



## Research paper

# Prediction of lymph node parasite load from clinical data in dogs with leishmaniasis: An application of radial basis artificial neural networks



Rafaela Beatriz Pintor Torrecilha<sup>a</sup>, Yuri Tani Utsunomiya<sup>b</sup>, Luís Fábio da Silva Batista<sup>c</sup>, Anelise Maria Bosco<sup>a</sup>, Cárís Maroni Nunes<sup>d</sup>, Paulo César Ciarlini<sup>a</sup>, Márcia Dalastra Laurenti<sup>c,\*</sup>

<sup>a</sup> São Paulo State University (Unesp). School of Veterinary Medicine, Araçatuba. Department of Clinics, Surgery and Animal Reproduction. São Paulo, Brazil

<sup>b</sup> São Paulo State University (Unesp). School of Agricultural and Veterinarian Sciences, Jaboticabal. Department of Preventative Veterinary Medicine and Animal Reproduction. São Paulo, Brazil

<sup>c</sup> USP – Universidade de São Paulo, Departamento de Patologia Veterinária, Faculdade de Medicina Veterinária e Zootecnia, São Paulo, Brazil

<sup>d</sup> São Paulo State University (Unesp). School of Veterinary Medicine, Araçatuba. Department of Support, Production and Animal Health. São Paulo, Brazil

## ARTICLE INFO

## Article history:

Received 24 September 2016

Received in revised form

16 December 2016

Accepted 18 December 2016

## Keywords:

*Canis lupus familiaris*

qPCR

Machine learning

*Leishmania* spp.

## ABSTRACT

Quantification of *Leishmania infantum* load via real-time quantitative polymerase chain reaction (qPCR) in lymph node aspirates is an accurate tool for diagnostics, surveillance and therapeutics follow-up in dogs with leishmaniasis. However, qPCR requires infrastructure and technical training that is not always available commercially or in public services. Here, we used a machine learning technique, namely Radial Basis Artificial Neural Network, to assess whether parasite load could be learned from clinical data (serological test, biochemical markers and physical signs). By comparing 18 different combinations of input clinical data, we found that parasite load can be accurately predicted using a relatively small reference set of 35 naturally infected dogs and 20 controls. In the best case scenario (use of all clinical data), predictions presented no bias or inflation and an accuracy (i.e., correlation between true and predicted values) of 0.869, corresponding to an average error of  $\pm 38.2$  parasites per unit of volume. We conclude that reasonable estimates of *L. infantum* load from lymph node aspirates can be obtained from clinical records when qPCR services are not available.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Visceral leishmaniasis (VL) is a zoonotic disease caused by the protozoan parasite *Leishmania infantum* (synonym *Leishmania chagasi*) (WHO, 2010). Over 90% of the global prevalence of VL is concentrated in India, Bangladesh, Sudan, South Sudan, Ethiopia and Brazil (Alvar et al., 2012). In the Americas, parasites are typically transmitted from host to host by the bites of female *Lutzomyia longipalpis* sand flies. After being inoculated in the skin, promastigote parasites are phagocytized by macrophages or dendritic cells, where they turn into amastigotes, multiply, and disseminate to various tissues (Solano-Gallego et al., 2009, 2011)

The domestic dog (*Canis lupus familiaris*) is a key animal model for VL. Dogs are the most important urban reservoirs of the disease (Dantas-Torres, 2007; Gramiccia and Gradoni, 2005) (Gramiccia

and Gradoni, 2005; Dantas-Torres, 2007) and are also highly susceptible to *L. infantum*. They can express a wide range of physical signs such as those of dermatitis, alopecia, hyperkeratosis, lymphadenomegaly, ophthalmic alterations, pallor of mucous membranes, splenomegaly, emaciation, fever, epistaxis and onychogryphosis (Baneth et al., 2008). They can also present high levels of circulating anti-*L. infantum* antibodies (Lima et al., 2005) and alterations of biochemical findings, such as uremia and hyperglobulinemia (Freitas et al., 2012). Importantly, the occurrence of such alterations have been previously associated with increased parasite load (PL) in tissues such as lymph nodes, bone marrow, spleen (Manna et al., 2009; Reis et al., 2006) and other peripheral tissues (Torrecilha et al., 2016).

Since PL is associated with clinical alterations, one can hypothesize that patients with similar levels of infection might also have similar clinical profiles, such that knowledge of PL may have a prognostic as well as an epidemiological value (Martínez et al., 2011), which may directly impact therapeutic conduct and decisions in disease control. This association could also be turned into the opposite direction, namely clinical data could be used to predict PL.

\* Corresponding author at: USP – Univ de São Paulo. Av Dr Arnaldo, 455–sala 1209, Cerqueira César, São Paulo, 01246–903 Brazil.  
E-mail address: [mdlauren@usp.br](mailto:mdlauren@usp.br) (M.D. Laurenti).

Given PL has typically been determined by limiting dilution (Titus et al., 1985) and real-time quantitative polymerase chain reaction (qPCR) (Francino et al., 2006), which is costly and demands specialized equipment and properly trained personnel, prediction of unobserved PL from clinical data could be an useful tool to generate information about the extent of infection in situations where direct quantification is impracticable.

One important issue in this prediction problem is that the precise nature of the relationships between various clinical parameters and PL is largely unknown. Therefore, devising an accurate statistical model, at least relying on current knowledge, seems intractable without heuristics. In this situation, machine learning techniques come in handy, since they require no explicit model of the relationships between predictand (output) and predictors (inputs), such that they can recognize (non)linear patterns arising from data structure implicitly (Hastie et al., 2009). Radial Basis Artificial Neural Network (RB-ANN) is an example of such technique (Orr, 1996), which has been successfully applied to a wide range of problems, from adaptive flight control (Yilmaz and Özer, 2009) to speech recognition (Venkateswarlu et al., 2011).

In order to test whether clinical data could reliably predict PL, we used a dataset comprised of 35 naturally infected dogs and 20 controls to assess the accuracy of predictions made by a RB-ANN machine trained on an example set of physical signs, serological test and biochemical markers.

## 2. Material and methods

### 2.1. Dog samples

A total of 55 owned dogs, including 35 naturally infected dogs and 20 controls, were sampled from 7 distinct endemic areas, distributed among the Brazilian states of Bahia, Minas Gerais, São Paulo and the Federal District. Breeds included English Bull-dog (n=2), Cocker Spaniel (n=1), Doberman Pinscher (n=1), Labrador Retriever (n=20), Golden Retriever (n=2), German Shepherd (n=8), Belgian Shepherd (n=6), White Swiss Shepherd (n=1), Toy Poodle (n=1), Rottweiler (n=11), German Spitz (n=1) and Dachshund (n=1). Samples were obtained according to the Ethical Principles in Animal Research by the Brazilian College of Animal Experimentation and approved by the Ethics Committee in the use of animals of the School of Veterinary Medicine and Animal Science from University of São Paulo (Protocol 2391/2011).

### 2.2. Parasite load data

Lymph node (popliteal or pre-scapular) fine needle aspiration biopsies were collected, maintained in NET buffer (0.15 M NaCl, 50 mM EDTA, 0.1 M Tris-HCl, pH 7.5) in a final volume of 200  $\mu$ L and stored at 4 °C. DNA was extracted using a commercial kit (NucleoSpin<sup>®</sup> Tissue, Machery Nagel, Germany), according to the manufacturer's recommendations. Quantity and quality of DNA were evaluated in a Biospectrometer (Eppendorf, German), and DNA samples were stored at -20 °C until used. For DNA amplification, primers targeting conserved regions of *Leishmania* kDNA were used, namely LEISH-1 (5'-AACTTTCTGGTCTCCGGGTAG-3') and LEISH-2 (5'-ACCCAGTTTCCGCC-3'). Additionally, eukaryotic 18S was used as internal reference of canine DNA (Francino et al., 2006). Amplification was performed in a final volume of 15  $\mu$ L that consisted of 5  $\mu$ L of total DNA (10 ng/ $\mu$ L), 7.5  $\mu$ L of KAPA SYBR<sup>®</sup> FAST Universal 2X qPCR Master Mix (Kapa Biosystems, USA), 0.5  $\mu$ L of each primer at a final concentration of 300 nM and 1.5  $\mu$ L of deionized water. PCR amplification was carried out in a Realplex<sup>®</sup> thermocycler (Eppendorf, German) using the following conditions: 95 °C for 4 min, followed by 35 cycles of denaturation

at 95 °C for 15 s, annealing of primers at 58 °C for 20 s and extension at 72 °C for 8 s. A standard curve (slope = -3.327, intercept = 28.25, efficiency = 1.00, R<sup>2</sup> = 0.929) was constructed with dilutions of *L. infantum* DNA (MHOM/BR/72/strain46) from  $0.5 \times 10^6$  to  $0.5 \times 10^{-4}$  parasites/ $\mu$ L. The amplifications were performed in duplicate and parasite load was obtained via regression of the cycle threshold (Ct) values against the standardized parasite concentrations (Aschar et al., 2016) in Realplex<sup>®</sup> Software (Eppendorf, German). The results were defined as the number of parasites in volume of lymph node aspirate which yielded 10 ng of DNA. Parasite load were then transformed to a log<sub>10</sub>(PL + 1) scale for analysis.

### 2.3. Clinical data

Dog owners provided data on sex, age, vaccination and the use of any form of repellent (e.g., collars or spot-ons). These variables were included as input in all prediction analyses (NULL input set).

All dogs were evaluated by veterinary experts to determine the presence/absence of common physical signs found in leishmaniasis, such as skin alterations (dermatitis, alopecia and hyperkeratosis), ophthalmic alterations (uveitis, blepharitis, ulcer and secretion), onychogryphosis, lymphadenomegaly, pallor of mucous membranes and emaciation. These physical signs were used as either separate inputs (SIGNS1) or encoded as a single indicator variable (SIGNS2). In the first case, indicator variables (0 = absence, 1 = presence) were used for each physical sign individually. In the second case, animals were classified as sick or clinically healthy based on the presence of at least one physical sign.

Serum samples were obtained from whole blood collected via venipuncture for biochemical markers analysis and quantification of anti-*L. infantum* IgG levels. The quantification of biochemical markers was performed in an automated spectrophotometer (BS 200, Shenzhen Mindray Bio-Medical Electronics Co., Nanshan, China), previously calibrated with serum control levels I and II (Biosystems, Barcelona, Spain). The biochemical markers and their respective quantification methods were: albumin (g/L) by bromocresol green method; urea (mg/dL) by the urease/glutamate dehydrogenase coupled with the UV enzymatic assay; creatinine (mg/dL) by the kinetic alkaline picrate assay; aspartate aminotransferase (AST, IU/L), alanine aminotransferase (ALT, IU/L) by enzymatic UV method, following the International Federation of Clinical Chemistry (IFCC) guidelines; bilirubin (mg/dL) by the diazotized sulfanilic method; uric acid (mg/dL) by the uricase/peroxidase enzymatic method and total plasma protein (g/L) by the biuret method. Globulin levels (g/L) were determined by the difference between total protein and albumin. Biochemical markers were treated as quantitative inputs (BIOCHEM1) or as indicator variables (BIOCHEM2) indicating levels above or below reference values, according to Kaneko et al. (1997).

Indirect enzyme-linked immunosorbent assay (ELISA) using crude *L. infantum* antigen (MHOM/BR/72/strain46) and anti-canine IgG (A40-123AP) conjugated to alkaline phosphatase (Bethyl, USA) was performed according to Laurenti et al., 2014 to evaluate the presence of anti *L. infantum* IgG in serum samples. Eleven dogs from non-endemic area, with negative cytological examination were used to establish the cut line (cut-off) of reaction that refers to the average of the negative controls added to three standard deviations.

### 2.4. Non-mathematical description of the prediction machine

The objective of this section is to provide a 'pictorial' description of the prediction problem treated in this paper, which may be useful to help grasping the general idea of a RB-ANN. We make clear, however, that this description is shallow and should be interpreted with caution when dissociated from its mathematical formalism. The problem is stated as follows: we have access to clinical data

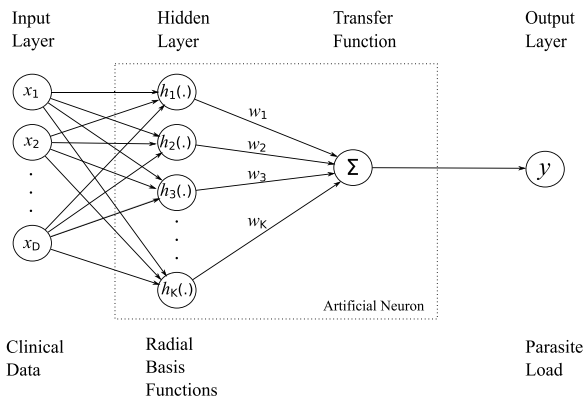


Fig. 1. Schematic representation of the radial basis artificial neural network.

of patients for which we wish to obtain PL data. Therefore, we want a machine that can take clinical data as input and output PL. We also have access to historical records of other patients with both clinical data and PL that could be used by this machine to train its predictions. Importantly, the machine should learn from training examples using the entire set of inputs, that is, without pre-selection of variables that humans might find informative. Here we used one particular type of such a machine, namely RB-ANN, which metaphorically operates similar to a network of neurons: stimuli (input data) are fed to several neurons (radial basis functions) which will be activated and react differently according to their previous experiences (learned weights) (Fig. 1). Appropriate responses can only be obtained if neurons have encountered similar stimuli in the past, such that training is a key step for a RB-ANN machine.

### 2.5. Mathematical description of the prediction machine

In this section we provide a self-sufficient description for the purpose of stating the technique applied in this paper, but more complete descriptions for RB-ANN and other types of neural networks can be found elsewhere, e.g. (Hastie et al., 2009; Orr, 1996).

In respect to a given patient, let  $y$  be a scalar representing its  $\log_{10}(\text{PL} + 1)$  value, which we wish to predict using some function of a  $D$ -dimensional vector of clinical inputs  $\mathbf{x}$ , i.e.  $y = f(\mathbf{x})$ . In short, we can approximate this function using a linear combination of basis functions:

$$f(\mathbf{x}) = \sum_{k=1}^k w_k h_k(\mathbf{x}) \quad (1)$$

where  $h_k(\mathbf{x})$  is a Gaussian radial function at basis  $k$ , and  $w_k$  is the weight of that function. The Gaussian function is defined as:

$$h_k(\mathbf{x}) = \exp \left\{ -\gamma \|\mathbf{x} - \boldsymbol{\mu}_k\|^2 \right\} \quad (2)$$

where  $\boldsymbol{\mu}_k$  represents the centers at basis  $k$ ,  $\|\cdot\|$  is the Euclidean norm,  $\gamma = 1/2r^2$  and  $r$  is the radius (bandwidth) parameter (i.e., standard deviation of the Gaussian function). Given training data, the objective of this analysis is to use each patient as a basis, i.e. define  $\boldsymbol{\mu}_k = \mathbf{x}_k$ , and finding weights for each basis.

After weights have been learned, predictions are conducted considering  $\mathbf{x}$  as the input layer, the collection of functions (neurons)  $h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_K(\mathbf{x})$  as the hidden layer with weights  $w_1, w_2, \dots, w_K$ , and the weighted sum of radial basis functions as the output layer (Fig. 1).

The problem of finding appropriate weights is solved as follows. In matrix notation, let  $\mathbf{y}$  be the  $K \times 1$  vector of  $K$  training observations,  $\mathbf{X}$  be the  $K \times D$  matrix of  $D$  training inputs,  $\mathbf{H}$  the  $K \times K$  matrix of radial basis functions, and  $\mathbf{w}$  the  $K \times 1$  vector of unobserved weights.

For the sake of numerical stability, both  $\mathbf{y}$  and  $\mathbf{X}$  have their columns centered on their means and scaled by their maximum, such that all variables have values ranging from  $-1$  to  $1$ . We seek solutions for  $\mathbf{w}$  that minimize the penalized sum of squared prediction errors:

$$\hat{\mathbf{w}} = \arg_{\mathbf{w}} \min \|\mathbf{y} - \mathbf{H}\mathbf{w}\|^2 + \delta \|\mathbf{w}\|^2 \quad (3)$$

where  $\delta$  is a regularization parameter that controls the amount of shrinkage towards zero exerted on the weights. Minimization is achieved by setting the partial derivatives of the cost function in respect to  $\mathbf{w}$  to zero, which yields the ridge regression equations:

$$\hat{\mathbf{w}} = (\mathbf{H}'\mathbf{H} + \mathbf{I}\delta)^{-1} \mathbf{H}'\mathbf{y} \quad (4)$$

The  $\delta$  and  $r$  parameters were chosen from a grid search in the space  $0-10$  and  $0.1-1.0$ , respectively. The combination of values presenting the lowest mean squared error (MSE) was adopted.

### 2.6. Accuracy of predictions

Predictions were carried out using a leave-one-out cross-validation procedure. Briefly, a given animal was tested by training the RB-ANN on the remaining dogs. Accuracy of parasite load value predictions was measured by mean squared errors and Pearson's product-moment correlation coefficient between observed and predicted values. Additionally, bias and inflation of predictions were measured by regressing observed values onto predicted values. The bias and inflation were defined as the intercept and the slope of that regression, respectively. An intercept close to zero indicates absence of bias. A slope greater than one indicates inflation, whereas a slope smaller than one indicates deflation of predictions.

A total of 18 different prediction scenarios were tested, which differed in terms of the set of inputs (ELISA, SIGNS1, SIGNS2, BIOCHEM1 and BIOCHEM2) combined to train the RB-ANN machine. These different scenarios were intended to evaluate the influence of specific sets of inputs, as well as their respective representations, on the accuracy of PL predictions. Additionally, the power of our study was assessed through bootstrap sampling of the training set. We tested the performance of the RB-ANN machine with numbers varying from 10 to 55 individuals. For each training set size, the expected accuracy of predictions was computed from the average across 100 bootstrap samples. An accuracy of at least 80% was then used to determine the minimum sample size required for reliable predictions. All analyses were carried out in R v3.3.0 (available at: <https://www.r-project.org/>) using customized code.

## 3. Results

The percentages of male, sick, vaccinated, repellent-protected and ELISA-positive dogs were 47.3%, 61.8%, 20.0%, 41.8% and 45.5%, respectively. Ages ranged from 11 months to 13 years, with an average of  $4.00 \pm 2.63$  years (median = 3 years, IQR = 3.5 years). Table 1 shows the summary statistics for the parasite load and the biochemical markers. Common impaired biochemical statuses included decreased serum levels of ALT (43.4%), AST (55.6%) and bilirubin (57.1%), increased creatinine (22.2%), urea (14.5%) and uric acid (14.6%), hypoalbuminemia (38.2%), hyperalbuminemia (21.8%), hyperproteinemia (43.6%), and hyperglobulinemia (47.3%). A small percentage of the age (3.6%), ALT (3.6%), AST (1.8%), creatinine (1.8%) and bilirubin (10.9%) data were missing and had their values imputed based on their respective averages prior to the prediction analysis.

After cross-validation of the RB-ANN machine with all possible combinations of clinical inputs (Table 2), the most accurate predictions were obtained when all sources of clinical data were

**Table 1**  
Summary statistics for parasite load and biochemical markers.

Variable	Minimum	Mean	Median	Standard deviation	Maximum	% Below reference	% Above reference	Reference values	Unit
Parasite load	0.00	2,310.52	3.54	13,515.15	99,458.00	–	–	–	Parasites per volume
Albumin	2.62	26.30	28.14	7.79	41.04	38.2	21.8	26–33	g/L
ALT	1.78	31.23	26.41	32.75	221.59	43.4	0	21–102	IU/L
AST	3.97	22.31	21.12	13.13	64.01	55.6	0	23–66	IU/L
Bilirrubin	0.02	0.09	0.09	0.04	0.18	57.1	0	0.1–0.5	mg/dL
Creatinine	0.58	1.35	1.30	0.40	3.01	0	22.2	0.5–1.5	mg/dL
Globulin	28.25	47.07	41.35	14.59	78.63	0	47.3	27–44	g/L
Ureia	20.66	41.12	37.52	16.32	130.23	0	14.5	10.03–50.03	mg/dL
Uric acid	0.24	1.38	1.20	1.20	7.70	0	14.6	0–2	mg/dL
Protein	50.67	73.37	70.03	12.37	101.50	0	43.6	54–71	g/L

**Table 2**  
Accuracy (mean squared error and correlation), bias and inflation of predictions.

Input set	Mean squared error	Correlation	Bias	Inflation
NULL <sup>a</sup>	1.379	0.192	0.514	0.468
ELISA <sup>b</sup>	0.891	0.595	0.106	0.879
SIGNS1 <sup>c</sup>	0.879	0.616	0.176	0.801
SIGNS2 <sup>d</sup>	0.991	0.526	0.057	0.911
BIOCHEM1 <sup>e</sup>	0.933	0.617	0.256	0.710
BIOCHEM2 <sup>f</sup>	0.702	0.718	0.185	0.807
ELISA + SIGNS1	0.678	0.716	0.092	0.884
ELISA + SIGNS2	0.608	0.745	–0.047	1.005
ELISA + BIOCHEM1	0.734	0.688	0.056	0.886
ELISA + BIOCHEM2	0.490	0.801	–0.016	0.983
SIGNS1 + BIOCHEM1	1.016	0.588	0.284	0.665
SIGNS1 + BIOCHEM2	0.936	0.631	0.277	0.686
SIGNS2 + BIOCHEM1	0.782	0.651	0.203	0.758
SIGNS2 + BIOCHEM2	0.739	0.681	0.096	0.908
ELISA + SIGNS1 + BIOCHEM1	0.708	0.715	0.125	0.814
ELISA + SIGNS1 + BIOCHEM2	0.695	0.729	0.167	0.786
ELISA + SIGNS2 + BIOCHEM1	0.339	0.869	0.000	0.950
ELISA + SIGNS2 + BIOCHEM2	0.488	0.804	0.019	0.937

<sup>a</sup> NULL = base model including only age, sex, vaccination and use of repellent.

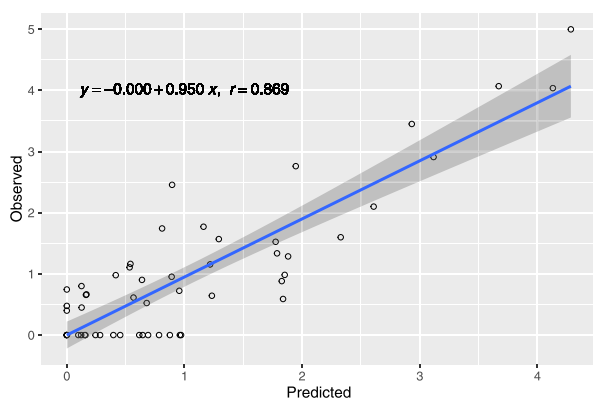
<sup>b</sup> ELISA = base model + positivity in the serological test.

<sup>c</sup> SIGNS1 = base model + presence/absence of separate physical signs. Only physical signs with at least three affected dogs were included, which comprised pallor of mucous membranes, lymphadenomegaly, skin alterations and ophthalmic alterations.

<sup>d</sup> SIGNS2 = base model + presence/absence of at least one physical sign.

<sup>e</sup> BIOCHEM1 = base model + biochemical markers.

<sup>f</sup> BIOCHEM2 = base model + biochemical markers recoded as dummy variables indicating whether the observed value was above or below the reference value reported by (Kaneko et al., 1997).



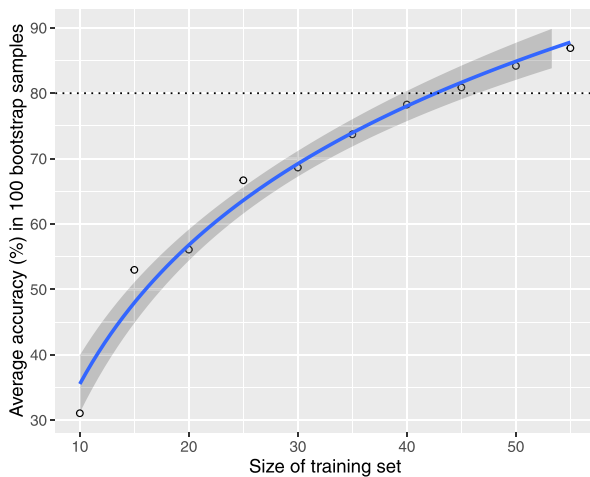
**Fig. 2.** Regression of observed (y-axis) onto predicted (x-axis) parasite load values. The shaded area surrounding the regression line represents the 95% confidence interval.

simultaneously considered, provided that physical signs were encoded by a single presence/absence variable. In this scenario, the correlation between predicted and observed PL was 0.869 (Fig. 2), and predictions were unbiased and non-inflated, with an average prediction error of  $\pm 38.2$  parasites per unit of volume (Supplementary Table 1). In general, all possible combinations of clinical data yielded accuracies of at least 0.526, which substantially dif-

fered from the lower boundary accuracy of 0.192 imposed by the null model. When sources of input were considered separately, the biochemical marker data presented the highest accuracy (0.718), whereas the serological test was the least reliable (0.595). Finally, our power assessment revealed that a minimum sample size of 43 individuals, including infected dogs and controls, would be required to achieve a prediction accuracy of at least 80% (Fig. 3). This result indicates that our sample of 55 dogs had sufficient power to assess prediction accuracies.

#### 4. Discussion

Lymph node aspirates present high sensitivity in the detection of *L. infatum* parasites by qPCR (Martínez et al., 2011; Moreira et al., 2007). Increased parasite load measured via qPCR have been found to correlate with impairment of serum biochemistry, occurrence of physical signs and positivity in serological tests (Manna et al., 2009; Reis et al., 2006; Torrecilha et al., 2016). Consequently, serology, physical signs and laboratory findings are used for disease stage classification and prognosis (Solano-Gallego et al., 2011, 2009). Although qPCR quantification of parasite load has been proposed as a practical tool for accurate diagnosis (Martínez et al., 2011; Solano-Gallego et al., 2011) and monitoring therapy (Solano-gallego et al., 2016), its routine use in clinics and in epidemiological surveillance is still hampered by financial and technical constrains. Here, we used a machine learning technique to predict parasite load from



**Fig. 3.** Power analysis. Power was measured from average prediction accuracies over 100 bootstrap samples for each size of training set. The shaded area around the regression line represents the 95% confidence interval. The dashed horizontal line marks the 80% threshold, corresponding to an approximate sample size of 43 dogs.

clinical data (serological test, biochemical markers and physical signs), as an alternative for quantification in the absence of qPCR.

Radial Basis Artificial Neural Network is a flexible, easy-to-use machine learning technique that can be trained to predict unobserved values using nothing more than example data (Orr, 1996) and an ordinary laptop computer. This technique is capable of universal approximation (Park and Sandberg, 1991), that is, artificial neurons can accurately represent virtually any underlying function mapping inputs to outputs. This feature is very convenient because unknown biological relationships can be approximated even in the absence of knowledge about the underlying phenomenon. Indeed, even without explicit modeling, the Radial Basis Artificial Neural Network was fairly accurate in estimating lymph node parasite load from a relatively small training set of clinical data. Overall, our analysis showed that exams that are routinely at the disposal of veterinary practitioners could be fed into a neural network to estimate parasitism in dogs from endemic areas where services and infrastructure required for qPCR are not available or are still too costly.

Our results indicate that the best prediction performance is achieved when quantitative biochemical marker data are combined with indicator variables (positive/negative) for the presence of physical signs and positivity in a serological test against *L. infantum*. Although the serological test has been shown to be highly sensitive (Lima et al., 2005), it performed poorer than physical sign and biochemical marker data in predicting parasite load. In fact, biochemical marker data were found to have the highest predictive power among all clinical inputs. Our results also showed that reasonable estimates could be obtained from different combinations of clinical inputs, which is convenient in scenarios where a given exam is not available.

Future questions to be addressed include whether predictions could be improved even further by increasing the number of animals in the training set and by inclusion of other clinical inputs. Assessment of predictions of parasite load in other tissues might be also of interest.

## 5. Conclusions

Radial basis artificial neural networks can accurately predict lymph node *L. infantum* load in dogs from physical sign, biochemical marker and serological data. Provided reference data is accessible for training, reasonable estimates of parasite load can be obtained

from an ordinary laptop computer when infrastructure for qPCR is not available.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.vetpar.2016.12.016>.

## References

- Alvar, J., Vélez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., de Boer, M., 2012. Leishmaniasis worldwide and global estimates of its incidence. *PLoS One*, <http://dx.doi.org/10.1371/journal.pone.0035671>.
- Aschar, M., de Oliveira, E.T.B., Laurenti, M.D., Marcondes, M., Tolesano, J.E., Hiramoto, R.M., Corbett, C.E.P., da Matta, V.L.R., 2016. Value of the oral swab for the molecular diagnosis of dogs in different stages of infection with *Leishmania infantum*. *Vet. Parasitol.* 225, 108–113, <http://dx.doi.org/10.1016/j.vetpar.2016.06.005>.
- Baneth, G., Koutinas, A.F., Solano-Gallego, L., Bourdeau, P., Ferrer, L., 2008. Canine leishmaniasis – new concepts and insights on an expanding zoonosis: part one. *Trends Parasitol.* 24, 324–330, <http://dx.doi.org/10.1016/j.pt.2008.04.001>.
- Dantas-Torres, F., 2007. The role of dogs as reservoirs of *Leishmania* parasites, with emphasis on *Leishmania (Leishmania) infantum* and *Leishmania (Viannia) braziliensis*. *Vet. Parasitol.* <http://dx.doi.org/10.1016/j.vetpar.2007.07.007>.
- Francino, O., Altet, L., Sánchez-Robert, E., Rodríguez, a, Solano-Gallego, L., Alberola, J., Ferrer, L., Sánchez, a, Roura, X., 2006. Advantages of real-time PCR assay for diagnosis and monitoring of canine leishmaniasis. *Vet. Parasitol.* 137, 214–221, <http://dx.doi.org/10.1016/j.vetpar.2006.01.011>.
- De Freitas, J.C.C., Nunes-Pinheiro, D.C.S., Lopes Neto, B.E., Santos, G.J.L., De Abreu, C.R.A., Braga, R.R., Campos, R.D.M., De Oliveira, L.F., 2012. Clinical and laboratory alterations in dogs naturally infected by *Leishmania chagasi*. *Rev. Soc. Bras. Med. Trop.* 45, 24–29, <http://dx.doi.org/10.1590/S0037-86822012000100006>.
- Gramiccia, M., Gradoni, L., 2005. The current status of zoonotic leishmaniasis and approaches to disease control. *Int. J. Parasitol.* 35, 1169–1180, <http://dx.doi.org/10.1016/j.ijpara.2005.07.001>.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning*, 2nd ed. Springer, California (764 pp).
- Kaneko, J.J., Harvey, J.W., Bruss, M.L., Vernau, W., Vernau, K.A., Sue Bailey, C., 1997. *Clinical Biochemistry of Domestic Animals*, 5th ed. Academic Press, California (932 pp).
- Laurenti, M.D., de Santana Leandro, M.V., Tomokane, T.Y., De Lucca, H.R.L., Aschar, M., Souza, C.S.F., Silva, R.M., Marcondes, M., da Matta, V.L.R., 2014. Comparative evaluation of the DPP® CVL rapid test for canine serodiagnosis in area of visceral leishmaniasis. *Vet. Parasitol.* 205, 444–450, <http://dx.doi.org/10.1016/j.vetpar.2014.09.002>.
- Lima, V.M.F., De Biazzone, L., Silva, A.C., Correia, A.P.F.L., Luvizotto, M.C.R., 2005. Serological diagnosis of visceral leishmaniasis by an enzyme immunoassay using protein A in naturally infected dogs. *Pesqui. Veterinária Bras.* 25, 215–218.
- Manna, L., Reale, S., Vitale, F., Gravino, A.E., 2009. Evidence for a relationship between *Leishmania* load and clinical manifestations. *Res. Vet. Sci.* 87, 76–78, <http://dx.doi.org/10.1016/j.rvsc.2008.12.009>.
- Martínez, V., Quilez, J., Sanchez, A., Roura, X., Francino, O., Altet, L., 2011. Canine leishmaniasis: the key points for qPCR result interpretation. *Parasit. Vectors* 4, 57, <http://dx.doi.org/10.1186/1756-3305-4-57>.
- Moreira, M.A.B., Luvizotto, M.C.R., Garcia, J.F., Corbett, C.E.P., Laurenti, M.D., 2007. Comparison of parasitological, immunological and molecular methods for the diagnosis of leishmaniasis in dogs with different clinical signs. *Vet. Parasitol.* 145, 245–252, <http://dx.doi.org/10.1016/j.vetpar.2006.12.012>.
- Orr, M., 1996. Introduction to Radial Basis Function Networks. University of Edinburgh, 67pp <http://www.cc.gatech.edu/~isbell/tutorials/rbf-intro.pdf>.
- Park, J., Sandberg, I.W., 1991. Universal approximation using radial-Basis-Function networks. *Neural Comput.* 3, 246–257, <http://dx.doi.org/10.1162/neco.1991.3.2.246>.
- Reis, A.B., Martins-Filho, O.a., Teixeira-Carvalho, A., Carvalho, M.G., Mayrink, W., França-Silva, J.C., Giunchetti, R.C., Genaro, O., Corrêa-Oliveira, R., 2006. Parasite density and impaired biochemical/hematological status are associated with severe clinical aspects of canine visceral leishmaniasis. *Res. Vet. Sci.* 81, 68–75, <http://dx.doi.org/10.1016/j.rvsc.2005.09.011>.
- Solano-Gallego, L., Koutinas, a, Miró, G., Cardoso, L., Pennisi, M.G., Ferrer, L., Bourdeau, P., Oliva, G., Baneth, G., 2009. Directions for the diagnosis, clinical staging, treatment and prevention of canine leishmaniasis. *Vet. Parasitol.* 165, 1–18, <http://dx.doi.org/10.1016/j.vetpar.2009.05.022>.
- Solano-Gallego, L., Miró, G., Koutinas, A., Cardoso, L., Pennisi, M., Ferrer, L., Bourdeau, P., Oliva, G., Baneth, G., 2011. LeishVet guidelines for the practical management of canine leishmaniasis. *Parasit. Vectors* 4, 86, <http://dx.doi.org/10.1186/1756-3305-4-86>.
- Solano-gallego, L., Filippo, L., Ordeix, Di, Planellas, L., Roura, M., Altet, X., Martínez-orellana, L., 2016. Early reduction of *Leishmania* infantum-specific antibodies and blood parasitemia during treatment in dogs with moderate or

- severe disease. *Parasit. Vectors* 1–9, <http://dx.doi.org/10.1186/s13071-016-1519-0>.
- Titus, R.G., Marchand, M., Boon, T., Louis, J.a., 1985. A limiting dilution assay for quantifying *Leishmania major* in tissues of infected mice. *Parasite Immunol.* 7, 545–555, <http://dx.doi.org/10.1111/j.1365-3024.1985.tb00098.x>.
- Torrecilha, R.B.P., Utsunomiya, Y.T., Bosco, A.M., Almeida, B.F., Pereira, P.P., Narciso, L.G., Pereira, D.C.M., Baptistioli, L., Calvo-Bado, L., Courtenay, O., Nunes, C.M., Ciarlini, P.C., 2016. Correlations between peripheral parasite load and common clinical and laboratory alterations in dogs with visceral leishmaniasis. *Prev. Vet. Med.* 132, 83–87, <http://dx.doi.org/10.1016/j.prevetmed.2016.08.006>.
- Venkateswarlu, R.L.K., Kumari, R.V., Jayasri, G.V., 2011. Speech recognition using radial basis function neural network. 2011 3rd Int. Conf. Electron. Comput. Technol 3, 441–445, <http://dx.doi.org/10.1109/ICECTECH.2011.5941788>.
- WHO Organization, Leishmaniases, C., 2010. *Control of the leishmaniases: report of a meeting of the WHO Expert Committee on the Control of Leishmaniases, Geneva, 22–26 March 2010*. World Health Organ. Tech. Rep. Ser. (xii).
- Yilmaz, A.S., Özer, Z., 2009. Pitch angle control in wind turbines above the rated wind speed by multi-layer perceptron and radial basis function neural networks. *Expert Syst. Appl.* 36, 9767–9775, <http://dx.doi.org/10.1016/j.eswa.2009.02.014>.