery in dogs with malignant mammary tumors: 79 cases (1998-2002). *Journal of American Veterinary Medical Association* 227, 1625-1629.

Clemente M., Pérez-Alenza M.D., Illera J.C., Peña L., 2010. Histological, immunohistological, and ultrastructural description of vasculogenic mimicry in canine mammary cancer. *Veterinary Pathology* 47, 265-274.

Elston C.W., Ellis I.O., 1998. Assessment of histological grade. In: *The breast*. Elston, C.W. and Ellis, I.O. (eds.), Vol. 13. Churchill Livingstone, Edinburgh; New York, USA, pp. 356-384.

Fujikawa K., Matsui Y., Kobayashi T., Miura K., Oka H., Fukuzawa S., Takeuchi H., 2000. Predicting pathological stage of localized prostate cancer using volume weighted mean nuclear volume. *Journal of Urology* 164, 1587-1590.

Goldschmidt M., Peña L., Rasotto R., Zappulli V., 2011. Classification and grading of canine mammary tumors. *Veterinary Pathology* 48, 117-131.

Gundersen H.J., Jensen E.B., 1985. Stereological estimation of the volume-weighted mean volume of arbitrary particles observed on random sections. *Journal of Microscopy* 138, 127-142.

Hellmén E., Bergström R., Holmberg L., Spångberg I.B., Hansson K., Lindgren A., 1993. Prognostic factors in canine mammary tumors: a multivariable study of 202 consecutive cases. *Veterinary Pathology* 30, 20-27.

Kamaki K., Sano N., Tangoku A., 2006. Problems in histological grading malignancy and its clinical significance in patients with operable breast cancer. *Breast Cancer* 13, 249-253.

- Karayannopoulou M., Kaldrymidou E., Constantinidis T.C., Dessiris A., 2005. Histological grading and prognosis in dogs with mammary carcinomas: application of a human grading method. *Journal of Comparative Pathology* 133, 246-252.
- Klopfleisch R., von Euler H., Sarli G., Pinho S.S., Gärtner F., Gruber A.D., 2011. Molecular carcinogenesis of canine mammary tumors: news from an old disease. *Veterinary Pathology* 48, 98-116.
- Ladekarl M., 1998. Objective malignancy grading: a review emphasizing unbiased stereology applied to breast tumors. *APMIS* 79, 1-34.
- Ladekarl M., Sørensen F.B., 1993. Quantitative histopathological variables in *in situ* and invasive ductal and lobular carcinomas of the breast. *APMIS* 101, 895-903.
- Lana S.E., Rutteman G.R., Withrow S.J., 2007. Tumors of mammary gland. In: *Small Animal Oncology*. Withrow S.J., MacEwen E.G. (eds.), Saunders Elsevier, Fourth Ed., St. Louis, pp. 619-636.
- Matos A.J.F., Faustino A.M.R., Lopes C., Rutteman, G.R., Gärtner F., 2006. Detection of lymph node micrometastasis in canine malignant mammary tumors with the use of cytokeratin immunostaining. *Veterinary Record* 158, 626-629.
- Matos A.J., Baptista C.S., Gärtner M.F., Rutteman G.R., 2012. Prognostic studies of canine and feline mammary tumours: the need for standardized procedures. *The Veterinary Journal* 193, 24-31.
- Misdorp W., 2002. Tumors of mammary gland. In: *Tumors of Domestic Animals*. Meuten D.J. (ed.), Iowa State Press, Fourth Ed., Iowa, pp. 575-606.
- Misdorp W., Else R.W., Hellmén E., Lipscomb T.P., 1999. Histological classification of mammary tumors of the dog and the cat. *World Health Organization International Histological Classification of Tumours of Domestic Animals*, 2nd edition series, volume VII. Washington DC: Armed Forces Institute of Pathology.
- Peña L, De Andrés PJ, Clemente M, Cuesta P, Pérez-Alenza MD., 2013. Prognostic value of histological grading in noninflammatory canine mammary carcinomas in a prospective study with two-year follow-up: relationship with clinical and histological characteristics. *Veterinary Pathology* 50, 94-105.
- Philibert J.C., Snyder P.W., Glickman N., Glickman L.T., Knapp D.W., Waters D.J., 2003. Influence of host factors on survival in dogs with malignant mammary gland tumors. *Journal of Veterinary Internal Medicine* 17, 102-106.
- Rakha E.A., Reis-Filho J.S., Baehner F., Dabbs D.J., Decker T., Eusebi V., Fox S.B., Ichihara S., Jacquemier J., Lakhani S.R., Palacios J., Richardson A.L., Schnitt S.J., Schmitt F.C., Tan P.H., Tse G.M., Badve S., Ellis I.O., 2010. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Research* 12, 207-219.
- Rasotto R., Zappulli V., Castagnaro M., Goldschmidt M.H., 2012. A retrospective study of those histopathologic parameters predictive of invasion of the lymphatic system by canine mammary carcinomas. *Veterinary Pathology* 49, 330-340

Santos A., Lopes C., Marques R.M., Amorim I., Ribeiro J., Frias C., Vicente C., Gärtner F., de Matos A., 2011. Immunohistochemical analysis of urokinase plasminogen activator and its prognostic value in canine mammary tumours. *The Veterinary Journal* 189, 43-48.

Santos A.A., Lopes C.C., Ribeiro J.R., Martins L.R., Santos J.C., Amorim I.F., Gärtner F., Matos A.J., 2013. Identification of prognostic factors in canine mammary malignant tumours: a multivariable survival study. *BMC Veterinary Research* 9:1. doi: 10.1186/1746-6148-9.

Soda T., Fujikawa K., Ito T., Sasaki M., Nishio Y., Miyakawa M., 1999. Volume-weighted mean nuclear volume as a prognostic factor in renal cell carcinoma. *Laboratory Investigation* 79, 859-867.

Sorenmo K., 2003. Canine mammary gland tumors. *Veterinary Clinical North America Small Animal Practice* 33, 573-596.

Sorenmo K.U., Kristiansen V.M., Cofone M.A., Shofer F.S., Breen A.M., Langeland M., Mongil C.M., Grondahl A.M., Teige J., Goldschmidt M.H., 2009. Canine mammary gland tumours; a histological continuum from benign to malignant; clinical and histopathological evidence. *Veterinary Comparative Oncology* 7, 162-172.

Sorenmo K.U., Rasotto R., Zappulli V., Goldschmidt M.H., 2011. Development, anatomy, histology, lymphatic drainage, clinical features, and cell differentiation markers of canine mammary gland neoplasms. *Veterinary Pathology* 48, 85-97.

Sørensen F.B., 1992. Quantitative analysis of nuclear size for objective malignancy grading: a review with emphasis on new, unbiased stereologic methods. *Laboratory Investigation* 66, 4-23.

Steiner A., Binder M., Mossbacher U., Wolff K., Pehamberger H., 1994. Estimation of the volume-weighted mean nuclear volume discriminates Spitz's nevi from nodular malignant melanomas. *Laboratory Investigation* 70, 381-385.

Vieira, A.J. and Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. *Family Medicine* 37, 360-363.

Yörükoglu K., Aktas S., Güler C., Sade M., Kirkali Z., 1998. Volume-weighted mean nuclear volume in renal cell carcinoma. *Urology* 52, 44-47.

Tables

Table 1: Results and concordance of the nuclear grading by the two observers.

		Nuclear grade observer A			Total
		1	2	3	
Nuclear grade observer B	1	8	4	0	12
	2	10	29	9	48
	3	1	6	22	29
	Total	19	39	31	89

Table 2: Nuclear grade scores, mean and standard deviation of volume-weighted mean nuclear volume (\bar{v}_v) and frequency of vascular invasion and/or lymph node metastases in each carcinoma subtype ('others' include 1 squamous cell and 2 anaplastic carcinomas).

Histological type				
Nuclear score	Solid (<i>n</i> =23)	Tubulopapillary (<i>n</i> =21)	Complex (<i>n</i> =42)	Others (<i>n</i> =3)
1	2	2	12	0
2	12	12	22	2
3	9	7	8	1
\bar{v}_v (μm^3)	311(\pm 98)	295(\pm 82)	282(\pm 49)	248(\pm 116)
Invasion/metastases	11	5	7	0

Table 3: Results of the univariable models and the final multivariable model using vascular invasion and/or lymph node metastases as the dependent variable in 89 canine mammary carcinomas.

<i>Independent variables</i>	Univariable analysis			Multivariable analysis		
	OR	95% CI	P	OR	95% CI	P
Histological type						
<i>All categories</i>						
<i>TP</i>	Ref		—			
<i>Solid</i>	2.94	0.80 - 10.74	0.103			
<i>Complex</i>	0.64	0.17 - 2.32	0.498			
<i>Others</i>	0	0 - >100	0.991			
<i>Two categories</i>						
<i>Solid and anaplastic</i>	Ref		—			
<i>TP and complex</i>	0.28	0.10 - 0.77	0.014	0.31	0.09 - 1.07	0.065
Tumour size						
<i>Continuum (cm)</i>	1.50	1.21 - 1.87	<0.001	1.63	1.26 - 2.10	<0.001
<i>WHO categories</i>						
< 3cm	Ref		—			
3 cm> and < 5 cm	8.67	2.24 - 33.53	0.002			
> 5cm	12.43	3.47 - 44.43	<0.001			
<i>Santos et al., 2011 categories</i>						
< 3cm	Ref		—			
≥ 3cm	9.58	3.20 - 28.72	<0.001			
Nuclear pleomorphism						
<i>Score 3</i>	Ref		—			
<i>Score 2</i>	0.42	0.15 - 1.20	0.101			
<i>Score 1</i>	0	>0.001 - 0	0.991			
<i>Two categories</i>						
<i>Score 1 and 2</i>	Ref		—			
<i>Score 3</i>	3.42	1.26 - 9.29	0.017	2.57	0.76 - 8.63	0.126
Nuclear \bar{v}_v (μm^3)	1.004	0.99 - 1.01	0.248			

OR, odds ratio; CI, confidence interval; TP, tubulopapillary; \bar{v}_v , volume-weighted mean nuclear volume, Ref, referent;

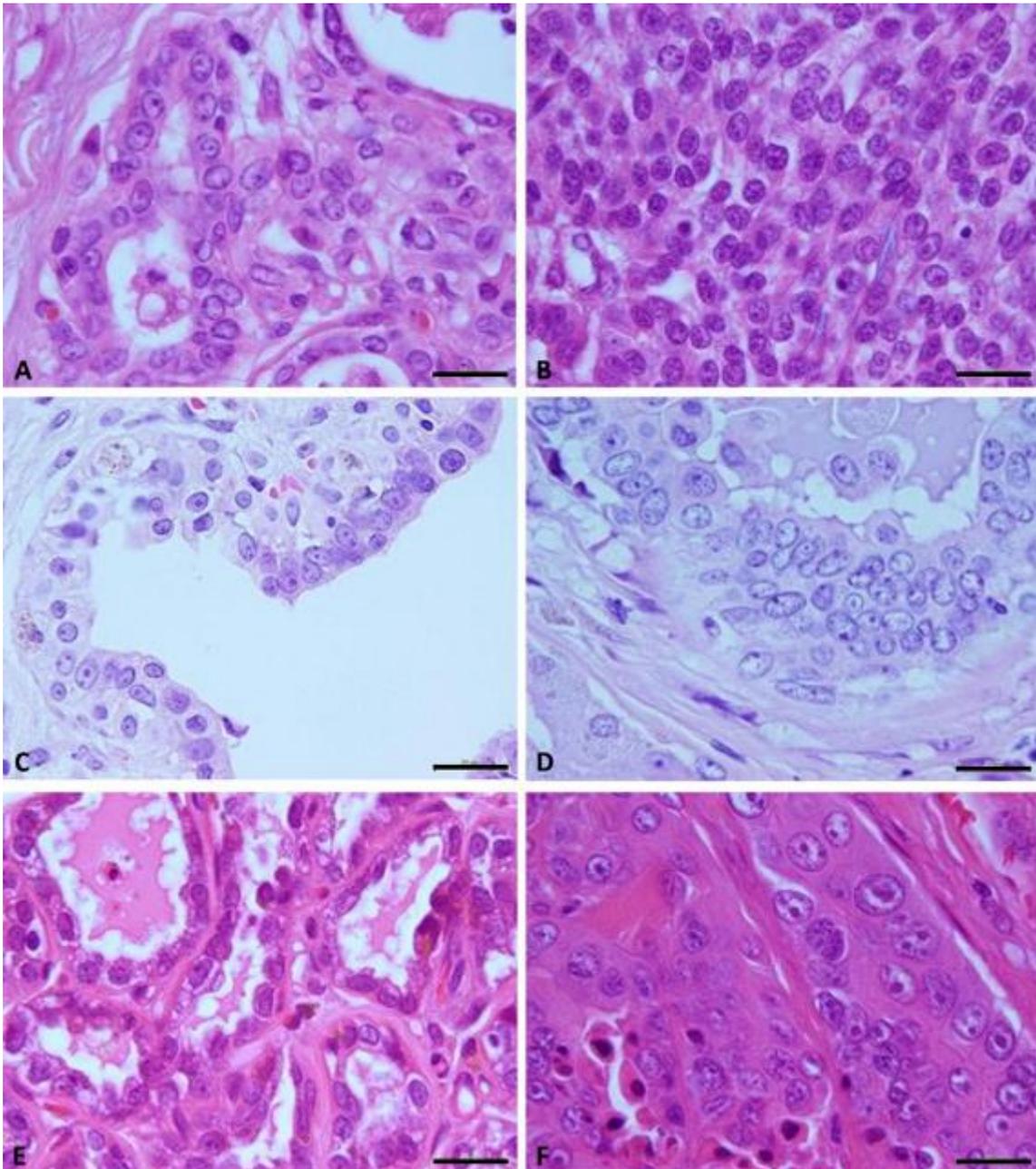


Figure 1: Histological normal surrounding mammary tissue (A, C, E) and tumour cells (B, D, F) in three cases of canine mammary carcinomas. Nuclear pleomorphism scoring followed the Nottingham method criteria: score 1 (image B) presents no increased size and shape variability comparing to normal epithelial cells (image A); score 2 (image D) has moderate variation and larger nuclei than normal ones (image C); whereas score 3 (image F) shows marked increased variation in size and shape, with very large nuclei and prominent nucleoli comparing with normal adjacent parenchyma (image E). Haematoxylin-eosin, Bar, 20 μ m.

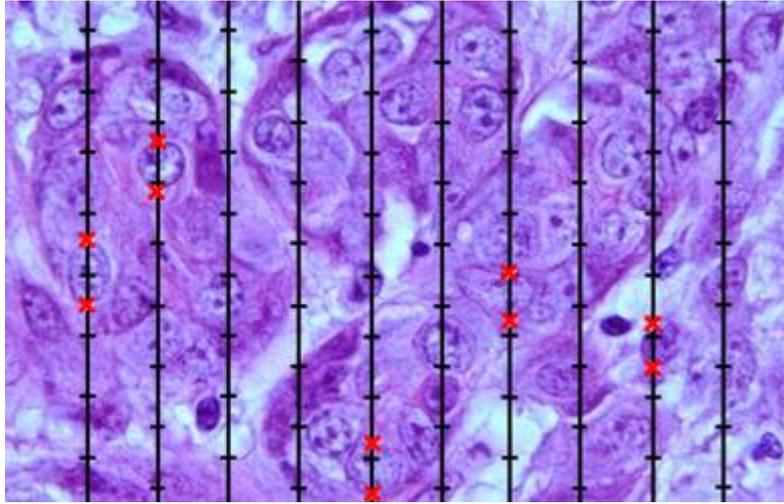


Figure 2: Estimation of the volume-weighted mean nuclear volume (\bar{v}_v) with the point sampled intercepts method. A test system made of parallel lines bearing a systematic pattern of points is generated by the software. Nuclei in focus that are hit by one of the points are selected. In these nuclei, the distance between the intersections of the nuclear boundaries with the lines (marked with red crosses) was measured. Haematoxylin-eosin, Bar, 13 μ m.

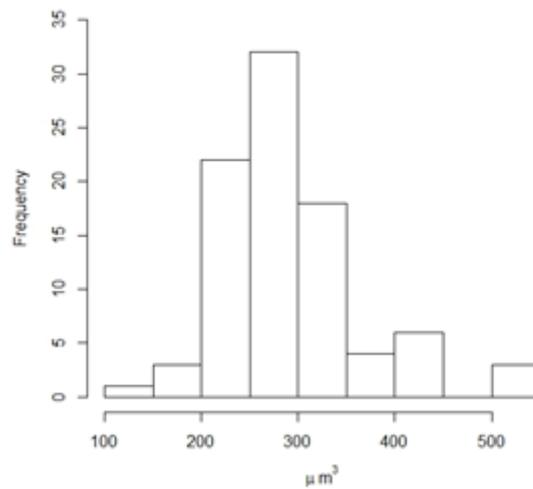


Figure 3: Histogram of the volume-weighted mean nuclear volume (\bar{v}_v) values of the 89 canine mammary carcinomas. The higher peak of the values is around 300 μm^3 .

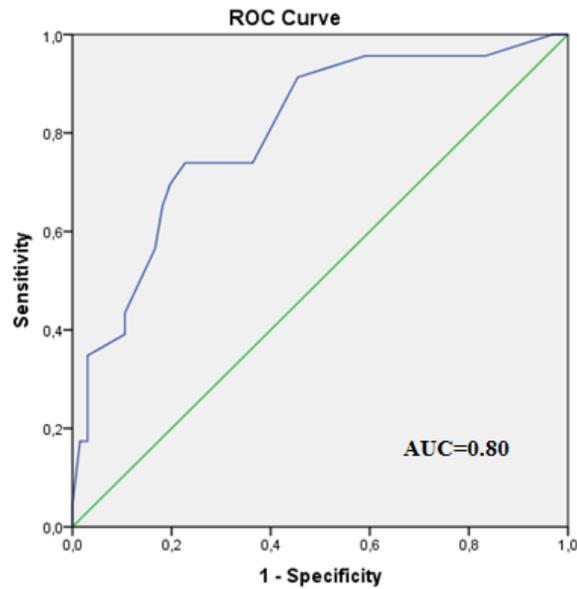


Figure 4: Receiver operating characteristics (ROC) curve for tumour size values. The area under the curve (AUC) is 0.80 which indicates that tumour size has high discriminatory power between tumours with and without evidence of vascular invasion and/or lymph node metastases. An optimal cut-off of 2.9 cm with 74% sensitivity and 77% specificity was defined.

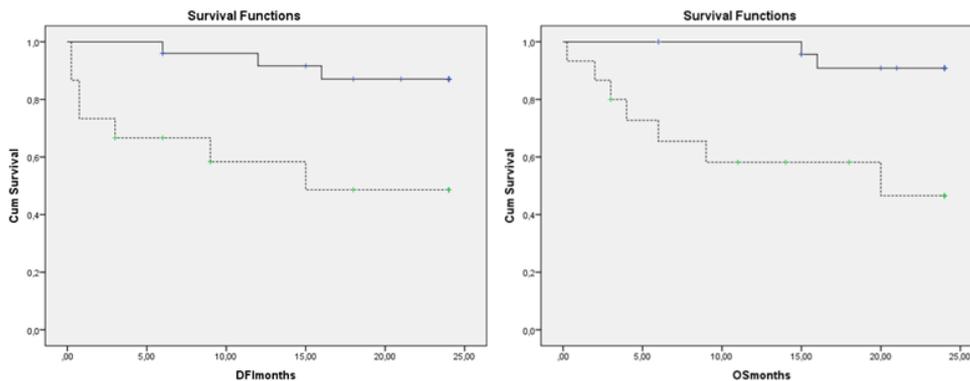


Figure 5– Kaplan-Meier disease free interval (DFI) and overall survival (OS) curves comparing cases with histological evidence of vascular invasion and/or lymph node metastases (dashed line) and cases without that evidence (continuous line) in 40 female dogs presenting a single mammary carcinoma. Vertical marks represent censored cases in each group. Histological evidence of vascular invasion and/or lymph node metastases was significantly associated to lower DFI and OS ($P = 0.004$ and $P = 0.001$, respectively).

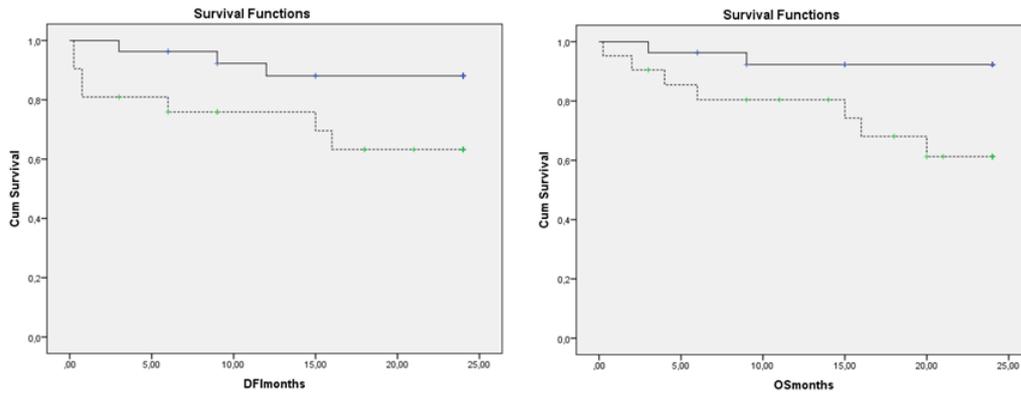


Figure 6 – Kaplan-Meier disease-free interval (DFI) and overall survival (OS) plots. Tumours were grouped on the basis of optimal cut-off size value determined by ROC curve analysis: size < 2.9 cm (continuous line) and size \geq 2.9 cm (dashed line). Vertical marks represent censored cases in each group. Size \geq 2.9 cm was associated to shorter DFI ($P = 0.04$) and OS ($P = 0.02$).