Original article manuscript 1

Which Dogs with Appendicular Osteosarcoma Benefit Most from Chemotherapy after Surgery? Results from an Individual Patient Data Meta-Analysis.

A F Schmidt[a,b,c,d*], R H H Groenwold[a,b], P Amsellem[e], N Bacon[g], O H Klungel[a,b], A W Hoes[a], A de Boer[b], K Kow[d], K. Maritato[f], J Kirpensteijn[h], M Nielen[c].

[a.] *Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, P.O. Box 85500, 3508 GA Utrecht, the Netherlands.*
[b.] *Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, P.O. Box 80082, 3508 TB Utrecht, The Netherlands.*
[c.] *Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, Utrecht 3584 CL, The Netherlands.*
[d.] *Institute of Cardiovascular Science, Faculty of Population Health, University College London, London WC1E 6BT, The United Kingdom.*
[e.] *Department of Companion Animals, Atlantic Veterinary College University of Prince Edward Island, Canada.*
[f.] *MedVet Medical and Cancer Centers for Pets, Cincinnati, Ohio, United States of America.*
[g.] *Department of Small Animal Clinical Sciences, University of Florida, Gainesville, Florida, United States of America.*
[h.] *Department of Clinical Sciences of Companion Animals, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 8, Utrecht 3584 CM, The Netherlands, for the VSSO investigators.*

***** *Corresponding author: Tel.: +44 20 3549 5625.*
E-mail address: amand.schmidt@ucl.ac.uk (A.F.Schmidt).

Target journal: Preventive Veterinary Medicine.

Running title: Chemotherapy in Canine Osteosarcoma

Word count text: 3412

Word count abstract: 167

Number of references: 40

Number of tables: 4

Number of figures: 3

(Web)appendices: 3

Abstract

Osteosarcoma (OS) is a malignant tumor of mesenchymal origin that produces osteoid. Given that the prognosis varies considerably between dogs, we explored whether treatment could be tailored towards prognostic subgroups of patients.For the current study, individual patient data from five nonrandomized studies were combined. Based on a multivariable prognostic model, the 5-month mortality risk was estimated. Subsequently, in surgically treated dogs, we explored whether 'any chemotherapy' compared to no chemotherapy influenced their 5-month mortality risk. After adjustment for potential confounders the main effect of any chemotherapy was odds ratio 0.48 (95%CI 0.30; 0.78). Testing for chemotherapy by 5-month mortality risk interaction revealed that the effects of any chemotherapy decreased with increasing risk, P-value = 0.04, indicating that dogs with a lower risk of 5-month mortality benefited most from chemotherapy. Results from individually comparing carboplatin, cisplatin, doxorubicin and doxorubicin combination therapy to no chemotherapy, were similar in magnitude and direction. These results indicate that the main treatment effects of chemotherapy do not necessarily apply to all patients.

MeSH/keywords: Canine; Personalized Medicine; Oncology, Bone tumour.

Introduction

Osteosarcoma (OS) is a malignant tumor of mesenchymal origin that produces osteoid. In dogs, OS most frequently occurs in large and giant breeds (Cooley and Waters, 1997; McNeill et al., 2007; Norrdin et al., 1989; Ru et al., 1998; Spodnick et al., 1992). Dogs that are treated with amputation have a median survival time of five months, with the majority succumbing to metastatic disease (Brodey and Abt, 1976; Straw and Withrow, 1996). Clinical studies have shown that on average survival in OS dogs can be extended by administrating chemotherapy (Bailey et al., 2003; Chun et al., 2005; Chun et al., 2000; Straw et al., 1991; Vail et al., 2002).

(Boerman et al., 2012)In a recent Individual Patient Data Meta-Analysis (IPDMA), we identified baseline variables that were associated with survival in dogs with osteosarcoma (Schmidt et al., 2013). Such a prognostic model can be used to predict a dog's risk of early mortality (Moons et al., 2012). This offers the possibility to identify subgroups of dogs according to their baseline prognosis and target treatment at those patients most likely to benefit. This can potentially prevent dogs from unnecessarily receiving treatment, which is relevant in terms of both costs and quality of life. Clearly, there is a need to obtain estimates of individualized treatment effects (Hayward et al., 2006; Kent et al., 2010; Rothwell and Warlow, 1999).

In the current paper, (Schmidt et al., 2013), chemotherapy effects were individualized by determining whether dogs with a different 5-month mortality risk, reacted differently to chemotherapy treatment. Specifically, using an adapted version of the previously published prediction model (Schmidt et al., 2013), we first predicted a dog's 5-month mortality risk based on age, weight, gender neuter status, serum alkaline phosphatase (SALP) level, breed,

and tumor location at time of surgery. Subsequently we evaluated what the effect was of "any chemotherapy" compared to no chemotherapy on the 5-month mortality incidence and if this effect differed between dogs with different predicted 5-month mortality risks. Finally, we repeated the analysis with separated groups for carboplatin, cisplatin, doxorubicin and doxorubicin combination therapy and estimated the effect on 5-month mortality incidence compared to dogs receiving no chemotherapy.

Materials and Methods

The effects of the different chemotherapeutics compared to no chemotherapy were determined using individual patient data (IPD). These IPD were used previously in an IPD meta-analysis (IPDMA) combining data of 20 studies to determine prognostic factors for early mortality in dogs with osteosarcoma (Schmidt et al., 2013). A detailed description of the data accrual can be found in the original publication (a review protocol is unavailable). Briefly, studies were collected via the Veterinary Society of Surgical Oncology (VSSO). In January 2012, a call for collaboration was sent out to VSSO members and other veterinary oncologic researchers. Data was deemed eligible if baseline patient characteristics of OS dogs and time to event (death or metastasis) were recorded. To reduce the possibility of publication bias (Easterbrook et al., 1991), published and unpublished studies were both eligible. All dogs in these studies were diagnosed with osteosarcoma. For the present analysis, dogs were excluded if they did not receive surgery; due to euthanasia (n = 197), who received limb-sparing surgery (n = 41), who received an infrequently used chemotherapeutic protocol (n = 13) or who received radiation therapy (n =11). Additionally, the study by Sottnik (Sottnik et al., 2010) only collected data on metastasis, not mortality, and was excluded. Data was available from 1295 dogs collected in 16 studies.

To answer our present questions, does chemotherapy effectiveness differ between dogs with different predicted 5-month mortality risk, we used the 1295 dogs to construct a logistic regression prediction model; predicting mortality at 5 months. Subsequently, from these 1295 dogs (16 studies), studies were selected that included at least five dogs on no chemotherapy and at least five dogs treated with one of the interventions of interest (i.e., carboplatin, cisplatin, doxorubicin or doxorubicin combination therapy). Five studies fulfilled this criterion; of these 5 studies, two were previously published (Amsellem et al., 2014; Kirpensteijn et al., 2002; Kow et al., 2008). After excluding dogs that received lobaplatin chemotherapy (n = 27) 400 subjects remained. Regrettably, none of these 5 studies randomly allocated chemotherapy hence chemotherapy associations are presumably confounded; an issue that will be addressed later. We will first briefly describe how the logistic regression prediction model was derived (using the 1295 dogs). Second, we describe in detail how the 5-month mortality risk was calculated for each individual dog. Third, we explain how individualized chemotherapy effect estimates were derived (based on the 400 dogs). Finally, a number of sensitivity analyses are discussed. Note that this study focused on 5-month mortality, because this is regarded as a clinical relevant endpoint (Brodey and Abt, 1976; Spodnick et al., 1992; Straw et al., 1991).

Data analysis: prediction model

Instead of using the Cox's proportional hazards prediction model described in Schmidt et al., (2013) to predict an individual dog's risk of dying at 5 months, the current analysis uses a prediction model based on a logistic regression model with random intercept for study. This prediction model used data from the previously described 1295 dogs and regressed a 5-month mortality indicator on the predictor's gender, neuter status, tumor location (proximal humerus, distal femur or proximal tibia, distal radius, versus other locations), age (years, continuous),

weight (kg, continuous), breed (Rottweiler, Golden Retriever, Labrador Retriever, Greyhound, Doberman, mixed breeds, versus other breeds) and serum alkaline phosphatase (SALP, using study specific cut-off values for high and normal SALP levels). Chemotherapy was included as a nuisance variable and was set to zero (no chemotherapy) when predicting the 5-month mortality risk. As in the original publication, all predictors were predefined and no model selection was used (Schmidt et al., 2013). However, linearity of the continuous predictors was assessed by comparing a model (using a likelihood ratio test) with restricted cubic splines (5 knots) to a model forcing linearity. Additionally, restricted cubic spline plots were created to visually inspect linearity. Besides, SALP which was dichotomized, no deviations from linearity were observed (Refer to Table 1 for the derived prediction model based on 1295 dogs with). additional(Chatfield, 1995)

In the following section, the performance of this prediction model is described. Discrimination, measured as a c-statistic, was 0.63 (95%CI 0.59; 0.67) indicating that the model could not perfectly discriminate survivors from those that died. However, calibration-in-the-large was 0.0005, indicating that the predicted and observed 5-month mortality risk agreed on average (p-value = 0.99). The calibration slope of 0.77 (95%CI 0.55; 1.00), showed that the predicted risk was too small for dogs with a low observed risk and too large for dogs with a higher observed risk, however these discrepancies were small (Appendix Figure A). All performance measures were corrected for optimism using 100 bootstrap samples (Steyerberg, 2009; Steyerberg et al., 2010).

In these 1295 dogs about 8% of the data was missing, information on 5-month mortality was missing for 4.2% of the observations and chemotherapy for 2.4% of the observations (see for more details Schmidt et al., 2013). Univariable tests showed that

missingness was associated with observed variables (results available from the first author) biasing a complete case analysis (Altman and Bland, 2007; Rubin, 1976). To adjust bias due to missing data, this dependency was taken into account by imputing missing observation using the aregImpute algorithm from the Hmisc package version 3.13-0 (Harrell, Jr. and Dupont, 2013). The aregImpute algorithm was implemented using 10 burn-in iterations, 100 approximate bootstrap samples and predictive mean matching. To get correct estimates of the standard errors 100 imputed datasets were created (i.e., multiple imputation). Results over all 100 imputed datasets were pooled using Rubin's rules (Little and Rubin, 2002; Marshall et al., 2009)

Data analysis: predicting 5-month mortality

An individual dogs' risk of 5-month mortality, under no chemotherapy, was predicted using the coefficient presented in Table 1 and by setting the chemotherapy to zero (no chemotherapy):

$$logit(5-month\ mortality\ risk) = logit(\hat{p}_i) = \hat{\beta}_0 + \hat{\beta}_1 * chemotherapy(0) + \ldots + \hat{\beta}_j x_{ij} \quad \text{[equation 1]}$$

Here $\hat{p}_i$ indicates an individuals' risk of being dead at 5 months. $\hat{\beta}_j$represent the coefficient for the jth variable presented in Table 1, note that $j \neq \{0,1,2,3,4,5\}$. Finally, x_{ij} represents an individuals' value for the jth variable. Note, that this is equal to calculating the linear predictors conditional on no chemotherapy. ,For these predictions, the random intercept (from the logistic regression model) was ignored because in clinical practice this prediction model would not be limited to the studies included in our meta-analysis. The logit(5-month mortality risk) can vary from minus to plus infinity, with zero referring to a risk of 50%. This logit(5-

month mortality risk) can be transformed to the 5-month mortality risk, bounded between 0 and 1, by the following equation:

$$\hat{p}_i = \frac{1}{1+e^{-logit(\hat{p}_i)}}$$ [equation 2]; see Table 1 for an example.

Applying equation 2 results in an estimate of the risk of 5-month mortality given that the patient did not receive chemotherapy. Note that because the logit(5-month mortality risk) is expected to be linearly related with the outcome this was included in all models. Where appropriate the logit(5-month mortality risk) was transformed to the 5-month mortality risk using equation 2; for example when graphing results.

Data analysis: estimating chemotherapy effectiveness

As indicated previously, first the association of "any chemotherapy" compared to no chemotherapy with 5-month mortality was estimated. If this association was significant we determined how the different chemotherapeutics carboplatin, cisplatin, doxorubicin or doxorubicin combination compared to no chemotherapy. These analyses used the previously defined subset of 400 subjects, collected by combining 5 studies (see Appendix Table A). Depending on the comparison, a different subset of these 5 studies was used (see Appendix Tables B through E): for any chemotherapy all studies were used, similarly for doxorubicin combination all studies were used, for carboplatin the study by Kirpensteijn was excluded, for doxorubicin the studies by Kirpensteijn and Amsellum were excluded and finally for the cisplatin comparison only the study by Bacon was used. This selection was based on whether the studies included any dog on the mentioned chemotherapeutic and prevents bias due to study specific influences.

Before determining whether chemotherapy effects differed between dogs with a different logit(5-month mortality risk) we first estimated the main effect of chemotherapy (i.e., a model regressing 5-month mortality on chemotherapy and co-variables without an interaction term with chemotherapy). These main effects provide an estimate of the average effect of chemotherapy and were derived using a logistic regression model including a random intercept for study. Specifically, a model was fitted, regressing 5-month mortality on the chemotherapy variable and a random intercept for study. A second model additionally included gender, neuter status, tumor locations, age, weight and SALP. The third model additionally adjusted for breed. These variables were included in an attempt to adjust for confounding and were selected based on prior knowledge (Hernan et al., 2002). To reduce the risk of residual confounding(Bland and Altman, 1995), no model reduction strategy was employed (i.e., backward selection) and no differentiation was made between predictors of the outcome and confounders.

After determining the main effects, we explored whether chemotherapy effectiveness depended on logit(5-month mortality risk) by including a chemotherapy by logit(5-month mortality risk) interaction term in the model.

$$logit(prob[y_{is} = 1]) = \hat{\alpha}_{0s} + \hat{\alpha}_1 * chemotherapy + \hat{\alpha}_2 * logit(\hat{p}_i) + \hat{\alpha}_3 * chemotherapy * logit(\hat{p}_i) + \epsilon_{is}$$

Here y_i represent an individual's mortality status at 5-months. Let $\hat{\alpha}_1$represent the estimated association of chemotherapy compared to no chemotherapy when all co-variables are zero [i.e., when the logit(5-month mortality risk) = 0], $\hat{\alpha}_2$ the association of logit(5-month

mortality risk) when a patients does not receive chemotherapy, and $\hat{\alpha}_3$ the association of the chemotherapy by logit(5-month mortality risk) interaction term. $\hat{\alpha}_3$ indicates how much the association of chemotherapy changes per unit increase or decrease of the logit(5-month mortality risk). In the absence of interaction, $\hat{\alpha}_3$ becomes zero and could be omitted. Note that epsilon indicates the amount of residual error from the intercept $\hat{\alpha}_{0s}$ per individual i and per study s.(Arbogast et al., 2008). Using these estimates an individualized effect of chemotherapy can be calculated:

$$\widehat{OR}_i = e^{\hat{\alpha}_1 * chemotherapy(1) + \hat{\alpha}_3 * chemotherapy(1) * logit(\hat{p}_i)} \quad \text{[equation 3]}$$

Here OR represent the estimated odds ratio of chemotherapy for the ith individual. This methodology has been previously applied in human medicine most notably in the SYNTAX trial (Farooq et al., 2013; van Klaveren D. et al., 2015). Note, that $logit(\hat{p}_i) = logit(5 - month\ mortality\ risk)$ and is calculated using equation 1.

For the subset of 400 dogs on average 12.6% of the information was missing; specifically, 5-month mortality 9%, chemotherapy 7.5%, tumor location 9.5%, gender 5.3%, neuter status 5.3%, age 6.5%, weight 26.3%, high SALP 57,5% and breed 4.5%, (see Table 2 for an overview). Again missing values were imputed as previously described.

Data analysis: sensitivity analyses

In the following section we describe a few sensitivity analyses evaluating the appropriateness of assumptions made.

Throughout a logistic regression model was used including a random intercept for study. Such a model assumes that the random intercept can sufficiently be described by a normal distribution. To evaluate this assumption a regular logistic regression model was used including study as a categorical factor, which does not assume any parametric distribution; results did not differ (see Appendix Table F).

In all 5 studies included, chemotherapy was not allocated randomly. Therefore it is possible that dogs did not receive chemotherapy because of a worse prognosis, which would overestimate any beneficial effects of chemotherapy. To explore this, all analyses were repeated using the subset of patients that survived the first month (30 days). While, this analysis potentially decreases bias, precision was decreased because 22 % of the events occurred in the first month.

Previously, we implicitly assumed that the association of chemotherapy by logit(5-month mortality risk) interaction term with the outcome was linear. The appropriateness of this assumption was evaluated by comparing a model with a restricted cubic spline (with five knots) for the interaction term to a model without splines, using a likelihood ratio test. Additionally, a model was compared that categorized the logit(5-month mortality risk) in quintiles. No significant deviations from linearity were observed.

All tests were applied using a significance level of 0.05, estimates are presented as odds ratios (OR) with 95% confidence intervals (95%CI). Analyses were carried out using the R statistical package for windows version 3.0.2 (R Development Core Team, 2013), the lme4 package version 1.1-7 for random effect models (Bates et al., 2012), and the metaphor

279 package (Viechtbauer, 2010). R codes are available upon request with the first author. A
280 PRISMA checklist is included as appendix.
281
282 **Results**
283 Baseline characteristics of the 406 included dogs are presented in Table 2, after
284 surgical amputation 227 received additional chemotherapy and 143 dogs did not, of these 87
285 were dead after 5 months. Information on chemotherapy was missing for 30 subjects and 5-
286 month mortality for 36 subjects. In general, dogs not receiving chemotherapy were older,
287 weighed less, were more often female, neutered and had high SALP. The range of the logit(5-
288 month mortality risk) was -1.91 to 1.03. Baseline characteristics for the other comparisons
289 (carboplatin, cisplatin, doxorubicin and doxorubicin combination) are presented in Appendix
290 Tables B through E.
291
292 The crude main effect estimates of "any chemotherapy" versus no chemotherapy on 5-
293 month mortality was OR 0.43 (95%CI 0.27; 0.70). After adjustment for potential confounders
294 the OR was 0.48 (95%CI 0.30; 0.78), for details refer to Table 3. Results for the other
295 comparisons were similar; with the possible exception of the cisplatin effect, which was non-
296 significant (Table 3).
297
298 Testing for chemotherapy by logit(5-month mortality risk) interaction revealed that the
299 effects of any chemotherapy (compared to no chemotherapy) decreased with increasing
300 logit(5-month mortality risk, Table 4); interaction OR 3.41 (95%CI 1.08; 10.79) P-value =
301 0.04. Figure 1 depicts how the OR of chemotherapy changes with logit(5-month mortality)
302 and 5-month mortality (i.e., on the risk scale) and shows that dogs' with a 5-month mortality
303 risk of approximately 0.43 or less benefit from chemotherapy. For dogs' at a higher risk,

chemotherapy effectiveness is uncertain because an OR of 1 is included in the 95% confidence interval. Results for the other comparisons and study specific estimates were consistent with the overall ORs presented here,(Figures 2 and 3).

The results of the sensitivity analysis of excluding those subjects who died within the first month are depicted in Table 4. The main effect estimates were closer to the null than in the entire sample; OR of "any chemotherapy" 0.81 (95%CI 0.47; 1.39). Interaction effects were also closer to the 1: OR 2.44 (95%CI 0.66; 8.97). The magnitude and direction of these interaction effects were in agreement with those estimated in the entire sample. Similar and consistent results were found for the individual comparisons (data not shown).

Discussion

This study showed that dogs with osteosarcoma and a relatively low 5-month mortality risk (< 0.43) benefited, more from "any chemotherapy" compared to no chemotherapy than those with a higher risk. Results from individually comparing the chemotherapeutics, carboplatin, cisplatin, doxorubicin or doxorubicin combination therapy, to no chemotherapy were similar in direction and magnitude, indicating consistency, however interaction effects did not attain significance.

Previous clinical studies showed that the effect of chemotherapy might be modified by other factors. One of the clearest examples of this in dogs is the synergistic effect between immunotherapy and chemotherapy (MacEwen and Kurzman, 1996; Vail et al., 1995). To the best of our knowledge, our study is the first to explore whether chemotherapy effects vary according to baseline mortality risk (using multiple variables).

The current study has some limitations. First, only data from nonrandomized studies were available. Therefore, chemotherapy effect estimates could be biased due to unobserved and residual confounding. To explore this, a sensitivity analysis was performed, including dogs who survived the first month. Results in this subgroup showed estimates closer to an OR of 1.Interaction effects remained in the same direction and of the same magnitude as the interaction effects using the entire sample, implying consistency. Furthermore, using an independent validation cohort study (jointly submitted) these results were replicated. Despite these reassuring results, we acknowledge that in lieu of randomization our results can possibly still be confounded. This shortcoming could be remedied in the future by replicating our results using (historical) RCT data. A second issue is that the prediction model was derived including the subset of studies which was used to test for the presence of an interaction. However, a recent simulation study showed that such an internally developed model only deviated slightly from externally derived models (Burke et al., 2014). Additionally, we note that results were replicated using an independent cohort study (see manuscript 2 jointly submitted). To some including, non-significant predictors may seem erroneous. Perhaps surprisingly, numerous studies have shown that focussing on significant predictors results in an overfitted model which does not generalize well to other settings (Steyerberg, 2009; Steyerberg et al., 1999; Steyerberg et al., 2011; Steyerberg et al., 2010). To remedy this, it has been suggested to use prior knowledge to select relevant predictors, which we have implemented here. Regardless, a validation study (see manuscript 2 jointly submitted) showed that our choice of non-significant predictors was appropriate to predict early OS mortality an independent setting. More importantly, we note that the discriminatory ability of the prediction model was limited (c-statistic 0.63). However, calibration (i.e., how well the predicted risk matched the observed risk) was very reasonable. For the current purpose, predicting chemotherapy effects in individual dogs, good calibration is perhaps more relevant.

Another limitation is that a number of observations were missing. Instead of focussing on complete cases we tried to address this problem using multiple imputation. While most studies suffer from missing data and using imputation methods is likely to decrease bias it possible that results were biased. We did not perform a comparison with a complete case analysis (dropping missing values), because our analyses showed missing data to associated with observed variables invalidating a complete case analysis. Furthermore, too many subjects (n = 231; 58%) would be excluded to allow a proper analysis.(Carpenter and Bithell, 2000).(Sofroniou and Hutcheson, 2002)(Sofroniou and Hutcheson, 2002) RegardlessSimilarly, due to the small sample size available we did not adjust for multiple testing. Additionally, in the present analyses we only focused on beneficially effects (or its absence), ideally adverse events should also be evaluated. Unfortunately, this was not systematically measured in the original studies included in this IPD meta-analysis. Finally, some researchers prefer absolute effect measure such as risk differences over odds ratios. Presumably, the risk difference is preferred because, when the outcome incidence is low, the odds ratio can be large while the risk difference is small. For example, if the incidence in an unexposed group of subjects is 0.05 and 0.01 in the exposed the odds ratio is 5.21 while the risk difference is only 0.04. However, 5-month mortality in OS dogs is very common (24% in our sample) making this distinction less relevant. Nevertheless, the risk difference can be calculated from the equations presented. (Kent and Hayward, 2007) (Sun et al., 2011)

Conclusions

In conclusion, surgically treated dogs with osteosarcoma which have a relatively low risk of 5-month mortality might benefit most from additional chemotherapy.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organisations that could inappropriately influence or bias the content of the paper.

Acknowledgements

We wish to thank the Veterinary Society of Surgical Oncology (VSSO) society for allowing us to approach their members for collaboration. We gratefully acknowledge the following researchers for their willingness to collaborate:

Professor Dr. John Berg; Tufts Cummings School of Veterinary Medicine, North Grafton, MA, US.

Dr. Ilene Kurzman; School of Veterinary Medicine University of Wisconsin, Madison, Wisconsin, US.

Dr. Antony Moore; Veterinary Oncology Consultants, Australia.

Dr. Emanuela Morello; School of Veterinary Medicine, Turin, Italy.

Dr Joe Sottnik; Animal Cancer Center, Colorado State University, Fort Collins, CO, US.

Professor Dr. David Vail; School of Veterinary Medicine, University of Wisconsin, Madison, Wisconsin, US.

Author contributions

AFS, RHHG and MN contributed to the idea and design of the study. AFS and JK approached and coordinated with researchers to collect data. AFS performed the analyses and drafted the manuscript. MN, OHK, AWH, AB, RHHG, PA, NB, KK, KM and JK provided guidance during initial planning of the paper and during critical revision. AFS had full access to all of the data and takes responsibility for the integrity of the data presented.

Funding

This work was supported by Research Focus Areas funding of the Utrecht University, which is a collaboration between the faculties of medicine, science, and veterinary medicine. The study sponsor was not involved in the design, collection, analysis or writing of this manuscript.

Prior postings and presentations

This study and its results were neither previously published. An abstract containing this work was presented at the 2014 International Evidence-Based Veterinary Medicine Network Conference.

Appendix A: Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi: …

Figure captions

Figure 1. Estimated effect of any chemotherapy compared to no chemotherapy on 5-month mortality in surgically treated dogs with osteosarcoma.

[figure 1]

Figure shows the odds ratio (OR) of any chemotherapy treatment (solid line) with 95% confidence intervals (dotted lines) for dogs with different predicted risks of 5-month mortality. The horizontal solid line indicates a neutral OR of 1.00. At the bottom a histogram is given, corresponding to the patient frequencies of the x-axis measurement.

Figure 2. Estimated effects of four different chemotherapeutics compared to no chemotherapy on 5-month mortality in surgically treated dogs with osteosarcoma.

[figure 2]

Figure shows the odds ratio (OR) of carboplatin, cisplatin, doxorubicin or doxorubicin combination therapy compared to no chemotherapy (solid line) with 95% confidence intervals (dotted lines) for dogs with different predicted risks of 5-month mortality . The horizontal solid line indicates a neutral OR of 1.00. At the bottom of each graph a histogram is given, corresponding to the patient frequencies of the x-axis measurement.

Figure 3. Forest plot of the estimated effect of any chemotherapy compared to no chemotherapy on 5-month mortality in surgically treated dogs with osteosarcoma.

[figure 3]

The left panel shows the main effects, excluding an interaction, the right panel shows the interaction effects. Effects are depicted as odds ratio (OR) with 95% confidence interval, and pooled across studies using fixed effects (random intercept only) and random effects (random intercept and random slope) models. Interaction effect is per unit increase in the logit(5-month mortality risk). Heterogeneity measured as the tau-squared was 0.01 95%CI (-0.30; 0.33) for the main effect and 0.01 95%CI (-0.19; 0.19) for the interaction effect.

454
455 References
456
457 Altman, D.G., Bland, J.M., 2007. Missing data. BMJ 334, 424.

458 Amsellem, P.M., Selmic, L.E., Wypij, J.M., Bacon, N.J., Culp, W.T., Ehret, N.P., Powers, B.E., Stryhn, H.,
459 Farese, J.P., 2014. Appendicular osteosarcoma in small-breed dogs: 51 cases (1986-2011). Journal of
460 American Veterinary Medical Association.

461 Arbogast, P.G., Kaltenbach, L., Ding, H., Ray, W.A., 2008. Adjustment for multiple cardiovascular risk
462 factors using a summary risk score. Epidemiology 19, 30-37.

463 Bailey, D., Erb, H., Williams, L., Ruslander, D., Hauck, M., 2003. Carboplatin and doxorubicin
464 combination chemotherapy for the treatment of appendicular osteosarcoma in the dog. Journal of
465 Veterinary Internal Medicine 17, 199-205.

466 Bates, D., Maechler, M., Bolker, B., 2012. lme4: Linear mixed-effects models using S4 classes.

467 Bland, J.M., Altman, D.G., 1995. Multiple significance tests: the Bonferroni method. BMJ 310, 170.

468 Boerman, I., Selvarajah, G.T., Nielen, M., Kirpensteijn, J., 2012. Prognostic factors in canine
469 appendicular osteosarcoma - a meta-analysis. BMC Veterinary Research 8, 56.

470 Brodey, R.S., Abt, D.A., 1976. Results of surgical treatment in 65 dogs with osteosarcoma. Journal of
471 the American Veterinary Medical Association 168, 1032-1035.

472 Burke, J.F., Hayward, R.A., Nelson, J.P., Kent, D.M., 2014. Using internally developed risk models to
473 assess heterogeneity in treatment effects in clinical trials. Circ. Cardiovasc. Qual. Outcomes. 7, 163-
474 169.

475 Carpenter, J., Bithell, J., 2000. Bootstrap confidence intervals: when, which, what? A practical guide
476 for medical statisticians. Stat. Med 19, 1141-1164.

477 Chatfield, C., 1995. Model Uncertainty, Data Mining and Statistical Inference. Journal of the Royal
478 Statistical Society. Series A (Statistics in Society) 158, 419-466.

479 Chun, R., Garrett, L.D., Henry, C., Wall, M., Smith, A., Azene, N.M., 2005. Toxicity and efficacy of
480 cisplatin and doxorubicin combination chemotherapy for the treatment of canine osteosarcoma.
481 Journal of the American Animal Hospital Association 41, 382-387.

482 Chun, R., Kurzman, I.D., Couto, C.G., Klausner, J., Henry, C., MacEwen, E.G., 2000. Cisplatin and
483 doxorubicin combination chemotherapy for the treatment of canine osteosarcoma: a pilot study.
484 Journal of Veterinary Internal Medicine 14, 495-498.

485 Cooley, D.M., Waters, D.J., 1997. Skeletal neoplasms of small dogs: a retrospective study and
486 literature review. Journal of the American Animal Hospital Association 33, 11-23.

487 Easterbrook, P.J., Berlin, J.A., Gopalan, R., Matthews, D.R., 1991. Publication bias in clinical research.
488 Lancet 337, 867-872.

489 Farooq, V., van, K.D., Steyerberg, E.W., Meliga, E., Vergouwe, Y., Chieffo, A., Kappetein, A.P.,
490 Colombo, A., Holmes, D.R., Jr., Mack, M., Feldman, T., Morice, M.C., Stahle, E., Onuma, Y., Morel,
491 M.A., Garcia-Garcia, H.M., van Es, G.A., Dawkins, K.D., Mohr, F.W., Serruys, P.W., 2013. Anatomical

492 and clinical characteristics to guide decision making between coronary artery bypass surgery and
493 percutaneous coronary intervention for individual patients: development and validation of SYNTAX
494 score II. Lancet 381, 639-650.

495 Harrell, F.E., Jr., Dupont, C., 2013. Hmisc: Harrell Miscellaneous. R package.

496 Hernan, M.A., Hernandez-Diaz, S., Werler, M.M., Mitchell, A.A., 2002. Causal knowledge as a
497 prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J
498 Epidemiol. 155, 176-184.

499 Kent, D.M., Hayward, R.A., 2007. Limitations of applying summary results of clinical trials to
500 individual patients: the need for risk stratification. Journal of the American Medical Association 298,
501 1209-1212.

502 Kirpensteijn, J., Kik, M., Rutteman, G.R., Teske, E., 2002. Prognostic significance of a new histologic
503 grading system for canine osteosarcoma. Veterinary Pathology 39, 240-246.

504 Kow, K., Thamm, D.H., Terry, J., Grunerud, K., Bailey, S.M., Withrow, S.J., Lana, S.E., 2008. Impact of
505 telomerase status on canine osteosarcoma patients. Journal of Veterinary Internal Medicine 22,
506 1366-1372.

507 Little, R.J.A., Rubin, D.B., 2002. Statistical Analysis with Missing Data. Wiley-Blackwell.

508 MacEwen, E.G., Kurzman, I.D., 1996. Canine osteosarcoma: amputation and chemoimmunotherapy.
509 Veterinary Clinics of North America: Small Animal Practice 26, 123-133.

510 Marshall, A., Altman, D.G., Holder, R.L., Royston, P., 2009. Combining estimates of interest in
511 prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Medical
512 Research Methodology 9, 57.

513 McNeill, C.J., Overley, B., Shofer, F.S., Kent, M.S., Clifford, C.A., Samluk, M., Haney, S., Van Winkle,
514 T.J., Sorenmo, K.U., 2007. Characterization of the biological behaviour of appendicular osteosarcoma
515 in Rottweilers and a comparison with other breeds: a review of 258 dogs. Veterinary and
516 Comparative Oncology 5, 90-98.

517 Moons, K.G., Kengne, A.P., Woodward, M., Royston, P., Vergouwe, Y., Altman, D.G., Grobbee, D.E.,
518 2012. Risk prediction models: I. Development, internal validation, and assessing the incremental
519 value of a new (bio)marker. Heart 98, 683-690.

520 Norrdin, R.W., Powers, B.E., Torgersen, J.L., Smith, R.E., Withrow, S.J., 1989. Characterization of
521 osteosarcoma cells from two sibling large-breed dogs. American Journal of Veterinary Research 50,
522 1971-1975.

523 R Development Core Team, 2013. R: A language and environment for statistical computing. R
524 Foundation for Statistical Computing, Vienna, Austria.

525 Ru, G., Terracini, B., Glickman, L.T., 1998. Host related risk factors for canine osteosarcoma. The
526 Veterinary Journal 156, 31-39.

527 Rubin, D.B., 1976. Inference and missing data. Biometrika 63, 581-592.

528 Schmidt, A.F., Nielen, M., Klungel, O.H., Hoes, A.W., de Boer, A., Groenwold, R.H., Kirpensteijn, J.,
529 2013. Prognostic factors of early metastasis and mortality in dogs with appendicular osteosarcoma

after receiving surgery: An individual patient data meta-analysis. Preventive Veterinary Medicine 112, 414-422.

Sofroniou, N., Hutcheson, G.D., 2002. Confidence Intervals for the Predictions of Logistic Regression in the Presence and Absence of a Variance- Covariance Matrix. Understanding Statistics 1, 3-18.

Sottnik, J.L., Rao, S., Lafferty, M.H., Thamm, D.H., Morley, P.S., Withrow, S.J., Dow, S.W., 2010. Association of blood monocyte and lymphocyte count and disease-free interval in dogs with osteosarcoma. Journal of Veterinary Internal Medicine 24, 1439-1444.

Spodnick, G.J., Berg, J., Rand, W.M., Schelling, S.H., Couto, G., Harvey, H.J., Henderson, R.A., MacEwen, G., Mauldin, N., McCaw, D.L., ., 1992. Prognosis for dogs with appendicular osteosarcoma treated by amputation alone: 162 cases (1978-1988). Journal of the American Veterinary Medical Association 200, 995-999.

Steyerberg, E.W., 2009. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer, New York.

Steyerberg, E.W., Eijkemans, M.J., Habbema, J.D., 1999. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J. Clin. Epidemiol. 52, 935-942.

Steyerberg, E.W., Schemper, M., Harrell, F.E., 2011. Logistic regression modeling and the number of events per variable: selection bias dominates. J. Clin. Epidemiol. 64, 1464-1465.

Steyerberg, E.W., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M.J., Kattan, M.W., 2010. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 21, 128-138.

Straw, R.C., Withrow, S.J., 1996. Limb-sparing surgery versus amputation for dogs with bone tumors. Veterinary Clinics of North America: Small Animal Practice 26, 135-143.

Straw, R.C., Withrow, S.J., Richter, S.L., Powers, B.E., Klein, M.K., Postorino, N.C., LaRue, S.M., Ogilvie, G.K., Vail, D.M., Morrison, W.B., ., 1991. Amputation and cisplatin for treatment of canine osteosarcoma. Journal of Veterinary Internal Medicine 5, 205-210.

Sun, X., Briel, M., Busse, J.W., You, J.J., Akl, E.A., Mejza, F., Bala, M.M., Bassler, D., Mertz, D., az-Granados, N., Vandvik, P.O., Malaga, G., Srinathan, S.K., Dahm, P., Johnston, B.C., onso-Coello, P., Hassouneh, B., Truong, J., Dattani, N.D., Walter, S.D., Heels-Ansdell, D., Bhatnagar, N., Altman, D.G., Guyatt, G.H., 2011. The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials: systematic review. BMJ 342, d1569.

Vail, D.M., Kurzman, I.D., Glawe, P.C., O'Brien, M.G., Chun, R., Garrett, L.D., Obradovich, J.E., Fred, R.M., III, Khanna, C., Colbern, G.T., Working, P.K., 2002. STEALTH liposome-encapsulated cisplatin (SPI-77) versus carboplatin as adjuvant therapy for spontaneously arising osteosarcoma (OSA) in the dog: a randomized multicenter clinical trial. Cancer Chemotherapy and Pharmacology 50, 131-136.

Vail, D.M., MacEwen, E.G., Kurzman, I.D., Dubielzig, R.R., Helfand, S.C., Kisseberth, W.C., London, C.A., Obradovich, J.E., Madewell, B.R., Rodriguez, C.O., Jr., ., 1995. Liposome-encapsulated muramyl tripeptide phosphatidylethanolamine adjuvant immunotherapy for splenic hemangiosarcoma in the dog: a randomized multi-institutional clinical trial. Clinical Cancer Research 1, 1165-1170.

van Klaveren D., Vergouwe, Y., Farooq, V., Serruys, P.W., Steyerberg, E.W., 2015. Estimates of absolute treatment benefit for individual patients required careful modeling of statistical interactions. J. Clin. Epidemiol. In press.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software 36, 1-48.