Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at *Vet. Sci.* **2018**, *5*, 77; doi:10.3390/vetsci5030077

Article

Estimating the size of dog populations in Tanzania to inform rabies control

Maganga Sambo[1, 2]*, Katie Hampson[2], Joel Changalucha[1], Sarah Cleaveland[2], Tiziana Lembo[2], Kennedy Lushasi[1,3], Eberhard Mbunda[4†], Zacharia Mtema[1], Lwitiko Sikana[1, 3], Paul Johnson[2]

1 Ifakara Health Institute, P.O.Box 53, Ifakara, Tanzania

2 Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, UK

3 Nelson Mandela African Institution of Science and Technology, P.O. Box 447, Arusha, Tanzania.

4 Department of Epidemiology, Ministry of Livestock and Fisheries, P.O. Box 2870, Dodoma, Tanzania.

* Correspondence: e-mail : smaganga@ihi.or.tz; Tel.: +44-740-810-2140

† Deceased

Abstract: Estimates of dog population sizes are a prerequisite for delivering effective canine rabies control. However, dog population sizes are generally unknown in most rabies-endemic areas. Several approaches have been used to estimate dog populations but without rigorous evaluation. We compare post-vaccination transects, household surveys and school-based surveys to determine which most precisely estimates dog population sizes. These methods were implemented across 28 districts in southeast Tanzania, in conjunction with mass dog vaccinations, covering a range of settings, livelihoods and religious backgrounds. Transects were the most precise method, revealing highly variable patterns of dog ownership, with human: dog ratios ranging from 12.4:1 to 181.3:1 across districts. Both household and school-based surveys generated imprecise and sometimes inaccurate estimates, possible due to low sample size. Transect data were subsequently used to develop a predictive model for estimating dog populations in districts lacking transect data. We predicted a dog population of 2,316,000 (95% CI 1,573,000-3,122,000) in Tanzania and an average human: dog ratio of 20.7:1. Our modelling approach has the potential be applied to predicting dog population size in other districts where mass dog vaccination is carried out, given census and livelihood data. We recommend transects as a rapid and effective method to refine dog population estimates across large geographic areas and guide dog vaccination programs.

Keywords: dog-mediated rabies; transects; dog ownership; mass dog vaccination

1. Introduction

Dog-mediated rabies is a serious zoonosis responsible for at least 59,000 human deaths every year, primarily in low-income countries in Asia and Africa where rabies is endemic [1]. In these areas, over 95% of human rabies deaths result from bites by domestic dogs. Empirical data and mathematical modelling have shown that annual dog vaccination campaigns that achieve a coverage of 70% are sufficient to eliminate rabies [2,3]. The global call for the elimination of dog-mediated human rabies by 2030 [4], has prompted many countries to invest in dog vaccinations. For example, Tanzania has developed a National Rabies Control and Elimination Strategy, aiming to control dog rabies and eliminate human rabies in the country by 2030 [5]. However, in low-income countries, dog population sizes are usually unknown or hard to estimate, making it difficult to implement and evaluate dog vaccination campaigns [6].

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

In most low-income countries where rabies is endemic, owned dogs are not registered by local authorities, dogs are free to roam and dog censuses are not conducted. Insufficient knowledge of dog population sizes for planning of vaccination campaigns was reported as one of the limiting factors for ineffective rabies control in Africa [7]. This lack of knowledge prevents countries from forecasting vaccine procurement needs and hinders assessment of the effectiveness of mass dog vaccination campaigns. It is therefore important to develop practical methods for estimating dog sizes.

Approaches for estimating dog population sizes include extrapolation from human: dog ratios derived from household surveys and, mark-resight methods (transects) using dog collars or paint sprays [7–12]. The accuracy of these methods has been questioned as they generate very different population estimates from the same geographical areas and can be imprecise unless large numbers of households are sampled [6,14]. Furthermore, household surveys are restricted to owned dogs, while transects capture only free-roaming (observable) dogs. Conducting a dog census, whereby each household in a community is visited, is the gold standard method to estimate the dog population, but requires considerable investment of time and resources.

Dog ownership patterns are not uniform across all communities. The distribution of dogs in different settings is linked to religious, cultural, geographical and socio-economic factors [9,10,15]. For example, there tend to be fewer dogs in predominantly Muslim communities than in Christian communities [15]. Tools to accurately estimate dog populations that take into account local cultural norms could therefore support the scaling up of mass dog vaccination programmes.

Our overall aim is to provide practical and effective approaches to estimate dog populations in different settings. Our first objective was to compare methods to determine which provides the most precise dog population estimates and explain why estimates differ according to method. From this comparison, we identified that post-vaccination transects, which involves counting both vaccinated and unvaccinated dogs, provides more reliable dog population estimates than either household or school-based surveys. Our second objective was to identify factors that predict dog ownership in different settings in Tanzania, using our estimates of dog population size from transects. The aim of identifying these factors was to enable prediction of dog population sizes and densities in other parts of the country not yet subject to vaccination campaigns, which was our third objective. We assessed the performance of these factors from known populations in our study and finally used these factors together with nationally available human census data to predict dog population sizes throughout Tanzania. Our findings should be valuable for both Tanzania and other countries as they develop, implement and monitor their national rabies control programmes.

2. Methods

2.1. Study sites

The study was conducted in 28 districts across Tanzania: 24 from southeast Tanzania and four from Pemba Island (Figure 1). In 2012, the national human population and housing census recorded 8,819,886 people (average population growth rate of 2.16% per year) living in 2,159,721 households in these districts, with 784,424 of these households in rural areas [16]. These districts span a wide range of settings, livelihoods and religious backgrounds [5] and cover rural, urban, coastal and inland areas. The districts cover a total area of 165,043 square kilometres (17% of the landmass in Tanzania) and are representative of much of Tanzania in terms of population sizes, area, and livelihoods (Figure A1).

2.2. Data collection

Since 2011 dog rabies vaccinations have been conducted in villages across the study districts using a central point approach [17], whereby owners bring their dogs to a centrally located vaccination point within their village. We used the number of vaccinated dogs in each district as the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

minimum possible dog population size from which to compare the performance of our estimates. Data from three different approaches were used to estimate dog populations: post-vaccination transects, household surveys and school-based surveys. These methods are described in full elsewhere [13] and outlined briefly below.

Post-vaccination Transects: Since 2013, following these vaccination campaigns post-vaccination transects have been conducted [14]. Transects were walked (or occasionally cycled) on the same day as campaigns from 4 to 6 p.m. when dogs were active and visible, counting all marked (vaccinated) and unmarked (unvaccinated) dogs. In rural areas, transects were conducted in two randomly selected sub-villages per village (villages ranged in size from 2 to 10 sub-villages, with a median of 4), aiming to representatively sample each village. In the first sub-village, enumerators started at the sub-village centre and headed to the outskirts, while in the other, enumerators walked from the edge toward the centre. Each transect was conducted by one enumerator for 1 hour per sub-village, taking 2 hours to complete the village. In urban areas, enumerators covered the jurisdiction of a street (a geographical area defined by the National Census, which covers a neighbourhood with several roads). Enumerators selected the direction at the start of transects, at the border of sub-villages/streets and at road junctions by spinning a pen. Numbers of vaccinated dogs per district were compiled from dog vaccination registers and used in conjunction with the transect data to estimate coverage achieved and numbers of dogs in each district. In Serengeti district, data from mass dog vaccinations and transects were used to validate the performance of transects whereas data from a complete census of dogs (covering all 35,867 households in Serengeti district) were used to calculate puppies to adult ratio.

Household Surveys: Household surveys were completed in all study districts between July and August 2011 with the aim of obtaining an initial assessment of coverage from the first phase of vaccination campaigns conducted between February and April, 2011. Six villages were randomly selected from all villages in each district. In each selected village, a landmark was identified (preferably a school, otherwise a dispensary, church, or mosque). From this starting point, interviewers randomly chose a direction for selecting households for interview by spinning a pen. Every third household was sampled, and interviews conducted until 30 households were completed in each village. For households that owned dogs, the questionnaire captured details of dogs owned (adults and puppies < 3 months) and their vaccination status on the basis of owner recall.

School-Based Surveys: School-based surveys were conducted from June 2014 to February 2015 in the two months following dog vaccination campaigns in each district. The surveys were used to ask questions on the number of adults and children (< 18 years of age) in households as well as the number of dogs and puppies (< 3 months of age), and their vaccination status. Six primary schools (one per village; most villages in Tanzania have a primary school) were randomly selected from each district but logistic and financial limitations meant that surveys were not conducted in four districts and were conducted in fewer than six schools per district as initially planned. We used total population purposive sampling with a target to interview 100 pupils per school [18]. Standard VII pupils (ages from 13-15 years) were selected to complete the questionnaire. If more than one pupil from a household was recruited, the oldest pupil was selected. If the school had fewer than 100 standard VII pupils, pupils were recruited from lower classes (Standard IV–VI, ages from 12-14 years). Schools were selected from the same villages where household surveys were conducted.

2.3. Characteristics of study districts:

Human demographic data for the study were extracted from the 2012 national population and housing census, including district-level population sizes, annual population growth rates, average household sizes (persons/household), numbers of livestock keeping households, and the percentage of the population living in rural areas [15]. Human populations were projected from the 2012 census to 2014 using annual growth rates for each district. The census also included district-level information on employment status among the working population (defined as those aged 10 years

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

or over), which showed that the most frequent occupations were farming and livestock keeping. We therefore also extracted the number of workers in each district employed as peasants and livestock keepers respectively. To examine whether geographical setting was associated with dog ownership, districts were also categorized as: 1) *Inland;* 14 districts from the mainland that did not border the ocean 2) *Coastal;* 10 districts that bordered the Indian Ocean, including both Tanzania mainland coastal districts and the four districts from Pemba island; and 3) *Island;* the four districts from Pemba island. We extracted the area of each district from [19] using the *unionSpatialPolygon* function from the *maptools* package in R [20] excluding water bodies and protected areas. These variables are summarised in Table 1 for study and non-study districts.

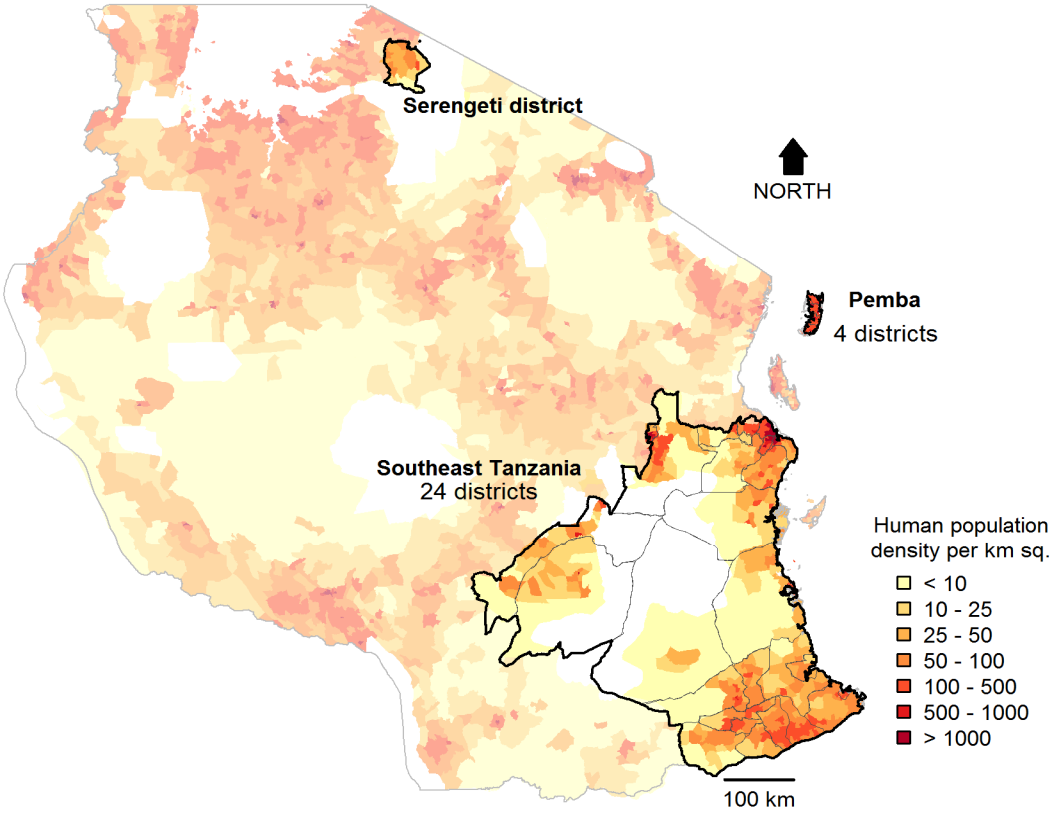

Figure 1. Districts in Tanzania where post-vaccination transects, household and school-based surveys were conducted. Data were collected in Southeast Tanzania and Pemba (28 districts) where mass dog vaccination campaigns have been conducted since 2011. In Serengeti district, mass dog vaccinations campaigns and transects were also conducted in the last quarter of 2015 to assess the performance of the model at predicting dog numbers outside the study area. White colored areas represent the protected areas.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Table 1. Characteristics of study and non-study districts in Tanzania. Continuous variables are summarised by the mean (range) and categorical variables by number (% proportion). Variables were either extracted from the national census [15] or from district shapefiles from the Tanzania National Bureau of Statistics (NBS) website [18].

Variable	Study districts (n = 28)	Non-study district (n = 140)
Human population size	307,676 (70,209, 1,775,049)	257,188 (39,242, 807,619)
Annual % human population growth rate	2.3 (-1.0, 7.0)	2.6 (-3, 7)
Number of households	75,452 (16,892, 441,240)	50, 636 (9,027, 134,608)
Average household size (n persons per household)	4.2 (3.5, 5.5)	5.1 (3.8, 7.8)
Area (km^2)¥	4,375 (15, 28,000)	4,375 (18.6, 28,244)
Setting:		
Inland	14 (50%)	128 (91%)
Coastal‡ [including Island]	14 (50%) [4 (14%)]	12 (9%) [6 (4.7%)]
Number of livestock-keeping households	18,317 (4,771, 35,829)	24,168 (2,258, 71,335)
Proportion (%) of persons employed* as:		
Peasants	60 (3, 88)	64 (4, 93)
Livestock keepers	1 (0, 6)	1 (0, 65)

¥Excluding protected areas and water bodies.

‡Coastal districts were defined as districts that border the Indian ocean.

*Defined as the main occupation on which persons aged 10 years and above spend most of their working time.

2.4. Statistical analysis

Estimation of dog population sizes: For transects, which involve counting both vaccinated and unvaccinated dogs, estimates of dog vaccination coverage and their 95% confidence intervals (CIs) were obtained from binomial generalized linear mixed models (GLMMs) with normally distributed (on the logit scale) random intercepts allowing coverage to vary among villages. Dog population estimates for each district were calculated by dividing the numbers of vaccinated dogs by the coverage estimates. The total number of dogs across the 28 districts was calculated as the sum of these estimates, with a 95% confidence interval estimated as the 2.5% and 97.5% quantiles from 10,000 bootstrap samples. Puppies under the age of three months are not often vaccinated during campaigns [21,22], and are less likely to be counted than adult dogs during transects [14,22,23]. However, puppies comprise approximately one quarter of the dog population [11]. Estimates of population sizes from transects therefore require further adjustment. We adjusted by calculating the missing ratio of pup:adult. The adjustment factor (pup:adult ratio) was calculated from a dog census conducted in Serengeti district in Northern Tanzania [14].

For households and school-based surveys, we estimated dog population sizes by multiplying the mean number of dogs per household recorded during these surveys by the total number of households within each district according to the 2012 national census. The district-level estimates were summed to give the overall dog population size in the study districts. For each district, the mean number of dogs per household and profile likelihood 95% confidence limits were estimated from a negative binomial GLM, where the response was the number of dogs in each household and the only fixed parameter was the intercept. Confidence limits were not calculated for districts where too few (≤ 3) surveyed households owned dogs to provide reliable estimates of uncertainty. We explored fitting a village random effect, but omitted it because of failure to converge for districts where few households owned dogs, and zero variance estimates otherwise.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Prediction of dog population sizes: We used a linear regression model to achieve our second objective of identifying factors associated with dog population sizes (Table 1) and our third objective of estimating the number of dogs in each district. Here we motivate our approach to modelling dog numbers.

First, rather than modelling dog numbers directly, we modelled the dog:human ratio, then multiplied estimated dog:human ratio by human population size to estimate dog numbers. A simple model for predicting the ratio of the number of dogs (D_i) to the number of humans (P_i) in district i is $\frac{D_i}{P_i} = K\varepsilon_i$, where every district is predicted to have the same dog: human ratio, K, from which individual districts are allowed to deviate by means of district-specific residuals, ε_i, where $\varepsilon_i > 0$ with geometric mean 1. We can add flexibility to this model by introducing continuous variables that are potential predictors of the dog: human ratio, for example the proportion of the population employed as livestock keepers, L_i/P_i, such that $\frac{D_i}{P_i} = K\left(\frac{L_i}{P_i}\right)^{\beta_1}\varepsilon_i$, where β_1 is a coefficient governing the influence of L_i/P_i on D_i/P_i. The model can be further extended to allow differences in D_i/P_i between categories of district, such as between districts with and without a coastline: $\frac{D_i}{P_i} = K\left(\frac{L_i}{P_i}\right)^{\beta_1}\exp(\beta_2 C_i)\varepsilon_i$, where C_i is a binary indicator variable with the value 1 for coastal districts and 0 otherwise, so that $\exp(\beta_2)$ represents the ratio (coastal dog: human ratio):(mainland dog:human ratio). Taking the natural log shows that this model can be viewed as a linear regression model: $\log\left(\frac{D_i}{P_i}\right) = \log(K) + \beta_1\log\left(\frac{L_i}{P_i}\right) + \beta_2 C_i + \log(\varepsilon_i)$, where $\log(\varepsilon_i) \sim \text{Normal}(0, \sigma^2)$. Further continuous and binary variables (listed in Table A2) were added in the same way (continuous variables were logged and categorical variables fitted as binary indicator variables) to give the full model prior to filtering using the variance inflation factor.

We used this ordinary least squares regression model to assess whether any of the variables that we collated were useful for predicting numbers of dogs. We partitioned our data into two sets of districts: a training (or model-building) set, which we used to develop the model, and a prediction set, for which we aimed to predict dog numbers from the model. The training set comprised the data from the 28 study districts where we collected data, whereas the prediction set included all districts in Tanzania (including the 28 study districts). The response variable was the estimated number of the estimated dogs in each of the study districts, log-transformed to fit the modelling assumptions of linearity and homoscedasticity. Before fitting the model, we conducted data exploration to detect collinearity among the variables using pairwise scatter plots and the variance inflation factor (VIF) as described by Zuur et al. [24]. The variable that with the highest VIF was removed and VIFs recalculated, repeating this process, stopping when all VIF values were below 5. The remaining variables were taken forward to model selection.

The aim of model selection was to identify the best-fitting model, that is, the model that most accurately predicted district dog population sizes. The best model was selected by fitting all possible subsets of the six candidate variables as main effects and choosing the model with the lowest corrected AIC (AIC_C) [25]. The predictive power of the best model was assessed by calculating R^2_{FPE}, where FPE stands for final prediction error [26]. The interpretation of R^2_{FPE} is similar to classical R^2 and adjusted R^2, except that unlike these two statistics, R^2_{FPE} does not overestimate predictive power [26]. We also calculated adjusted R^2 for comparison. The contribution of each selected variable to the predictive power of the best model was gauged by calculating partial R^2_{FPE}, which estimates the proportional reduction in prediction error when a given variable is included in the final model. Partial R^2_{FPE} is calculated as $1 - (1 - R^2_{FPE}) / (1 - R^2_{FPE}{}^*)$, where R^2_{FPE} is calculated from the best model, and $R^2_{FPE}{}^*$ is calculated from a reduced version of the best model where the variable under investigation has been dropped.

Two methods were used to assess the validity of the final model. First, we assessed the propensity for the model selection procedure to select spurious variables ("false discoveries"). We permuted the response variable 1000 times to simulate 1000 data sets in which none of the variables

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

are associated with the response. We applied the model selection procedure to each data set and calculated the mean number of variables selected. The ratio of mean number of false discoveries to the actual number of variables selected is an estimate of the false discovery rate (FDR). Second, we used bootstrapping to assess the stability of the selected variables to sampling error. We performed the model selection procedure on 1000 bootstrapped data sets (data sets of equivalent size but sampled with replacement from the original data set), recording the proportion of data sets from which each variable was selected. The final model was then used to predict the dog population size, density and 95% prediction interval for each district in Tanzania.

To assess year-to-year variation in our estimates, we compared dog population sizes estimated from data that were collected in 2014-2015 against those estimated from 2015-2016 data.

When making predictions from the models of coverage and the number of dogs per 1000 humans, it was necessary to back-transform from the linear predictor scale using a nonlinear inverse "link" function (inverse logit for coverage and exponential for dog numbers). Because these models have normally distributed errors, predictions would be biased, potentially severely, if calculated simply by back-transforming without adjusting for the error variance, as a consequence of Jensen's inequality [27]. Log-scale predictions from the model of district dog population size were adjusted for Jensen's inequality by back-transforming using $\exp(\mu + 0.5\sigma^2)$, and logit-scale predictions of vaccination coverage were back-transformed using $\text{logit}^{-1}(\mu + 0.5\sigma^2 \tanh(\mu(1 + 2 \exp(-0.5\sigma^2))/6))$, where μ is the link-scale prediction and σ^2 is the total error variance [28].

All statistical analyses were conducted using R version 3.4.1 [29]. GLMMs were fitted using the lme4 [30] and *glmmTMB* packages [31]. Models selection was performed using the *dredge* function from the *MuMIn* package [32].

2.5. Ethics

The study protocol was approved by the Medical Research Coordinating Committee of the National Institute for Medical Research of Tanzania (NIMR/HQ/R.8a/Vol.IX/2109), the Institutional Review Board of the Ifakara Health Institute and the Tanzania Commission for Science and Technology (COSTECH). Before administering any questionnaires, participants were informed about the background and purpose of the study. Participants were informed that they were free to participate in the study and withdraw from the study at any time during the interview, and that their answers would be kept confidential.

3. Results

During the 2014-2015 mass dog vaccination campaigns, 86,361 dogs were vaccinated in the 28 study districts, and 86,142 dogs were vaccinated in the 2015-2016 campaign. The following data collection activities were completed: (i) post-vaccination transects in ~2,100 villages in 2014-2015 and in ~2,600 villages in 2015-2016; (ii) household surveys in 4,488 households in 2011, from 160 randomly selected villages; and (iii) school-based surveys of 8,254 primary school pupils (each representing a unique household) within 115 randomly selected schools following the 2014-2015 campaign. During the 2014-2015 transects, about 18,436 dogs were counted, of which 63% were observed with collars indicating that they were vaccinated. From the school-based surveys 2,198 owned dogs were reported, corresponding to a mean of 0.7 dogs per household and from the household surveys 731 dogs were recorded, corresponding to a mean of 0.6 dogs per household (Table A1).

Estimation of dog population sizes: The overall dog population estimated in these 28 districts varied according to the survey method used (Figure 2). From transects, using the best fitting model we estimated a total dog population in the study districts of 174,000 (95% CI 138,000–257,000 reported to three significant digits) and an overall human:dog ratio of 50.6:1 in the study districts. In contrast, using household and school-based surveys we estimated the dog population in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

study districts to be 412,000 (CI 225,000 -789,000) and 403,000 (CI 251,000 - 752,000) respectively. District-level estimates of dog numbers from household and school-based surveys tended to have wide 95% confidence intervals, with a mean ratio of upper to lower confidence limit of 3.5 for household and 3.0 for school-based surveys. Transect estimates were considerably more precise, with a mean upper:lower confidence limit ratio of 1.2 (Figure A2). Household and school-based surveys were also sometimes highly inaccurate, giving estimates that were lower than the number of vaccinated dogs (Figure 2). Transect estimates by contrast cannot possibly be lower than the number of vaccinated dogs. We therefore considered transect estimates to be more reliable and used them to fit the model for predicting dog population sizes.

Prediction of dog population sizes and densities: Using our transect estimates we investigated the influence of district-level variables on dog population sizes. The pairwise plots between the log-scale continuous variables investigated showed that the number of people living in rural areas was highly correlated with the human population (Pearson's $r = 0.96$). We therefore dropped the number of people living in rural areas from the model. A further two variables, the number of households and the number of livestock-owning households, were dropped in order to reduce all VIFs to below 5 (Table A2). These variables were also highly correlated with the human population size.

All 64 possible models were fitted from the combination of the six retained variables and the models were ranked by δAIC_C (Table 2). The top three models were almost the equivalent in predictive power (R^2_{FPE} =58%) and were very close in δAIC_C which ranged from 0-1.27. Our best fitting model retained three variables: the proportions of livestock keepers and of peasants, and the geographic setting (inland versus coastal and island). We also found that there were 103% (95% CI: 38%, 113%) more dogs per person in inland districts than island and coastal districts.

Two predictor variables: the proportion of livestock keepers and the setting (inland versus coastal|island) were consistently retained in the best-ranked models (Table 2). Dropping either the setting variable or the proportion of persons employed as peasants also reduced the variance explained by 16% (R^2_{FPE} fell from 58% to 42%). Excluding the proportion of persons employed as livestock keepers in the final model reduced R^2_{FPE} to 30%, showing the substantial predictive power of this variable. The model with only two variables had a much lower R^2_{FPE} (~6%) compared to the model with three variables R^2_{FPE} (58%).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

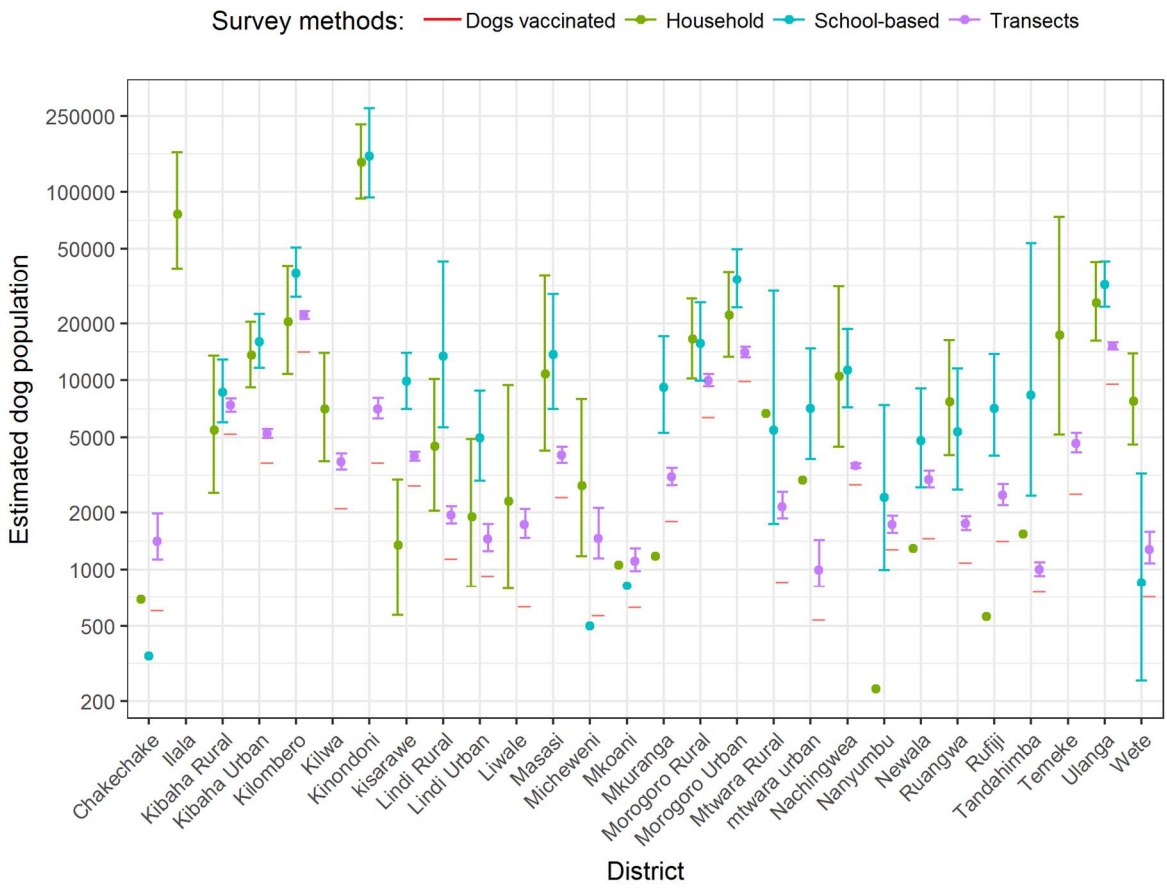

Figure 2. Comparison of the estimated number of dogs in the 28 districts from the three methods: household surveys, school-based surveys and transect surveys. Estimates are shown on a log scale (the actual scale is shown in Figure A3). Error bars show 95% confidence intervals (CI) around the mean estimate. Household surveys were conducted in 2011 while school-based surveys and transects were conducted in 2014 and 2015 following dog vaccination campaigns. The number of vaccinated dogs from the 2014-2015 campaign is also plotted as a horizontal line for each district. CIs are not plotted for household or school-based surveys from districts where fewer than three households were recorded to own dogs among the randomly selected households in each district.

We used permutations and bootstrapping to assess the reliability of the selected best-fitting model. Using 1,000 permutations we estimated an FDR of 28%. Bootstrapping the model selection procedure showed that two of the three variables were highly robust: proportion of livestock keepers and coastal setting were selected in 97% and 91% of bootstrapped models, respectively (Table 2). Two other variables were selected with around 50% frequency, proportion employed as peasant (49%) and human population size (51%). Based on δAIC_C, R^2_{FPE} and the reliability analysis, we concluded that two of the three variables selected are highly robust, that at least one other variable is required to maximise predictive power. Because the top three models are almost equivalent in terms of predictive power (R^2_{FPE}) and δAIC_C, we selected the one with only three variables selected, which also happened to be the best-fitting model. Therefore this combination of three variables (the proportions of livestock keepers and of peasants, and coastal setting) was chosen for our final model and used to predict the dog population across Tanzania.

We predicted considerable variability in dog population sizes across all 168 districts in Tanzania (Figure 3). Dog population estimates ranged from just 630 dogs in Kusini district on the island of Zanzibar, to 45,000 dogs in the inland district of Nzega district. Overall, the final model predicted a dog population of 2,316,000 (95% CI 1,573,000 - 3,122,000) in Tanzania in 2014-2015. In 2014, a human population of 47,831,000 was projected from the Tanzania population census [16].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Taken together with our predicted dog population size (2,316,000 dogs) we obtain an overall human:dog ratio of 20.7:1 in Tanzania. Generally, the highest dog ownership was predicted from inland districts especially those dominated by rural livestock keepers. Human:dog ratios for each study district are presented in Table A1.

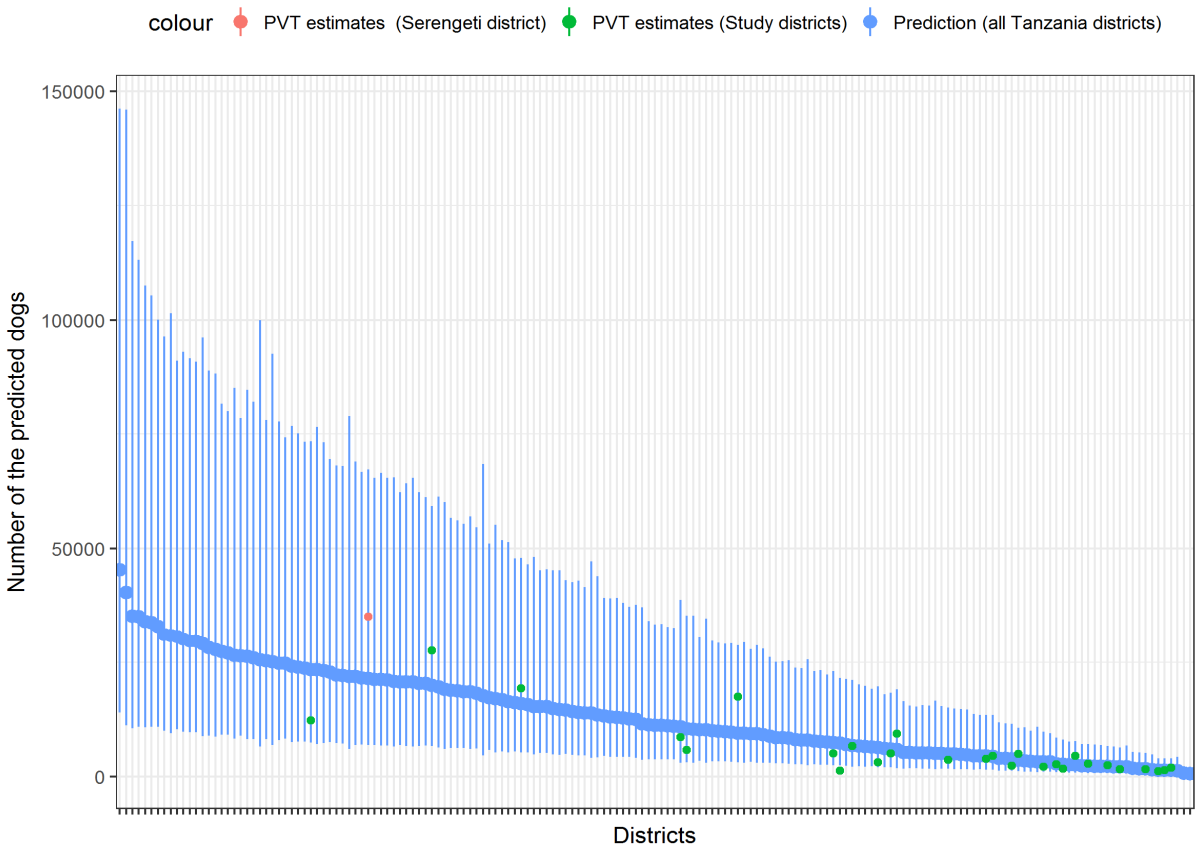

Figure 3. The predicted number of dogs in districts across Tanzania. The predictions and prediction intervals are shown in blue and compared to the dog population estimated directly from transects for a non-study district (Serengeti) and for the study districts. Most of the green points (27 out of 28) were within the prediction intervals suggesting that the fit of the model is reasonably good, with no outlier districts for whom the model is making very wrong predictions. Districts are ordered according to those with the largest predicted dog population (left-hand side) to the smallest population (right-hand side). District names are shown in Figure A4. Serengeti transects were conducted in July to December, 2015.

Dog density from the predicted dog population sizes for each of the 168 districts also varied greatly, ranging from 0.14 to 113 dogs per square kilometer (Figure 4). Liwale district (an inland study district which has an area of ~15,000 km^2) had the lowest dog density of 0.14 km^2 while Bukoba urban (an inland district with an area of 30 km^2) had the highest dog density of 113 dogs per km^2. Based on the Tanzanian's total land area and predicted dog population, we determined the mean density of dogs to be 8.81 per square kilometer (interquartile range: 2.24 - 9.23). Predicted densities were highest in northern and north-eastern Tanzania and lowest in the southern and central-west parts of the country (Figure 4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at *Vet. Sci.* **2018**, *5*, 77; doi:10.3390/vetsci5030077

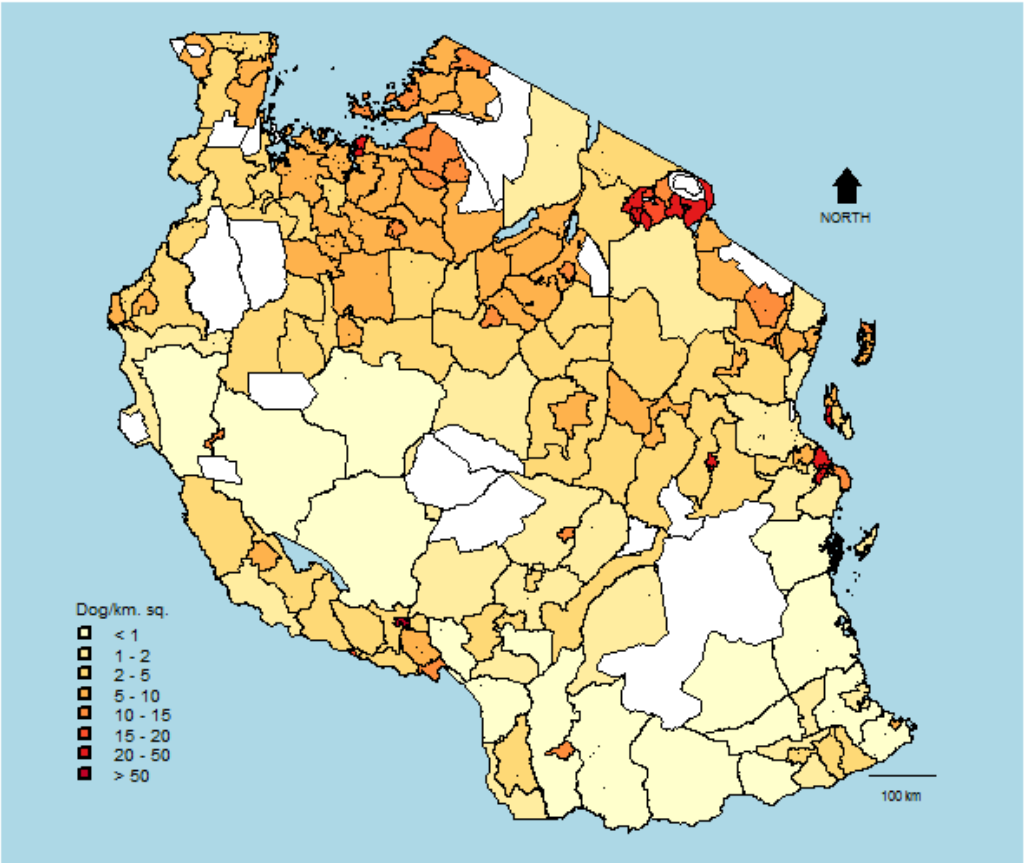

Figure 4. Estimated dog densities at the district-level across Tanzania. White areas represent water bodies, forest reserves or wildlife protected areas..

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Table 2. Regression coefficients and model fit statistics for models predicting dog populations in Tanzania. Details are shown for all models with $\delta AIC_C < 4$, ranked in order of decreasing fit as gauged by δAIC_C. We report the predictor variable regression coefficient estimates, model fit statistics including the partial R^2 (R^2), adjusted $R^{2,}$, the degrees of freedom (df), the log-likelihood (LL), a second-order akaike information criterion (AICc), delta AIC (δAIC_C), akaike weights, and R^2_{FPE}. We also report the importance of each predictor variable (variable robustness).

Regression coefficient estimates							**Model fit statistics**							
Intercept	***Area (km²)***	***Inland vs Coastal & Island***	***Mainland vs Island***	***Number of people***	***Employed as livestock keepers***	***Employed as peasants***	R^2	***Adjusted R^2***	***df***	***LL***	***AICc***	***δAIC_C***	***Weight***	***R^2_{FPE} (%)***
5.60		-0.708			0.358	0.448	0.658	0.724	5	-18.65	50.03	0.00	26.0%	57.6%
9.51	0.124	-0.758		-0.447	0.303		0.688	0.763	6	-17.49	50.98	0.94	16.2%	58.0%
8.07		-0.760		-0.219	0.335	0.300	0.682	0.760	6	-17.65	51.30	1.27	13.8%	57.5%
9.99		-0.907		-0.412	0.281		0.640	0.757	5	-19.39	51.52	1.48	12.4%	55.3%
11.03		-0.767	-0.422	-0.489	0.302		0.661	0.703	6	-18.54	53.08	3.05	5.7%	54.7%
Variable robustness	33%	91%	13%	51%	97%	49%								

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

4. Discussion

Knowledge of the size of dog populations is critical for the planning and implementation of effective canine rabies control strategies. Before implementation of vaccination campaigns, this information is useful for determining the personnel required, and for procurement of vaccines and other supplies. After implementation, this information is required to evaluate the intervention in terms of the vaccination coverage achieved. Our study uses data from well-studied dog populations in Tanzania where vaccination campaigns have been conducted to predict the size of dog populations in new areas, where vaccination campaigns will hopefully be scaled up in future.

Although several methods have been used to calculate the size of dog populations [33], these methods have not been comprehensively compared to assess which generates the most precise dog population estimates. Here we compared methods used to estimate dog population sizes to inform rabies control in Tanzania. We found that post-vaccination transects generated more precise and reliable estimates than either household or school-based surveys.

The imprecision of the estimates from household and school-based surveys is due to their relatively low sampling effort compared with post-vaccination transects [14]. Although sample size at the household level was not low (30 households per village and ~100 families per school), only six out of an average of 95 villages per district were sampled, and the precision of estimates from hierarchical sampling designs is expected to be dominated by sample size at the highest level [34,35]. The overall proportions of the population surveyed was also small (4,488/441,000 or ~1% from household surveys versus 8,254/441,000 or ~2% from school-based surveys), despite the large numbers of surveys conducted, compared with post-vaccination transects, where ~2500 villages and >70% of the population were sampled. The practical consequence of imprecision in the estimates of from household and school-based surveys for the implementation of mass dog vaccination is that upper bound of the range of plausible dog population sizes is on average 3 times the lower bound, compared with 1.18 for transects.

In addition to being imprecise, estimates from household and school-based surveys were often inconsistent (Figure 2), either with each other (for example, Kisarawe and Wete districts), or with the number of dogs vaccinated (for example, Mkuranga and Chakechake districts).

The dog population may therefore be either over or under-estimated by projecting from the mean number of dogs per household, given limited sampling and considerable house-to-house variation in dog ownership. For example, Kinondoni district in Dar es Salaam has over 400,000 households with 0.32 dogs per household estimated from the household survey, which would suggest a dog population of 141,197 which is highly over-estimated (Table A1).

The number of vaccinated dogs (numerators) together with the size of dog populations derived from household and school-based surveys (denominators) were used to calculate the vaccination coverage, and compared against coverage estimated directly from these surveys. We found large discrepancies depending upon the method used. For example, in 2014-2015 around 86,000 dogs were vaccinated but school-based surveys estimated a dog population of 399,000, which would have led to just 22% coverage (Table A1). In contrast, direct estimates of dog vaccination coverage from school-based and household surveys were both around 56% (Table A1). In some districts, the discrepancy was clearly implausible, for example in Chake Chake, Pemba, where more than double the number of dogs estimated from school-based surveys were vaccinated (Table A1). These discrepancies have practical impacts on rabies control and demonstrate the need for careful post-vaccination evaluations, without reliance on dog population survey estimates only.

Our multivariable regression analysis identified the proportion of livestock keepers and geographical setting as robust variables for predicting dog population sizes, consistent with previous studies on factors influencing dog ownership in different parts of the world [33,36–38]. In most African and Asian countries, dogs are reported to play a role in protecting livestock [9,12,13]. Other factors reported to influence dog ownership in Africa include socio-economic status,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

livelihood (which could include livestock keeping), culture, and religious beliefs linked to different settings. The reason for fewer dogs in Tanzanian coastal and island areas could be linked to predominately Muslim communities. Muslims are reported to own fewer dogs [9,15]. Our study demonstrates the need to consider such factors in planning vaccination campaigns, and in understanding dog rabies incidence, control and prevention more generally.

We found that despite the robustness of these two variables, by themselves they accounted for only about two thirds of the predictive power of the best-fitting model. A third variable was also needed to improve predictive power. Our model validation showed that the proportion of peasants or the human population were almost equivalent in the final model (R_{2FPE} values of 58% and 55% respectively). If applying this model to new settings, the decision as to which variable should be used will depend on the availability of data.

Our study was consistent with previous findings in Tanzania which reported more dogs in mainland compared to island and coastal areas [6,8]. However, our overall human:dog ratio of 20.7:1 was higher than previous studies in Africa, which ranged from 3:1 to 15:1 [6,9,15,36,39,40]. This suggests that human:dog ratios extrapolated from household or school-based surveys could be unreliable when extrapolated to district or national level. Initially, the study area was estimated to have about 400,000 dogs [5,41], based on reported human:dog ratios [9], which was much higher than the number of dogs subsequently estimated with post-vaccination transects (174,000, 95% CI 138,000–257,000). The lower number of dogs in our study suggests that dog vaccine requirements in Africa might be less than previous estimates. However, our data were largely collected from south-east Tanzania where there are fewer pastoralists who tend to own more dogs [42], and in coastal or island districts (~50% of study districts) which tend to have fewer dogs. This suggests that additional data (dog vaccination and transect surveys) from other populations (inside and outside of Tanzania) would be valuable to further refine and validate this predictive approach.

Several household surveys have been conducted to estimate human to dog ratios in Tanzania. In Iringa urban, the HDR was estimated to be 14 (CI 13 -15) [8], whereas our HDR estimated from transect survey for Iringa urban were 34 (11-106). In Serengeti and Ngorongoro districts HDR were estimated to be 7 and 8 respectively [17], versus our HDR estimates of 13 (4 - 40) and 7 (1 - 28) respectively. In Kilombero and Ulanga districts HDR ratios were estimated to be 12 and 29 respectively [15], in contrast to our estimate of 21 (7, 63) and 18 (6, 53) respectively. Variation was also reported by Geographical setting, with HDR ratios estimated to be 7.6:1 in rural-inland areas, 10.8:1 in rural-coastal areas, 27.1:1 in urban-coastal areas and 14.4:1 in urban-inland areas [9]. These estimates were all derived from household surveys and are likely to be affected by limited sampling. Our study suggests that transects in all villages within a district better captures variability, but also can only be conducted in association with mass dog vaccinations, hence the need for predictive methods for working in areas where dog vaccinations have yet to be conducted.

Our overall estimates of dog density were higher those reported from elsewhere in Africa [36,43–46]. This was probably because we excluded water bodies and protected areas when calculating densities. Our dog density map highlights districts with high dog densities that should be prioritized in the scaling up of dog vaccinations (only two districts, Moshi urban and Zanzibar mjini had densities exceeding 120 dogs/km2, Table A1), and districts where dog densities might be too low to support rabies transmission without importation from other districts [45,47,48].

Our study had several limitations. The main limitation was that we didn't assess the our predictions using external data (data from outside of the study areas). In addition, our study areas did not cover many inland livestock keeping districts. Our surveys were also completed at different times, with household surveys conducted in 2011, transects immediately following dog vaccination campaigns in 2014/15 and school-based surveys within 2 months of these dog vaccinations. These differences may have affected our estimates. Transects do have several limitations. They may result in the recounting of dogs [49], but we tried to avoid this in the design of transect paths. Larger villages also require more time to complete and those with more sub-villages were less well

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

sampled. Only observable dogs are counted from transects, which may also result in systematic biases, such as poor observation of pups [6], but we also tried to adjust for this using pup:adult ratio. Finally, we could not externally validate our predictive model due to a lack of reliable data on dog numbers outside the study area, with the exception of Serengeti district. Notwithstanding these limitations, we found that transects were fast and relatively low cost to complete at scale, sampling populations more representatively than other approaches that were limited in spatial scope. We suggest that transects should be completed immediately after vaccination campaigns, aiming to cover the center and the periphery of villages, as vaccination coverage has been reported to decrease with the distance to the vaccination point [17,22,50,51], and that estimates should be adjusted to account for not observing pups.

5. Conclusions

Our study underlines the importance of knowledge of dog population sizes and distribution in different settings in rabies endemic countries, with methods of estimating the number of dogs needed to plan, monitor and evaluate the performance of rabies control efforts. We demonstrated that post-vaccination transects together with dog vaccination data can be used to rapidly generate and refine dog population estimates in areas with ongoing vaccination campaigns and to generate tentative predictions of dog populations in districts without dog vaccination interventions. We also show that data derived from smaller scale sampling of the dog population could lead to substantial under or overestimation of the population in areas with considerable village-to-village variation leading to poor rabies control. We conclude that post-vaccination transects may be a useful tool for the global initiatives for rabies elimination, taking advantage of data on vaccinated dogs that are routinely collected through implementation of mass dog vaccinations.

Author Contributions: Conceptualization, Maganga Sambo, Sarah Cleaveland, Tiziana Lembo and Katie Hampson; Data curation, Maganga Sambo; Formal analysis, Maganga Sambo and Paul Johnson; Funding acquisition, Katie Hampson, Sarah Cleaveland and Tiziana Lembo; Investigation, Maganga Sambo, Joel Changalucha, Kennedy Lushasi, Eberhard Mbunda, Zacharia Mtema and Lwtiko Sikana ; Methodology, Maganga Sambo; Project administration, Eberhard Mbunda; Supervision, Katie Hampson and Paul Johnson; Validation, Maganga Sambo; Writing – original draft, Maganga Sambo; Writing – review & editing, Maganga Sambo, Katie Hampson, Joel Changalucha, Sarah Cleaveland, Tiziana Lembo, Kennedy Lushasi, Zacharia Mtema, Lwtiko Sikana and Paul Johnson.

Funding: We acknowledge the Bill & Melinda Gates Foundation for funding rabies mass dog vaccination campaigns in Tanzania as part of an initiative coordinated by the World Health Organization, the Wellcome Trust (grants 095787/Z/11/Z and 207569/Z/17/Z) for supporting KH and MS. MS received an international student fee waiver from the University of Glasgow. The UBS Optimus Foundation for funding surveillance through a grant to the University of Glasgow and the CARE project coordinated by the Global Alliance for Rabies Control. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments: We thank the District Veterinary Officers, Livestock Field Officers and local community members who provided invaluable assistance with data collection. We are grateful to the Ministry of Health, Community Development, Gender, Eldery and Children and the Ministry of Livestock and Fisheries, Ifakara Health Institute, the National Institute for Medical Research and the Tanzanian Commission for Science and Technology for permission to carry out the study. Appreciation is extended to Francesco Ventura and Rachel Steenson (University of Glasgow) for technical support.

Conflicts of Interest: The authors declare that no conflict of interest exist. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Hampson, K.; Coudeville, L.; Lembo, T.; Sambo, M.; Kieffer, A.; Attlan, M. et al. Estimating the global burden of endemic canine rabies. *PLoS Negl Trop Dis*. **2015**;9: e0003709. doi:10.1371/journal.pntd.0003709

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

2. Coleman, P.G.; & Dye, C. Immunization coverage required to prevent outbreaks of dog rabies. *Vaccine.* **1996**;14: 185–186. doi:10.1016/0264-410X(95)00197-9

3. Hampson, K.; Dushoff, J.; Cleaveland, S.; Haydon, D.T.; Kaare, M.; Packer, C. et al. Transmission dynamics and prospects for the elimination of canine Rabies. *PLoS Biol*. **2009**;7: 0462–0471. doi:10.1371/journal.pbio.1000053

4. Abela-Ridder, B.; Knopf, L.; Martin, S.; Taylor, L .; Torres, G.; De Balogh, K. 2016: the beginning of the end of rabies? *Lancet Glob Heal* **2018**;4: e780–e781. doi:10.1016/S2214-109X(16)30245-5

5. Mpolya, E.A.; Lembo, T.; Lushasi, K.; Mancy, R.; Mbunda, E.M .; Makungu, S. et al. Toward Elimination of Dog-Mediated Human Rabies: Experiences from Implementing a Large-scale Demonstration Project in Southern Tanzania. *Front Vet Sci*. **2017**; doi:10.3389/fvets.2017.00021

6. Downes, M.J .; Dean, R.S.; Stavisky, J.H.; Adams, V.J.; Grindlay, D.J.; Brennan, M.L. et al. Methods used to estimate the size of the owned cat and dog population: a systematic review. *BMC Vet Res*. **2013**;9: 121. doi:10.1186/1746-6148-9-121

7. Lembo, T.; Hampson, K.; Kaare, M.; Ernest, E.; Knobel, D.; Kazwala, R. et al. The feasibility of canine rabies elimination in Africa: Dispelling doubts with data. PLoS Negl Trop Dis. 2010; doi:10.1371/journal.pntd.0000626

8. Bögel, K.; & Meslin, F.X.; Economics of human and canine rabies elimination: guidelines for programme orientation. Bull World Health Organ,1990; 68 (3), p.281.

9. Gsell, S.; Knobel, D.; Kazwala, R.; Vounatsou, P.; Zinsstag, J. Domestic dog demographic structure and dynamics relevant to rabies control planning in urban areas in Africa: the case of Iringa, Tanzania. *BMC Vet Res*. **2012**;8: 236–245. doi:10.1186/1746-6148-8-236.

10. Knobel, D.; Laurenson M.K .; Kazwala, R.; Boden, L.A.; Cleaveland, S. A cross-sectional study of factors associated with dog ownership in Tanzania. *BMC Vet Res*. **2008**;4: 5. doi:10.1186/1746-6148-4-5

11. Kongkaew, W.; Coleman, P.; Pfeiffer, D.U.; Antarasena, C.; Thiptara, A. Vaccination coverage and epidemiological parameters of the owned-dog population in Thungsong District, Thailand. *Prev Vet Med*. **2004**;65: 105–115. doi:10.1016/j.prevetmed.2004.05.009

12. Tenzin, T.; McKenzie, J.S.; Vanderstichel, R .; Rai, B.D.; Rinzin, K.; Tshering, Y. et al. Comparison of mark-resight methods to estimate abundance and rabies vaccination coverage of free-roaming dogs in two urban areas of south Bhutan. *Prev Vet Med.* **2015**;118: 436–448. doi:10.1016/j.prevetmed.2015.01.008

13. Tenzin, T.; Ahmed, R .; Debnath, N.C.; Ahmed, G.; Yamage M. Free-roaming dog population estimation and status of the dog population management and rabies control program in Dhaka City, Bangladesh. *PLoS Negl Trop Dis*. **2015**;9: e0003784. doi:10.1371/journal.pntd.0003784

14. Sambo, M.; Johnson, P.C.; Hotopp, K.; Changalucha, J.; Cleaveland, S.; Kazwala, R . et al. Comparing Methods of Assessing Dog Rabies Vaccination Coverage in Rural and Urban Communities in Tanzania. *Front. Vet. Sci*. **2017**;4: p.33. doi:10.3389/fvets.2017.00033

15. Bardosh, K.; Sambo, M.; Sikana, L.; Hampson, K.; Welburn, S.C. Eliminating Rabies in Tanzania? Local Understandings and Responses to Mass Dog Vaccination in Kilombero and Ulanga Districts. *PLoS Negl Trop Dis*. **2014**;8: e2935. doi:10.1371/journal.pntd.0002935

16. NBS, *Tanzanian Housing and Population Census*. Planning Office, Dar Es Salaam, Tanzania; **2012**.

17. Kaare, M.; Lembo, T.; Hampson, K.; Ernest, E.; Estes, A.; Mentzel, C . et al. Rabies control in rural Africa: evaluating strategies for effective domestic dog vaccination. *Vaccine.* **2009**;27: 152–60. doi:10.1016/j.vaccine.2008.09.054

18. Teddlie, C.; & Yu, F. Mixed methods sampling: A Typology With Examples. *J Mix Methods Res*. **2007**; 1(1), 77-100 doi:10.1177/2345678906292430

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

19. Tanzania National Bureau of Statisctics, Available online: http://www.nbs.go.tz/nbstz/index.php/english/statistics-by-subject/population-and-housing-census/258-2012-phc-shapefiles-level-one-and-two (accessed on 17 June 2018).

20. Bivand, R.; & Lewin-Koh, N. maptools: Tools for reading and handling spatial objects. *R package version* **2013**. 0.8-29.

21. Durr, S., Mindekem, R., Kaninga, Y., Moto, D. D., Meltzer, M. I., Vounatsou, P., & Zinsstag, J. Effectiveness of dog rabies vaccination programmes: comparison of owner-charged and free vaccination campaigns. *Epidemiol Infect.* **2009**;137: 1558–1567. doi:10.1017/S0950268809002386

22. Minyoo, A.B.; Steinmetz, M .; Czupryna, A .; Bigambo, M.; Mzimbiri, I.; Powell, G. et al. Incentives increase participation in mass dog rabies vaccination clinics and methods of coverage estimation are assessed to be accurate. *PLoS Negl Trop Dis*. **2015**;9: e0004221.

23. Arief, R.A .; Hampson, K.; Jatikusumah, A .; Widyastuti, M.D.W.; Basri, C.; Putra, A.G. et al. Determinants of vaccination coverage and consequences for rabies control in Bali, Indonesia. *Front Vet Sci*. **2017**;3: 123.

24. Zuur, A.F.; Ieno, E.N.; Elphick, C.S. A protocol for data exploration to avoid common statistical problems. *Methods Ecol Evol*. **2010**;1: 3–14. doi:10.1111/j.2041-210X.2009.00001.x

25. Wagenmakers, E.J.; Farrell, S. AIC model selection using Akaike weights. *Psychonomic Bulletin and Review*. **2004**. 11(1), 192-196. doi:10.3758/BF03206482

26. Rousson, V.; & Goşoniu, N.F. An R-square coefficient based on final prediction error. *Stat Methodol.* **2007**;4: 331–340. doi:https://doi.org/10.1016/j.stamet.2006.11.004

27. Nakagawa, S.; Johnson, P.C.D.; Schielzeth, H. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J R Soc Interface*. **2017**; 14(134), p.20170213. doi:10.1098/rsif.2017.0213

28. McCulloch, C.E.; & Searle, S.R. *Generalized, Linear, and Mixed Models*. John Wiley & Sons, Ltd, 2001. doi:10.1198/tech.2003.s13

29. R Core Team. R: A Language and Environment for Statistical Computing. **2017.** R Foundation for Statistical Computing , Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

30. Bates, D.; Maechler, M.; Bolker, B.; Walker, S. lme4: Linear mixed-effects models using Eigen and S4. *R package*. 2014:(1)7:1-23 31.

31. Magnusson, A.; Skaug, H.; Nielsen, A.; Berg, C.; Kristensen, K.; Maechler, M.; van Bentham, K.; Bolker, B.; & Brooks, M., 2017. glmmTMB: generalized linear mixed models using Template Model Builder. *R package version.* **2016**; *(0) 1. 0.*

32. Bartoń, K. MuMIn: multi-model inference. *R package version.* **2014**; (1)10. http//CRAN R-project org/package= MuMIn

33. Davlin, S.L & Vonville, H.M. Canine rabies vaccination and domestic dog population characteristics in the developing world: a systematic review. *Vaccine.* **2012**;30: 3492–502. doi:10.1016/j.vaccine.2012.03.069

34. Maas, C.J.; & Hox, J. Sufficient Sample Sizes for Multilevel Modeling. *J Res Methods Behav Soc Sci*. **2005**; 1(3), 86. doi:10.1027/1614-1881.1.3.86

35. Snijders, T.A.B.; Power and Sample Size in Multilevel Linear Models. *Enc Stat Beh Sci*. **2005.** doi:10.1002/0470013192.bsa492

36. Kitala, P.; McDermott, J.; Kyule, M.; Gathuma, J .; Perry, B.; Wandeler, A. Dog ecology and demography information to support the planning of rabies control in Machakos District, Kenya. *Acta Trop*. **2001**;78: 217–230. doi:10.1016/S0001-706X(01)00082-1

37. Acosta-Jamett, G.; Cleaveland, S .; Cunningham, A.A.; Bronsvoort, B.M. Demography of domestic dogs in rural and urban areas of the Coquimbo region of Chile and implications for disease transmission. *Prev Vet Med.* **2010**;94: 272–281. doi:10.1016/J.PREVETMED.2010.01.002

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

38. Mbilo, C.; Léchenne, M.; Hattendorf, J.; Madjadinan, S.; Anyiam, F. Rabies awareness and dog ownership among rural northern and southern Chadian communities—Analysis of a community-based, cross-sectional household survey. *Acta Trop.* **2017**;175: 100–111. doi:10.1016/J.ACTATROPICA.2016.06.003

39. Awoyomi, O.J.; Adeyemi, I.G.; Awoyomi, F.S. Socioeconomic Factors Associated With Non-Vaccination Of Dogs Against Rabies In Ibadan, Nigeria. *Niger Vet J.* **2008**; 28(3), 59-63. doi:10.4314/nvj.v28i3.3569

40. El-Yuguda, A.; Baba, A.A.; Baba, S.A. Dog population structure and cases of rabies among dog bite victims in urban and rural areas of Borno state, Nigeria. *Trop Vet*. 2007; (25), 34-40

41. Hatch, B.; Anderson, A.; Sambo, M.; Maziku, M .; Mchau, G .; Mbunda, E. et al. Towards Canine Rabies Elimination in South-Eastern Tanzania: Assessment of Health Economic Data. *Transbound Emerg Dis*. **2017**; 64(3), pp.951-958 . doi:10.1111/tbed.12463

42. Malele, I.; Nyingilili, H .; Msangi, A. Factors defining the distribution limit of tsetse infestation and the implication for livestock sector in Tanzania. *African J Agric Res*. **2011**;(6) 2341–2347

43. Knobel, D.L.; Cleaveland, S.; Coleman, P.Gv Fevre, E.M.; Meltzer, M.I .; Miranda, M.E . et al. Re-evaluating the burden of rabies in Africa and Asia. *Bull World Health Organ.* **2005**;(83) 360–368. doi:/S0042-96862005000500012

44. Brooks, R. Survey of the dog population of Zimbabwe and its level of rabies vaccination. *Vet Rec.* **1990**; 127(24), pp.592-596. doi:10.1136/vr.127.24.592

45. Dye, C.; & Cleaveland, S. Maintenance of a Microparasite Infecting Several Host Species Rabies in the Serengeti. *Parasitology.* **1995**; 111(S1), S33-S47. doi:10.1017/S0031182000075806

46. Kitala, P.M.; McDermott, J.J.; Coleman, P.G.; Dye, C. Comparison of vaccination strategies for the control of dog rabies in Machakos District, Kenya. *Epidemiol Infect.* **2002**;129, 215–222. doi:10.1017/S0950268802006957

47. Cleaveland, S.; Kaare, M.; Tiringa, P.; Mlengeya, T.; Barrat J. A dog rabies vaccination campaign in rural Africa: impact on the incidence of dog rabies and human dog-bite injuries. *Vaccine.* **2003**;21, 1965–1973. doi:10.1016/S0264-410X(02)00778-8

48. Lembo, T.; Hampson, K .; Haydon, D.T .; Craft, M.; Dobson, A.; Dushoff, J. et al. Exploring reservoir dynamics: A case study of rabies in the Serengeti ecosystem. *J Appl Ecol.* **2008**; 45 (4), pp.1246-1257. doi:10.1111/j.1365-2664.2008.01468.x

49. Rinzin, K .; Tenzin,T.; Robertson, I. Size and demography pattern of the domestic dog population in Bhutan: Implications for dog population management and disease control. *Prev Vet Med.* **2016**; 126, pp.39-47 doi:10.1016/j.prevetmed.2016.01.030

50. Matter, H.C.; Wandeler, A.I.; Neuenschwander, B.E.; Harischandra, L.P.A.; Meslin, F.X.; Study of the dog population and the rabies control activities in the Mirigama area of Sri Lanka. *Acta Trop.* **2000**;75: 95–108. doi:http://dx.doi.org/10.1016/S0001-706X(99)00085-6

51. Mazeri, S.; Gibson, A.D.; Meunier, N.; Barend, M.; Handel, I.G.; Mellanby, R.J. et al. Barriers of attendance to dog rabies static point vaccination clinics in Blantyre, Malawi. *PLoS Negl Trop Dis*. **2018**;12: e0006159

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

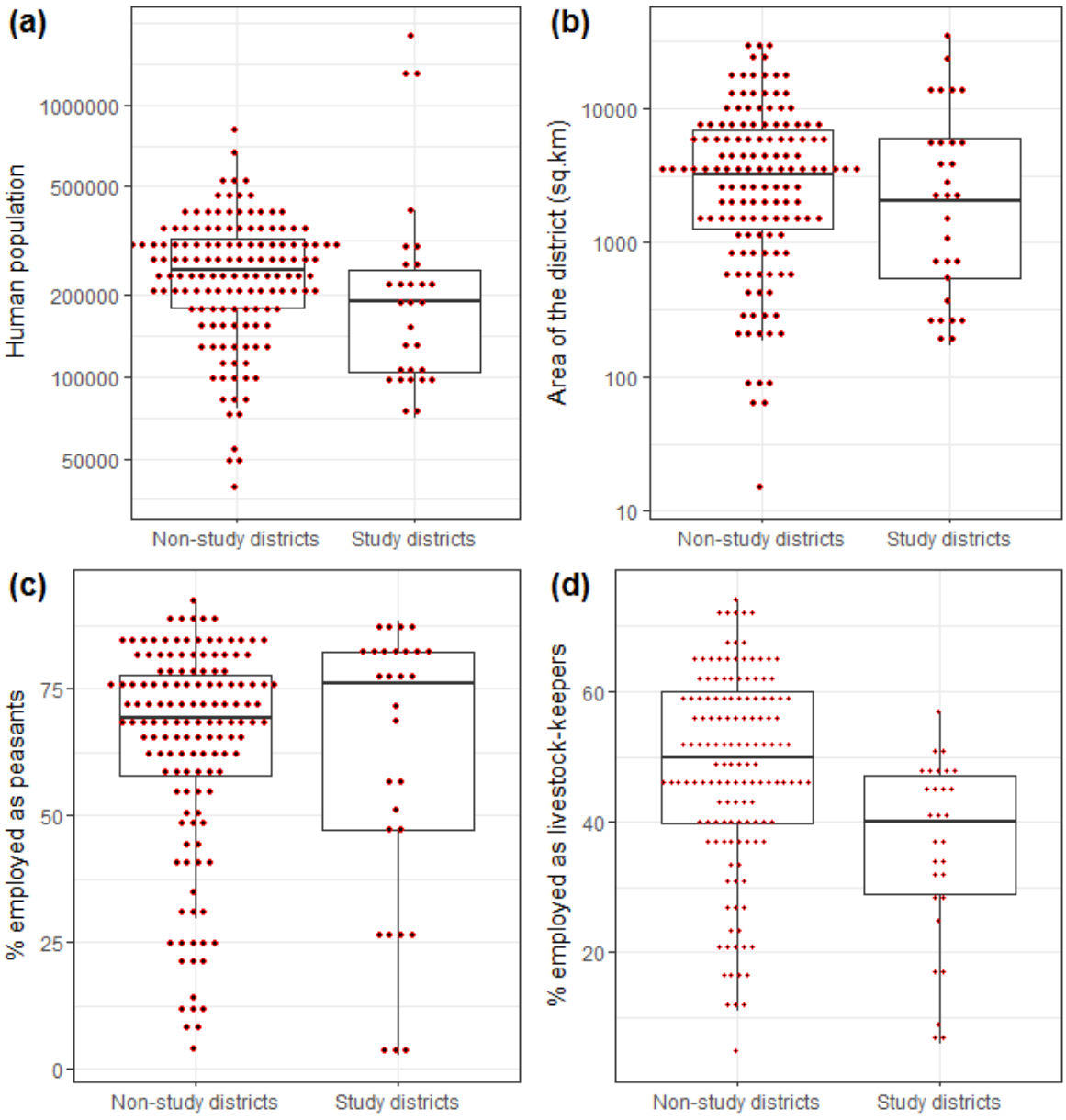

Figure A1. The distribution of the non-study districts and study districts in Tanzania. Red points represent data from each district. **(A)** human population from the National census in 2012 **(B)** geographical area of the district in square kilometres **(C)** percentage of population (aged >10 years and above) employed as peasants and **(D)** percentage of population (aged >10 years and above) employed as livestock keepers.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

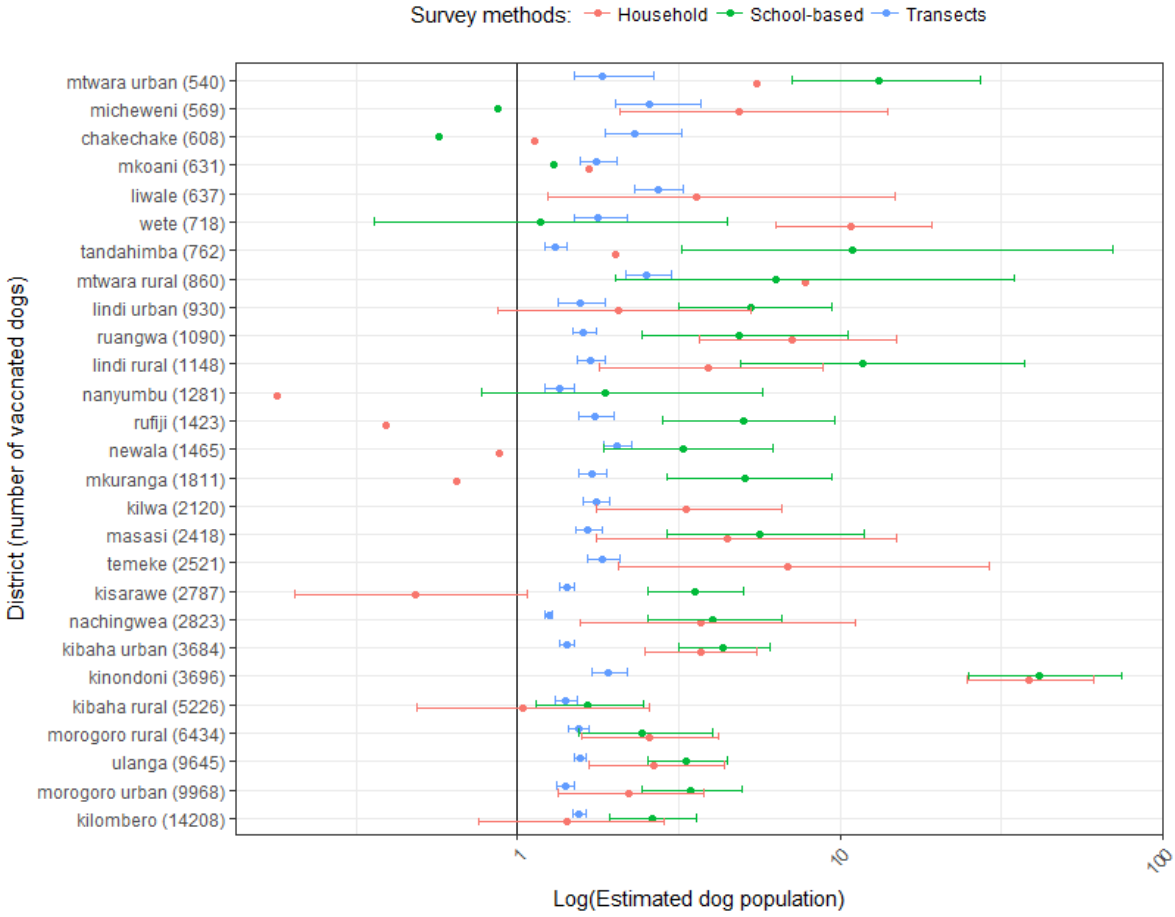

Figure A2. Dog population size estimates for the 27 districts. The ratio is on actual scale while the estimates are on log-scale. The points that are on the left side of the second vertical line means that the estimated number of dogs were fewer than number of vaccinated dogs. Points are the mean estimate and error bars show the 95% confidence intervals (CI) around the mean.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

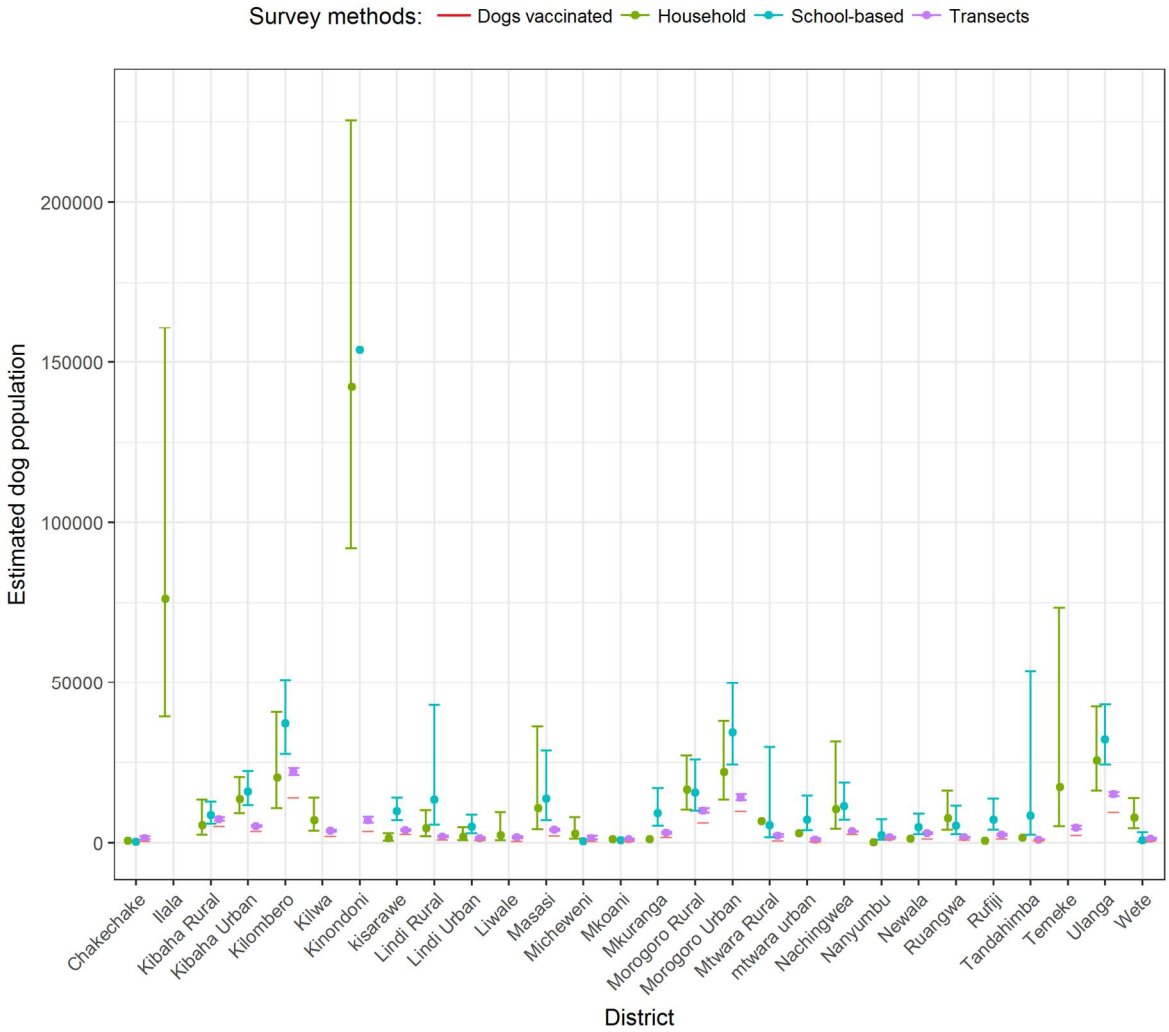

Figure A3. Dog population size estimates for the 28 districts (on actual scale). Data for the household survey were collected in 2011 (different dog vaccination campaigns) while school-based survey and transects were conducted in 2014 and 2015 following the dog vaccination campaigns. Points are the mean estimate and error bars show the 95% confidence intervals (CI) around the mean. We did not report CIs for districts which counted less than three dog-owning households or where surveys were not administered.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at *Vet. Sci.* **2018**, *5*, 77; doi:10.3390/vetsci5030077

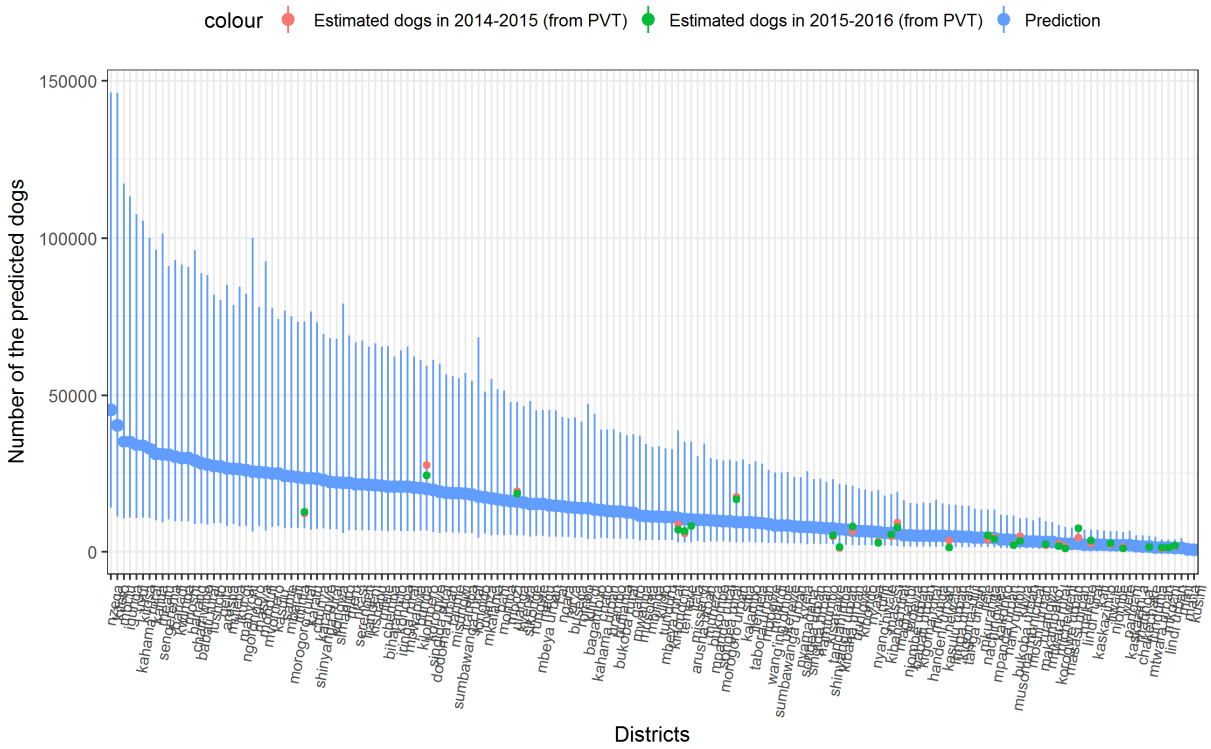

Figure A4. The predicted number of dogs in districts across Tanzania (blue points and their prediction interval) compared to dog populations estimated directly from transects were conducted in 2014-2015 (red points) against 2015-2016 transects (green points).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Table A1. Descriptive characteristics of the study districts. Number of households and employment status were obtained from the 2012 national census [16]. Dogs counted during surveys, mean dogs per household and vaccination coverage were compiled or calculated from household- or school-based surveys. Human: dog ratios (HDRs) were calculated from transects and numbers of vaccinated dogs were compiled from mass dog vaccination campaigns, both conducted in 2014-2015. HHS = Household survey. SBS = School-based survey. CI = Confidence interval. NA = Not applicable

			Employment status			HHS		SBS				Estimated dogs	
District	**Setting**	**Number of households (% rural)**	**Peasants**	**Livestock keepers**	**Others**	**Dogs (mean dogs/ HH)**	**Coverage (%)**	**Dogs (mean dogs /pupil)**	**Coverage (%)**	**HDR (CI)***	**Vaccinated dogs****	**HHS**	**SBS**
Chakechake	Coastal	17,551 (47)	51.1	0.5	48.4	7 (0.04)	0	3 (0.02)	100	59 (20, 177)	608	702	351
Ilala	Coastal	297,750 (9)	4	1.2	94.8	34 (0.26.)	50	NA	NA	NA	4,218	77,415	NA
Kibaha Rural	Inland	16,892 (28)	47.2	5.8	47	30 (0.32)	23	199 (0.50)	59	13 (4, 40)	5,226	5,405	8,446
Kibaha Urban	Inland	31,092 (25)	27.7	2.7	69.6	66 (0.44)	44	237 (0.51)	59	21 (7, 65)	3,684	13,680	15,857
Kilombero	Inland	93,331 (37)	78.7	0.7	20.6	32 (0.22)	47	218 (0.40)	66	21 (7, 63)	14,208	20,533	37,332
Kilwa	Coastal	42,596 (48)	71.6	0.4	28	26 (0.16)	15	NA	NA	55 (18, 166)	2,120	6,815	NA
Kinondoni	Coastal	441,240 (8)	2.9	0.8	96.3	59 (0.32)	44	163 (0.35)	52	181(50, 648)	3,696	141,197	154,434
Kisarawe	Inland	25,475 (48)	78.7	2.3	19	9 (0.05)	89	109 (0.39)	80	14 (4, 43)	2,787	1,274	9,935
Lindi Rural	Coastal	52,821 (42)	87.2	0.1	12.7	15 (0.08)	60	60 (0.24)	55	83 (27, 257)	1,148	4,226	12,677
Lindi Urban	Coastal	22,344 (29)	55.8	0.3	43.9	17 (0.10)	41	70 (0.20)	50	68 (23, 205)	930	2,234	4,469
Liwale	Inland	21,084 (50)	76.2	0.1	23.7	19 (0.11)	16	NA	NA	43 (14, 132)	637	2,319	NA
Masasi	Inland	67,872 (45)	82.2	0.1	17.7	27 (0.15)	30	32 (0.20)	69	42 (14, 127)	4,558	10,181	13,574
Micheweni	Coastal	19,257 (52)	57.3	1.2	41.5	25 (0.14)	0	4 (0.03)	25	41 (14, 125)	569	2,696	578
Mkoani	Coastal	18,067 (57)	26	0.6	73.4	9 (0.06)	100	8 (0.05)	100	75 (25, 224)	631	1,084	903
Mkuranga	Coastal	51,101 (34)	68.5	0.5	31	4 (0.02)	100	58 (0.18)	62	52 (17, 157)	1,811	1,022	9,198
Morogoro Rural	Inland	67,671 (46)	81	3	16	41 (0.24)	56	103 (0.25)	65	12 (4, 39)	6,434	16,241	16,918

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Table A1. Cont.

			Employment status			HHS		SBS				Estimated dogs	
District	**Setting**	**Number of households (% rural)**	**Peasants**	**Livestock keepers**	**Others**	**Dogs (mean dogs/ HH)**	**Coverage (%)**	**Dogs (mean dogs /pupil)**	**Coverage (%)**	**HDR (CI)***	**Vaccinated dogs****	**HHS**	**SBS**
Morogoro Urban	Inland	76,039 (16)	25.2	0.7	74.1	49 (0.29)	76	225 (0.40)	67	35 (12, 108)	9,968	22,051	30,416
Mtwara Rural	Inland	58,602 (37)	83.2	0.1	16.7	16 (0.11)	50	31 (0.09)	61	85 (27, 262)	860	6,446	5,274
Mtwara Urban	Coastal	27,968 (18)	27.7	0.4	71.9	14 (0.09)	64	69 (0.24)	35	84 (28, 252)	540	2,517	6,712
Nachingwea	Inland	48,145 (45)	85.7	0.1	14.2	37 (0.22)	32	84 (0.25)	8	41 (12, 125)	2,823	10,592	12,036
Nanyumbu	Inland	40,746 (33)	88.4	0.1	11.5	1 (0.01)	100	28 (0.06)	18	41 (13, 123)	1,281	407	2,445
Newala	Inland	58,035 (49)	81.2	0.1	18.7	4 (0.02)	100	55 (0.09)	53	42 (14, 128)	1,465	1,161	5,223
Ruangwa	Inland	37,326 (47)	83.7	0.1	16.2	37 (0.21)	35	24 (0.14)	50	42 (14, 126)	1,090	7,838	5,226
Rufiji	Coastal	48,164 (31)	77.1	1.3	21.6	2 (0.01)	100	61 (0.14)	49	35 (11, 109)	1,423	482	6,743
Tandahimba	Inland	60,872 (44)	86.2	0.2	13.6	2 (0.03)	100	24 (0.14)	8	32 (11, 95)	762	1,826	8,522
Temeke	Coastal	344,391 (6)	4.6	0.8	94.6	8 (0.05)	88	NA	NA	147 (44, 496)	2,521	17,220	NA
Ulanga	Inland	53,290 (42)	83.2	1.1	15.7	85 (0.48)	45	326 (0.58)	61	18 (6, 53)	9,645	25,579	30,908
Wete	Coastal	20,151 (40)	47.2	0.8	52	56 (0.38)	57	7 (0.04)	100	52 (17, 155)	718	7,657	806
Overall						731 (0.6)	56	2,198 (0.7)	56		86,321	410,800	398,983

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 July 2018 doi:10.20944/preprints201807.0090.v1

Peer-reviewed version available at Vet. Sci. 2018, 5, 77; doi:10.3390/vetsci5030077

Table A2. A stepwise comparison of Variance Inflation Factors (VIFs) for variables that may influence dog populations in Tanzania. Predictor variables with highest VIF values were removed and the stepwise comparison repeated until all VIF values were below 5. NA=not applicable after the predictor was dropped.

Variables	**VIF after step:**			
	1	2	3	4
Human population in 2014	972.8	9.5	2.4	2.9
Number of people living in rural areas	9.5	NA	NA	NA
Proportion of livestock keeping households	8.8	7.8	5.4	NA
Number of households	948.9	2.9	NA	NA
Proportion of persons employed as livestock keepers	1.6	1.6	1.4	1.3
Setting (inland vs coastal)	1.9	1.9	1.8	1.5
Setting (Mainland vs island)	9.2	5.3	4.7	2.1
Proportion of persons employed as peasant	9.4	8.8	5.4	4.7
Area	8.2	4.3	4.1	3.1