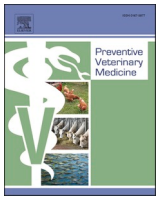
Preventive Veterinary Medicine 248 (2026) 106783

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed



Relative accuracy of point-of-care tests to rule-in heartworm infection in clinically suspected dogs using Bayesian latent class modelling

P.J. Atkinson [a,*,1], C. Quimby [b], A. Datt [b], T.D. Nielsen [a,1], C.G.B. Caraguel [a,c,1]

[a] School of Animal and Veterinary Sciences, The University of Adelaide, Australia
[b] Animals Fiji, Northern and Western Division, Fiji
[c] Sydney School of Veterinary Science, The University of Sydney, Australia

ARTICLE INFO

Keywords:
Veterinary science
Canine heartworm
Evidence-based diagnosis
Bayesian latent
Class model

ABSTRACT

Our prospective cross-sectional survey of dogs from Western Fiji aimed to evaluate the fitness of common diagnostic test modalities to rule-in *Dirofilaria immitis* infection in patients with suggestive clinical history or signs. In the absence of a perfect reference standard, we used latent class modelling to evaluate the relative diagnostic accuracy of two point-of-care (POC) detection modalities – the modified Knott's test (MKT) to detect circulating microfilaria and four antigen lateral flow immunoassays (LFI): Anigen Rapid Canine HW Ag Test® (Bionote Co.), SNAP® Heartworm RT Test (IDEXX Laboratories), trūRapid Heartworm (Heska), WITNESS Dirofilaria® (Zoetis Inc.). The tests' fitness for ruling-in infection were compared using the likelihood ratios of a positive result (LR^+). The performances of the MKT and the Anigen Rapid LFI to rule-in infection on fresh blood of clinically suspected dogs were moderate to strong (LR^+=13.4, 95 %PCI: 6.7–114.6; LR^+=20.2, 95 %PCI: 5.4–138.2; respectively) but not consistently different from each other. The Anigen Rapid, SNAP and truRapid tests consistently provided the strongest evidence to rule-in infection. The LR^+ of the WITNESS test was approximately twelve times, nine times and two times lower than the SNAP, truRapid and Anigen Rapid LFIs respectively (Bayesian p-value 0.002, 0.004 and 0.02 respectively). Overall, a positive result from MKT or LFIs is suitable to rule-in infection in dogs raising clinical suspicion and would increase the post-test probability of infection similarly. If veterinarians are choosing between LFIs, they should favour either Anigen Rapid, SNAP and truRapid over WITNESS.

1. Introduction

Dirofilaria immitis is a parasitic nematode that mainly infects domestic canids and is the causative agent of canine heartworm disease (Bowman and Atkins, 2009). Adult worms live in an infected host's pulmonary arteries, causing vascular inflammation and turbulent blood flow (McCall et al., 2008). The clinical manifestation of infection, dirofilariosis, ranges in severity from mild disease demonstrating coughing and exercise intolerance, to severe disease resulting in dyspnoea, syncope, ascites, right-sided congestive heart failure and death, although pre- and/or sub-clinical phases are most common (McCall et al., 2008; Bowman and Atkins, 2009).

Suspicion of infection by veterinarians starts with suggestive clinical history (e.g., living in an endemic area, lack of dirofilariosis prevention) or clinical signs of dirofilariosis. However, those alone are unlikely to be sufficient to rule-in infection and trigger consideration of therapy. Therefore, complementary *D. immitis* specific diagnostic modalities are required to rule-in infection in clinically suspected individuals.

Until the mid-1980s, the only available ante-mortem modality was the visual detection of an immature larval stage (microfilaria, L1) circulating in an infected canid's blood. Several variations of microfilaria detection have been developed, requiring differing levels of technical expertise, time and equipment and offering different degrees of detection performance (Martini et al., 1991; Mylonakis et al., 2004). However, all are bounded by the natural fluctuations to microfilaraemic concentrations between infected dogs, as well as within dogs due to diurnal periodicity, leading to possible false-negative findings according to the time of blood sampling (Evans et al., 2017). Although less likely

* Correspondence to: University of Adelaide – Roseworthy Campus, Mudla Wirra Rd, Roseworthy, SA 5371, Australia.
E-mail addresses: peter.atkinson@adelaide.edu.au (P.J. Atkinson), casey@animalsfiji.org (C. Quimby), ashmita@animalsfiji.org (A. Datt), torben.nielsen@adelaide.edu.au (T.D. Nielsen), charles.caraguel@adelaide.edu.au (C.G.B. Caraguel).
[1] Equal contribution to the supervision of this work

https://doi.org/10.1016/j.prevetmed.2026.106783
Received 8 October 2025; Received in revised form 29 December 2025; Accepted 5 January 2026
Available online 5 January 2026

than false-negatives, microfilarial testing may also return false-positive results if another filarial species is mis-identified, due to trans-placental/transfusion transfer of microfilaria, persistent micro-filaraemia after adult worm death (up to two years) or sample preparation errors (Bowman and Mannella, 2011; Little et al., 2018; Constantinoiu et al., 2023). Microfilarial detection methods also require a skilled operator and time, which may be limited in some test settings.

Modalities to detect various *D. immitis* antigens circulating in the blood of infected hosts were therefore developed to address these limitations, requiring less resources (at the point of care) to perform. Nowadays, several manufacturers offer commercial point-of-care (POC) antigen detection tests, which use lateral-flow immunoassay (LFI) technology to detect antigens released by the uterus of the adult female *D. immitis* worm (Goodwin, 1998; Rohrbach and Patton, 2013). Less susceptible to natural variability, the detection of this antigen is reported to offer an improved diagnostic sensitivity compared to microfilarial detection (Courtney and Zeng, 1993). Detection of the uterine antigen suggests the presence of at least one adult female worm, but antigens may not be present in immature female- or (any) male-only infections. Given that the potential severity of dirofilariosis is in part related to adult worm burden (Knight, 1987), with high burdens requiring multiple host (re-)infections, the antigen is expected to be present in dogs with clinical disease, although it may also be detected in pre- or sub-clinical cases.

The LFI modality is commonly used in clinical practice and many diagnostic test evaluation studies have been reported to assess their diagnostic accuracy (Atkinson et al., 2023). However, all studies reported diagnostic accuracy estimates relative to one or a combination reference standard, including necropsy, PCR or plate-based ELISA (also targeting the same female *D. immitis* uterine antigen). The authors of those studies assumed their reference standards to be perfectly accurate which may not necessary hold true according to their diagnostic purpose (Wilks, 2001), leading to inaccuracies in LFI performance (Naaktgeboren et al., 2013). Of note is necropsy, which some argue could be a perfect diagnostic test. Whilst its specificity may approach 100 %, its sensitivity is only as high as the scrutiny of pulmonary artery dissection. Unless every pulmonary artery were dissected during necropsy, a definitive infection-free status would be impossible to determine, and may lead to misclassification impacting the performance of any tests under evaluation in diagnostic test accuracy studies.

We implemented latent class modelling (LCM) to evaluate the fitness of current POC heartworm diagnostic modalities (microfilarial detection and LFIs) in domestic dogs that raised clinical suspicion of infection in practice. We selected this population as it reflects the most clinically relevant context in which these tests are used by veterinarians (i.e., to rule-in infection) (Atkinson et al., 2023). We compared the performances of both modalities (microfilarial detection versus LFI) to rule-in two-sex infections, and the performances across commercial LFIs to rule-in the presence of the target (uterine) antigen from a female worm. We anticipated that the commercial LFIs would show consistently similar performance, which would be better than microfilarial detection. We combined the estimates of diagnostic sensitivity and specificity to obtain each modality's and LFI's likelihood ratio of a positive result (LR^+), comparing their fitness to rule-in infection or antigen presence. By incorporating a discussion of each modality's relative costs, we could also assess their overall suitability to various clinical settings, and by ranking LFIs of different brands we could advise veterinarians to select the test/s providing the strongest diagnostic information.

2. Materials and methods

This report complies with the Standards for the Reporting of Diagnostic Accuracy Studies using LCM (STARD-BLCM) reporting guidelines (Kostoulas et al., 2017). Ethical approval for this study was obtained from The University of Adelaide's animal ethics committee (S-2023–081). Identifiable owners of study dogs were informed of the aim of the study and provided written consent. For dogs without an identifiable owner, consent was provided by the manager of the study site (author CQ).

2.1. Study subjects

Our target population were canids living where *D. immitis* infection is endemic with a suspicion of infection based on either suggestive clinical history or report/presence of clinical signs, and broadly applicable to contexts suitable for rule-in testing against *D. immitis* infection in practice. Adapting the outline reported in Atkinson et al. (2023), we defined a dog as 'suspect' if it:

i. was at least six months old (minimum time before a female worm develops into a sexually mature adult worm), AND
ii. presented at least one clinical sign consistent with dirofilariosis (see below), regardless of its prophylaxis history; OR
iii. was reported with at least a three-month interruption in preventative usage, at least six months prior to presentation; OR
iv. had no available history of preventative use (e.g., unidentifiable owner).

From September 2023, we prospectively sampled every suspect dog visiting the Animals Fiji veterinary facility located in the Western Division of Fiji's largest island, Viti Levu, where *D. immitis* is deemed endemic (Symes, 1960; Mataika et al., 1971; Olver, 2022) until our minimum sample size was achieved. We aimed to recruit a minimum of 100 infected or non-infected dogs to achieve acceptable precision for our diagnostic sensitivity or specificity estimates. We assumed that at least 20 % of dogs in this endemic zone would be infected, but the majority of the accessible dogs would be non-infected. Therefore, we focused on sampling a minimum of 100 dogs with evidence of infection, i.e., a positive result to either detection modality, although acknowledge our estimation of prevalence determined the overall sample size goal, and was an approximation. We excluded eligible dogs from testing if testing resources (i.e., labour or consumables) were unavailable at the time of presentation.

Data collected at presentation included domicile location ('unknown' for stray subjects), prophylaxis history (for dogs with identifiable owners) and presence of clinical history or signs associated with *D. immitis* infection: coughing (history and clinical examination), exercise intolerance (history), syncopal episodes (history), dyspnoea (clinical examination), abnormal lung and heart sounds (clinical examination), ascites or abdominal organomegaly (clinical examination), or caval syndrome – acute onset severe weakness with associated haemoglobin-aemia/uria (history and clinical examination) (McCall et al., 2008). Each subject was classified as 'clinically abnormal' if at least one of the listed clinical signs was present or 'apparently healthy' otherwise.

2.2. Specimen collection

Up to 5 mL of whole blood was collected from recruited subjects from either the jugular or cephalic vein/s, of which 1 mL was immediately placed into an EDTA tube to perform the modified Knott's test (MKT) and one of the lateral flow immunoassay (LFI) tests. The remaining sample was placed in one serum clot tube, and was refrigerated at 4℃. After overnight vertical storage, the serum was pipetted for separation and stored at −20℃ at Animals Fiji. Serum samples were later transferred to The University of Adelaide upon completion of sampling. Cold chain was maintained during transportation using the Bio-Freeze® phase-change system.[2]

[2] https://www.bio-bottle.com.au/

2.2.1. Modified Knott's test

MKT was performed within three hours of blood collection and storage at room temperature. A volume of 0.1 mL of EDTA-blood was mixed with 0.9 mL of 2 % buffered formalin solution. The sample was centrifuged for five minutes at 1500 rpm. One drop of 0.1 % methylene blue stain was added after discarding the supernatant. The resuspended specimen was transferred to a standard glass slide and examined entirely under a binocular light microscope using a 10x objective to screen and count microfilaria (mf) (Knott, 1939; Zajac and Conboy, 2012).

MKT result was classified as positive if at least one mf was identified, and we recorded load (mf/mL) by multiplying the mf count by a factor of ten. The operator performing the MKT was not purposively blinded to the history and clinical examination of the subject but was blinded to any LFI results.

2.2.2. LFI testing

Four commercially available LFIs were investigated including Anigen Rapid Canine HW Ag Test® (Bionote Co., South Korea), SNAP® Heartworm RT Test (IDEXX Laboratories, USA), trūRapid Heartworm (Heska, France) and WITNESS Dirofilaria® (Zoetis Inc., USA) test kits, referred to as Anigen Rapid, SNAP, truRapid and WITNESS for the remainder of this report.

One run of the Anigen Rapid test was performed on EDTA blood (denoted Anigen Rapid$_{blood}$) at a similar time as the MKT. Subsequent runs of LFI testing were performed in June 2024 when specimens were between three and nine months old. Archived sera were allowed to defrost at room temperature, then centrifuged for five minutes at 1500 rpm. Serum samples were tested simultaneously using the Anigen Rapid, SNAP, truRapid and WITNESS test (denoted Anigen Rapid $_{serum}$, SNAP $_{serum}$, truRapid $_{serum}$ and WITNESS test $_{serum}$ respectively).

Manufacturer instructions were followed, i.e., results were read at five (truRapid), eight (SNAP) and ten (Anigen, WITNESS) minutes after sample application. Tests were interpreted as either negative (no visible test band/spot in the viewing window) or positive test (any visible test band/spot development, regardless of intensity). Tests where the control band/spot failed to develop fully were considered faulty, and therefore discarded and rerun, regardless of any development of the test band/spot.

The operator performing the Anigen Rapid test on fresh blood was blinded to results of the MKT, but not to the history or clinical examination findings of the subject (as clinical suspicion triggered testing). LFIs on sera were run and interpreted by an operator blinded to previous test results, but were aware of the eligibility criteria for subject inclusion (not blinded to clinical suspicions).

During LFI procurement, we experienced inconsistencies with test packaging. The truRapid LFIs from Heska were externally labelled to refrigerate immediately, despite internal labels indicating a storage temperature of 25–30℃. Of the two batches of WITNESS LFIs delivered from Zoetis Inc., one had external instructions to refrigerate and one not, although the labelled storage temperature range is 2–25℃. Representatives from each company indicated that test performance would not be impacted by these storage conditions.

2.3. Latent class modelling

Only subjects for which we had results for all six tests across both modalities (MKT, Anigen Rapid$_{blood}$, Anigen Rapid $_{serum}$, SNAP $_{serum}$, truRapid $_{serum}$ and WITNESS test $_{serum}$) were included in the analyses. We constructed our latent class models (LCMs) within a Bayesian framework to enable customisation of the model structure. We constructed a schematic of the latency structure comparing the relationship of possible latent classes/s to our diagnostic target and the analytical targets measured by the modalities under evaluation (Fig. 1). For each model, we used Fig. 1 to define its latent class as the 'highest common

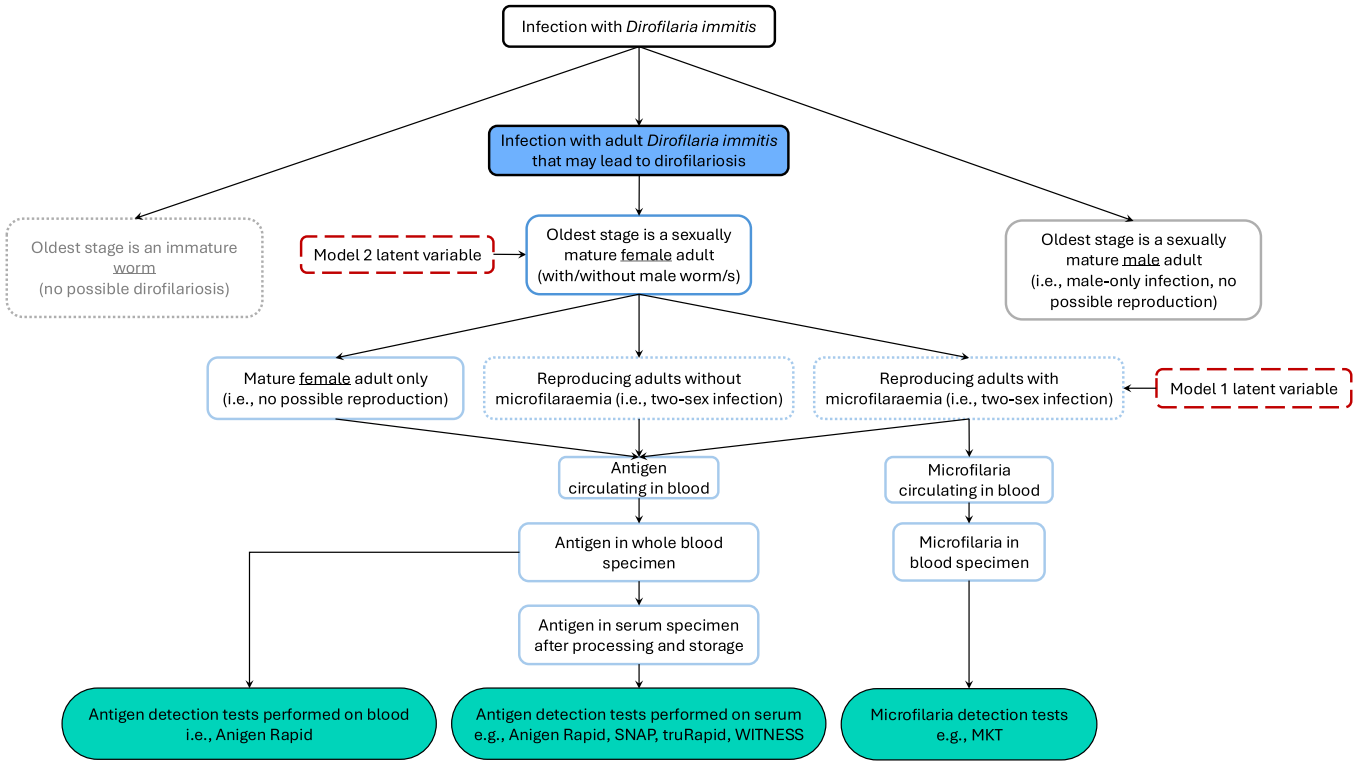


Fig. 1. Directed acyclic graph used to describe the latency structure in the Bayesian latent class model of canine heartworm diagnostic test evaluation. The target condition is presented in the light blue filled box, tests included in the green boxes, and implied latent variable in Models 1 and 2 are indicated by red dashed boxes. Infection stages relevant to this investigation of point-of-care test accuracy are bordered in blue, potentially infectious stages of infection are bordered with dashed lines and non-identifiable stages are bordered in grey. MKT; modified Knott's test.

ancestor' of any tests included (Denwood et al., 2024).

The models were executed using JAGS through R version 4.4.1 (R Core Team, 2025) using functionality from the contributed *runjags*, *mcmcplots*, *rjags* and *R2jags* packages (Denwood, 2016; Curtis, 2018; Plummer, 2023; Su and Yajima, 2024). Our R code and likelihood functions used for these models are available at https://doi.org/10.25909/28605821.

2.3.1. Model 1 – Three-test two-population

The first LCM (Model 1) compared MKT, Anigen Rapid$_{blood}$ and Anigen Rapid$_{serum}$, which allowed us to compare the two modalities (microfilarial detection and LFIs) to detect *D. immitis*. The latent variable implied in this model was 'Reproducing adults (i.e., two-sex infection with microfilaraemia)' infections (Fig. 1). Notably, this latent variable is different to the most relevant target condition i.e., 'infection with adult *D. immitis* that may lead to dirofilariosis' as neither modality can detect male-only infections. In this model, the LFI runs were considered conditionally dependent to each other due to common intermediate branches between them and the latent class (Gardner et al., 2000; Denwood et al., 2024), and covariance terms were included for both positive and negative test results. As the diagnostic target between the LFIs (*D. immitis* uterine antigen) and the MKT (microfilaria) is different, these were considered conditionally independent (Weil, 1987; Georgiadis et al., 2003; Bowman and Mannella, 2011), also supported in the absence of common branches between the latent variable and the tests in Fig. 1 (Denwood et al., 2024). A two population model with three tests under evaluation provided fourteen degrees of freedom ($df=(2^k-1)*p$, where k represents the number of tests and p the number of populations) to estimate ten parameters (two prevalences, three pairs of diagnostic sensitivity (DSe) and specificity (DSp), and two between-Anigen Rapid-runs covariance terms, one for infected and one for non-infected dogs), making the model identifiable and requirements for informative priors not essential (Cheung et al., 2021), i.e., uniform beta distributions (*beta (1,1)*) were used for all prior probabilities. Conditional dependence was modelled using covariance terms as outlined by Dendukuri and Joseph (2001). The recruited subjects were split into pairs of sub-populations based on clinical presentation, sex and location, with further details outlined below (Section 2.3.5).

2.3.2. Model 2 – Four-test one-population

Model 2 compared the four LFIs performed on archived sera, allowing us to rank their comparative performance. The latent class for this model was 'oldest stage is a sexually mature female adult' (Fig. 1). As the precise nature of the target antigen and capture antibodies used by each LFI are not disclosed by the companies that commercialise them, it was not possible to further refine Fig. 1 with intermediate branches between any of the three types of *D. immitis* infections within the latent variable and the target antigen, capture antibodies or other antigen detection processes specific to an individual LFI brand. Therefore, there was no intermediate branches possible to include in Fig. 1 on which to model conditional dependence, so we firstly modelled the LFIs as though they were conditionally independent on the latent class (Gardner et al., 2000; Denwood et al., 2024). However, to appraise the possibility of conditional dependence between these LFIs, we produced six subsequent models, each investigating the pairwise conditional dependence (in both positive and negative results) for each pair of LFIs (e.g., SNAP versus truRapid). We noted which posterior estimates of covariance (for either the positive or negative results, or both) did not overlap with zero, and considered these pair/s of LFIs as being conditionally dependent on an unknown intermediate branch between the latent class and the LFI result. Finally, we produced a model that incorporated any conditionally dependent LFI results, to simultaneously model all conditional dependence by adapting the model code provided by Nérette et al. (2008) if multiple conditional dependence terms were required. For all models (conditionally independent, pairwise conditional dependence and multiple conditional dependence), we monitored the deviance information criterion (DIC) (Spiegelhalter et al., 2002) to support the selection of the most appropriate model.

For reporting, we selected the model with the lowest DIC, but when the absolute difference in DIC between competing models was less than two, we preferred the simpler model (i.e., with minimal conditional dependence). If the most appropriate model (as suggested by the lowest DIC) was a complex model, we appraised any differences to model outputs between this and a simple model. In the absence of apparent differences, we preferenced reporting the simplest model (despite perhaps a higher DIC).

Considering four conditionally independent tests under evaluation in one population, a total of 15 degrees of freedom were available to identify nine parameters (one prevalence and four pairs of DSe and DSp). For each pairwise LFI comparison, a total of 11 parameters were modelled (one positive and one negative covariance term). As these models remained identifiable, uniform prior probabilities were used for all parameters. For the final model, the count of parameters depended on any apparent conditional dependence between LFI results. If there were less than seven covariance terms, the models would remain identifiable without informative priors. If the count of covariance terms exceeded this, then we would consider informative priors on the SNAP LFI parameters, sourced from the meta-analysis of its accuracy (Atkinson et al., 2023).

2.3.3. Model outputs

For each Markov chain Monte-Carlo (MCMC) iteration, the model provided a single set of values for each parameter – population prevalence/s, DSe and DSp, and pairs of covariance terms (when included) for each test under evaluation – that was supported by the data and model structure. Posterior parameter estimates were reported as the median and 95 % posterior credibility intervals (95 % PCI), generated from 160,000 effective iterations after a burn-in of 40,000 iterations. To assess and compare the tests' fitness to rule-in infection, we calculated the likelihood ratio of a positive result (LR^+) distribution by using the corresponding test's DSe and DSp at each iteration ($LR^+ = \frac{DSe}{1-DSp}$ (Dujardin et al., 1994)). To compare the LR^+ estimates between test pairs, we coded a step function in the models to calculate the proportion of iterations where one test's estimate was higher than another (i.e., the Bayesian p-value (Meng, 1994)). LR^+ of tests were considered consistently different if one was higher than the other in at least 95 % of iterations.

2.3.4. MCMC convergence assessment

Convergence and autocorrelation were assessed by visual inspection of three MCMC chains using distinct starting values, and the autocorrelation and trace diagnostic plots. Autocorrelation was diagnosed when there was at least 20 % correlation remaining at the end of MCMC sampling, or when the effective sample size used to generate the 95 % PCI was less than 10,000 values. This was managed by thinning to every fifth sampled value, and multiplying the total iteration count by five. We increased the total iteration count to 500,000 and extended burn-in up to 200,000 iterations based on visual inspection of the running mean plot if necessary to achieve a stable mean.

2.3.5. Sensitivity analysis

We performed sensitivity analysis of Model 1 (three-test two-population) by exploring the impact of the study population splitting on the posterior estimates, using sub-populations based either on i) clinical presentation (clinically abnormal vs apparently healthy), ii) subject sex (male vs female) and iii) geographical location (subjects from a specific town, Lautoka, vs elsewhere). We also explored potential discrepancies in the infection spectrum across populations by comparing microfilarial load distributions between the subpopulations, with a Mann-Whitney *U* test interpreted at the 5 % level of significance.

We also performed sensitivity analysis between conditional

dependence modelling of Model 2 as previously described.

2.4. Test fitness for purpose evaluation and extension

We considered the strength of evidence provided by each positive test result using the following LR^+ cutoffs – strong if ≥ 10, moderate if ≥ 5 but < 10, weak if ≥ 2 but < 5 and negligible if ≤ 2 (Hayden and Brown, 1999), and the larger the value, the stronger the evidence provided by a positive result. If the 95 % PCIs overlapped one (or multiple) cutoffs, we reported the range of strengths e.g., moderate to strong evidence.

We also used the LR^+ estimates from Model 1 to calculate the post-test probabilities of the target condition across the full range of pre-test probabilities (0–100 %) for a positive test result (Caraguel and Colling, 2021). Pre-test probabilities were first converted to pre-test odds ($Odds_{pre-test} = \frac{Probability_{pre-test}}{1-Probability_{pre-test}}$), adjusted to post-test odds based on the LR of a given test result ($Odds_{post-test} = Odds_{pre-test} \times LR_{test\ result}$) and finally converted back to a post-test probability ($Probability_{post-test} = \frac{Odds_{post-test}}{1+Odds_{post-test}}$). These calculations were performed on the median and 95 % PCI bounds of each LR to provide the median and 95 % PCI of the post-test probability of *D. immitis* infection. We then used both the probability modifying plot and the relative costs of each test to compare their clinical application/s.

3. Results

The minimum sample size of 100 dogs testing positive to either diagnostic modality was achieved by March 2024. In total, we recruited 333 subjects into the study that fit the eligibility criteria of suspect cases. Two subjects had incomplete data and were excluded from the evaluation leaving 331 analysed dogs (115 with either modality, or both, positive). Table 1 shows the counts of subjects in each set of test result combinations, subject demographics are outlined in Supplementary Materials Table S1, and the full raw dataset is provided at https://doi.org/10.25909/28605821.

Table 1
Combinations of positive (+) and negative (–) test results, and subject count found for each combination, in the diagnostic test evaluation of tests against *Dirofilaria immitis* infection, including modified Knott's test (MKT), one canine heartworm lateral flow immunoassay (LFI) performed on fresh blood and four LFIs performed on serum after processing, transport and archive. Test result combinations not shown had zero subjects.

Whole blood		*Stored and transported sera*				
MKT	*Anigen Rapid*[a]	*Anigen Rapid*[a]	*SNAP*[b]	*truRapid*[c]	*WITNESS*[d]	*Total*
+	+	+	+	+	+	50
+	+	+	+	+	-	0
+	+	+	+	–	+	2
+	+	+	–	–	–	2
+	–	+	+	+	+	1
+	–	–	–	–	–	11
–	+	+	+	+	+	38
–	+	+	+	–	–	1
–	+	+	–	+	+	2
–	+	+	–	–	+	2
–	+	–	–	–	–	6
–	–	+	+	+	+	2
–	–	–	+	+	+	1
–	–	–	–	+	+	1
–	–	–	–	–	+	4
–	–	–	–	–	–	208
All other test result combinations						0

[a] Anigen Rapid® (Bionote Co.)
[b] SNAP® (IDEXX Laboratories)
[c] trūRapid Heartworm (Heska)
[d] WITNESS® (Zoetis Inc.)

3.1. Bayesian latent class modelling

For all models run, the three MCMC chains converged similarly and satisfactorily with no effect of the starting value.

3.1.1. Model 1 – Three-test two-population

The sub-populations of apparently healthy and clinical dogs had a significantly different spectra of microfilarial loads ($p = 0.004$), potentially violating the assumption of constant DSe across populations (Table 2). There was no strong evidence the microfilarial load distribution differed significantly between dog sex (Run ii)), or dog location (Run iii)). Ultimately, the subdivision of the subjects had minimal impact of the model estimates (Table 2 and Fig. 2). MKT and LFI had similar performances to rule-in 'reproducing adults with microfilaraemia', providing moderate to strong evidence when returning a positive result (Fig. 2), and were not significantly different from each other (Table 2). There was minimal covariance between either test result of $Anigen\ Rapid_{blood}$ and $Anigen\ Rapid_{serum}$.

3.1.2. Model 2 – Four-test one-population

There was evidence of conditional dependence between the positive results of $truRapid_{serum}$ and $WITNESS_{serum}$, and $SNAP_{serum}$ and $truRapid_{serum}$, as well as the negative results of $Anigen\ Rapid_{serum}$ and $WITNESS_{serum}$ (Table 3). The model incorporating conditional dependence between these sets of test results had an absolute DIC of at least two less than any other model (Supplementary Materials Table S2). Therefore, we presented the results of LFI performance using a LCM incorporating covariance terms between these pairs of LFI results (Table 4), which provided different results to the simplest model assuming independence between all LFI results (Supplementary Materials Table S3).

All LFIs provided strong evidence to rule-in the presence of 'Oldest stage is a sexually mature female adult' infections (Table 4). However, the $WITNESS_{serum}$ LFI consistently provided weaker evidence to rule-in the latent class relative to the other LFIs tests (Bayesian p-value = 0.02, 0.002 and 0.004 compared to Anigen Rapid, SNAP and truRapid respectively; Table 4). Anigen Rapid provided a marginally lower LR^+ to SNAP (Bayesian p-value = 0.95). The prevalence of the modelled latent class was 29.7 % (95 % PCI: 24.9–34.9 %).

3.2. Test fitness for purpose

The probability modifying plots revealed non-distinguishable post-test probabilities for positive MKT, $Anigen\ Rapid_{blood}$ and $Anigen\ Rapid_{serum}$ (Fig. 3). For instance, for a pre-test probability of 30 %, the median post-test probability of infection was 85.2 % (95 %PCI: 74.0–96.4 %), 89.7 % (95 %PCI: 70.9–97.4 %) and 92.8 % (95 %PCI: 71.1–99.3 %) after a positive result from MKT, $Anigen\ Rapid_{blood}$ and $Anigen\ Rapid_{serum}$, respectively (Fig. 3).

4. Discussion

We evaluated two diagnostic modalities to rule-in *Dirofilaria immitis* infection in a population of dogs that would raise suspicion of infection, due to suggestive clinical history or signs, using latent class modelling. Our analysis included the modified Knott's test (MKT) and the four canine heartworm lateral flow immunoassay (LFI) tests Anigen Rapid Canine HW Ag Test® (Bionote Co.), SNAP® Heartworm RT Test (IDEXX Laboratories), trūRapid Heartworm (Heska) and WITNESS Dirofilaria® (Zoetis Inc.), referred to as Anigen Rapid, SNAP, truRapid and WITNESS respectively. We wish to here acknowledge the absence of the AbboScreen (Abbott Laboratories), Accuplex® (Antech Diagnostics) and ALERE Dirofilariose (Bionote Co.) LFIs in our evaluation, as they are not presently commercially available in Australia. Although the Speed Diro™ (Virbac) and VETSCAN® (Zoetis Inc) LFIs are available, logistical limitations during test procurement prevented their inclusion. Finally,

Table 2
Likelihood ratio of a positive result (LR^+), diagnostic sensitivity (DSe) and specificity (DSp) for the modified Knott's test (MKT) and Anigen Rapid Canine HW Ag Test® (Bionote Co.) conducted on either fresh blood or archived serum, as estimated by Bayesian latent class modelling in two populations for the latent class 'Reproducing adults with microfilaraemia'. Values are reported as medians with 95 % posterior credibility intervals (PCI) in parentheses. Each run divided the study population as indicated, and the covariance between positive and negative Anigen Rapid tests in each run is shown. The prevalence of the latent class is reported as the median (95 % PCI). Microfilarial load per mL of blood (mf/mL) is reported as the median and range (restricted to microfilaraemic subjects), and the p-value of a Mann-Whitney *U* test comparing the mf/mL between each group is shown.

		Run i)		*Run ii)*		*Run iii)*	
% Prevalence (95 % PCI)		*Clinically abnormal* 46.1 (36.8, 55.8)	*Apparently healthy* 19.6 (13.3, 26.7)	*Female* 26.3 (18.3, 35.1)	*Male* 36.2 (26.9, 46.0)	*Lautoka* 30.5 (21.3, 40.0)	*Elsewhere* 31.2 (22.6, 40.7)
mf/mL median (range)		9365 (10 – 280,080)	639 (10 – 8320)	40 (10 – 7220)	435 (10 – 280,080)	130 (10 – 8960)	70 (10 – 280,080)
p-value		0.004		0.27		0.39	
Test							
MKT	LR^+	11.7 (6.3, 35.1)		13.4 (6.7, 114.6)		13.7 (6.7, 141.7)	
	DSe	54.4 (44.3, 65.0)		55.0 (44.7, 66.8)		55.2 (44.9, 67.6)	
	DSp	95.3 (91.7, 98.4)		95.9 (92.0, 99.5)		95.9 (92.1, 99.6)	
Anigen Rapid$_{blood}$	LR^+	23.1 (6.7, 171.7)		20.2 (5.4, 138.2)		18.9 (5.2, 130.2)	
	DSe	95.7 (84.0, 99.6)		93.5 (78.7, 99.4)		93.5 (78.2, 99.4)	
	DSp	95.9 (85.7, 99.5)		95.4 (83.3, 99.3)		95.1 (82.6, 99.3)	
Anigen Rapid$_{serum}$	LR^+	33.0 (6.6, 700.8)		30.0 (5.4, 658.5)		26.4 (5.2, 502.5)	
	DSe	92.1 (80.5, 97.8)		91.5 (76.8, 98.8)		91.4 (76.2, 98.8)	
	DSp	97.2 (86.0, 99.9)		97.0 (83.5, 99.9)		96.6 (82.8, 99.8)	
Covariance between Anigen Rapid runs	*Positive result*	0.02 (0.00, 0.11)		0.03 (0.00, 0.14)		0.03 (0.00, 0.14)	
	Negative result	0.02 (0.00, 0.11)		0.02 (0.00, 0.13)		0.02 (0.00, 0.14)	

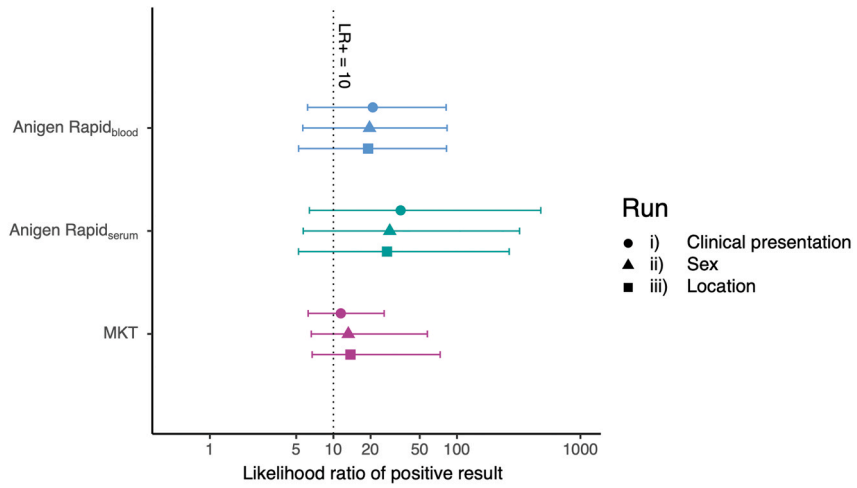


Fig. 2. Likelihood ratios for positive result (LR^+) of the modified Knott's test (MKT) and Anigen Rapid test performed on blood or archived serum to detect *Dirofilaria immitis* infection. Estimates were produced from a three-test two-population Bayesian latent class model incorporating conditionally dependence between the two Anigen Rapid tests. Points represent medians, and whiskers indicate 95 % posterior credibility intervals of each parameter in each model, plotted on a logarithmic (base 10) scale. Each run represents a distinct population stratification: Run i) By clinical presentation, Run ii) By sex and Run iii) By location. The vertical dotted line represents $LR^+ = 10$, above which a positive test result is considered to give strong evidence to rule-in infection (Caraguel and Colling, 2021).

the previously available Solo Step® CH (Heska) LFI has been replaced by the trūRapid Heartworm LFI.

The LFIs (proxied by Anigen Rapid) and MKT both provided approximately equivalent, moderate to strong, evidence of *D. immitis* infection in a positive result to rule-in a microfilaraemic infection. Between the LFIs, SNAP and truRapid had similar performance, followed by Anigen Rapid, with WITNESS showing consistently weaker performance than all other LFIs. For clinicians selecting an LFI to rule-in *D. immitis* infection, we can therefore recommend Anigen Rapid, SNAP or truRapid in preference to WITNESS.

As the performance of LFIs and MKT was not significantly different (Model 1), the veterinarian's modality selection may consider the relative costs of each, including the modality's price, welfare, safety and accessibility implications. The welfare and safety of LFIs and MKT are deemed equal, as both require a blood specimen, making their relative cost difference associated to their purchase price and accessibility. Once the necessary equipment for MKT (microscope and centrifuge) is in place, the running cost is primarily operator time for preparation and reading. In contrast, LFIs require no (additional) equipment and reduce operator labour requirements, although each test must be purchased for a single use. Higher labour costs in high-income countries (Australian Government, 2025) may make MKT either a similar (or higher) price than LFIs, whereas in low- and middle-income countries (LMICs) such as Fiji (Fijian Government, 2015), lower labour costs may make MKT the more economical and practical option. Furthermore, the single-use nature of LFIs and their relatively short shelf-life (up to one year), make accessibility reliant on a continuous and sustained supply chain. In contrast, very few consumable or perishable resources are needed for MKT, making it an attractive option in settings with limited supply reliability or infrequent testing. Therefore, the relative cost of MKT and

Table 3
Covariance between canine heartworm lateral flow immunoassay point-of-care antigen detection tests after Bayesian latent class modelling in one population for the latent class 'Oldest stage is a sexually mature female adult' with pairwise covariance comparisons between the test labelled in the column and the test labelled in the row. Median and 95 % posterior credibility interval (in parentheses) of the covariance terms are shown, and intervals not overlapping with zero are highlighted.

	Covariance in truly infected			*Covariance in truly non-infected*		
	SNAP[c]	*truRapid*[d]	*Witness*[e]	*SNAP*[c]	*truRapid*[d]	*Witness*[e]
Anigen Rapid[b]	0.008 (0.00, 0.025)	0.003 (-0.002, 0.024)	0.003 (0.00, 0.024)	0.003 (0.00, 0.013)	0.001 (0.00, 0.010)	0.007 (0.001, 0.023)[a]
SNAP[c]		0.013 (0.008, 0.032)[a]	0.002 (0.00, 0.016)		0.001 (0.00, 0.007)	0.001 (0.00, 0.007)
truRapid[d]			0.009 (0.001, 0.037)[a]			0.004 (0.00, 0.017)

[a] posterior credibility interval not overlapping with zero
[b] Anigen Rapid® (Bionote Co.)
[c] SNAP® (IDEXX Laboratories)
[d] trüRapid Heartworm (Heska)
[e] WITNESS® (Zoetis Inc.)

Table 4
Likelihood ratio of a positive result (LR^+), diagnostic sensitivity (DSe) and specificity (DSp) for canine heartworm lateral flow immunoassay point-of-care antigen detection tests performed on archived serum, after Bayesian latent class modelling in one population for the latent class 'Oldest stage is a sexually mature female adult'. Values are reported as their median and 95 % posterior credibility intervals (PCI) are shown in parentheses. Proportion of model iterations in which the LR^+ of the LFI labelled in the columns was higher than the LR^+ of the LFI labelled in the rows (i.e., Bayesian p-value) is shown. Proportions less than 5 % are highlighted.

LFI	*DSe (95 % PCI)*	*DSp (95 % PCI)*	*LR^+ (95 % PCI)*	*Bayesian p-value*		
				SNAP[b]	*truRapid*[c]	*Witness*[d]
Anigen Rapid[a]	97.6 (93.1, 99.6)	98.1 (94.6, 99.5)	51.7 (17.9, 184.4)	0.95	0.92	0.02*
SNAP[b]	95.4 (85.1, 98.9)	99.7 (98.4, 100.0)	313.5 (59.2, 8436.8)		0.43	0.002*
truRapid[c]	95.4 (84.7, 98.5)	99.6 (98.1, 100.0)	232.8 (48.8, 6193.8)			0.004*
Witness[d]	98.3 (94.4, 99.8)	96.2 (92.5, 98.3)	25.9 (13.1, 57.2)			

* $p < 0.05$
[a] Anigen Rapid® (Bionote Co.)
[b] SNAP® (IDEXX Laboratories)
[c] trüRapid Heartworm (Heska)
[d] WITNESS® (Zoetis Inc.)

LFIs are variable by geographical and economic situations, and given their similar performance, veterinarians may proceed to select the least costly modality for a given situation and can expect a similar rule-in performance from either.

Veterinarians in low-income settings may have limited testing resources, and they would probably preference the detection and management of infected individuals that may already have disease. This makes the outputs of Model 1 most applicable, and our findings suggest the most appropriate modality can be made by an assessment of their relative costs, and we wish to reassure clinicians that microfilarial detection is as suitable as LFIs. In contrast, a veterinarian in a high-income setting may be most interested in detecting *D. immitis* infection after a suspected (or reported) lapse to preventative medication during an infection-risk period. In this setting, investigation resources may have fewer limitations, and the testing goal may therefore be to detect an infected individual before disease is apparent. The outputs of Model 2 are therefore most applicable, and can be used to select the most brands of LFI with the strongest performance.

To apply diagnostic test results to an individual suspected to be infected, a veterinarian should consider the minimum probability of infection at which the testing effort would be stopped and intervention initiated i.e., the intervention threshold (Pauker and Kassirer, 1980). There is no guidance on the appropriateness of any intervention threshold/s relevant to *D. immitis* infections. Consequentially, the various guidelines outlining the use and interpretation of diagnostic tests for *D. immitis* (ESDA, 2017; Korman et al., 2017; CPEP, 2019; TroCCAP, 2019; CAPC, 2020; ESSCAP, 2022; AHS, 2024) provide vague recommendations about the selection or order of modalities to rule-in infection to direct intervention. In addition, none of the guidelines consider the relative cost of each modality in their recommendations. Whilst our investigation provides a framework to inform some improvements to these guidelines, we suggest robust analysis of different intervention thresholds would be required to then facilitate consequent guideline review, and may also include formal financial analysis.

The latent variable implied in Model 1 ('Reproducing adults with microfilaraemia') has a different clinical relevance to Model 2 ('Oldest stage is a sexually mature female worm'). Whereas Model 1 was useful to identify the accuracy of these modalities to detect microfilaraemic individuals, Model 2 allowed us to compare the relative performance of each included LFI, as well as being more closely related to infections likely to lead to heartworm disease (closer to the target condition in Fig. 1). These different latency structures probably are responsible for the apparently different performance of Anigen $Rapid_{serum}$ between Models 1 and 2.

Notably, the applicability of Model 2 to the target condition was limited due to our inability to include a test targeting the 'Oldest stage is a sexually mature male adult i.e., male-only infection' phase (Fig. 1). However, as the probability of a single-sex infection reduces with the count of reinfections and therefore with infection pressure (McCrea et al., 2021), and given the reasonably high prevalence of the latent class in both models (and therefore the likely high *D. immitis* infection prevalence in the population), the impact of this mis-detection would be (at most) negligible. Furthermore, as single-sex infections are less likely to cause clinical disease in dogs and are not infectious (Knight, 1987; Polizopoulou et al., 2000; Ames and Atkins, 2020), they have a lower clinical and epidemiological relevance. The inclusion of this rare stage in Fig. 1 was for completeness only. Additionally, no included modality could detect immature infections, hence our selection considerations only targeting detectable infections that were acquired at least six months prior to testing. Investigating dogs that may become infectious is a distinct and different testing context to what we presented, fitting

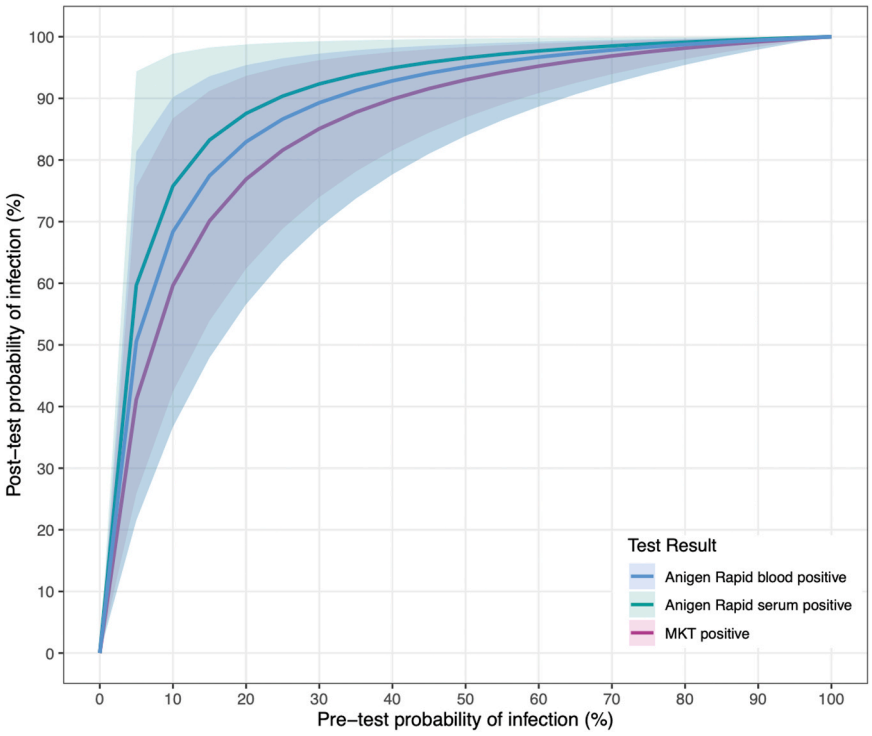


Fig. 3. Probability modifying plots for a positive test result provided by modified Knott's test (MKT; purple) and Anigen Rapid canine heartworm lateral-flow-immunoassay performed on fresh blood or archived serum (blue and green respectively). Median (line) and 95 % posterior credibility interval (ribbon) post-test probability of *Dirofilaria immitis* infection for given pre-test probabilities (0–100 %) after a positive test result.

more with population level control and prevention to notify stakeholders about possible future occurrence of disease i.e., ruling-out infection. Investigation of the rule-out performance of these modalities could become a diagnostic test evaluation of interest if the management of *D. immitis* by authorities becomes necessary or desirable.

We wish to briefly discuss the prevalence of *D. immitis* infection reported in our models. The source population was subjects raising clinical suspicion in veterinary practice, and therefore the reported prevalence of infection is likely an overestimate of the overall Fijian dog population. However, our subjects represented individuals in the population that were possible to be infected (based on their age, interrupted/absent preventative medication history or presence of clinical signs), and the prevalence in this group reflects the most clinically relevant prevalence in this region (or comparable regions).

As we included conditional dependence between Anigen $Rapid_{blood}$ and $_{serum}$ in Model 1, any disagreement between MKT and (either) LFI result were equally distributed. We would anticipate the sensitivity of MKT to be increased if specimens were purposively collected during peak microfilaraemia (i.e., during mosquito biting times) (Evans et al., 2017), thereby likely improving the agreement between MKT and the LFIs. We expect this to have also translated to an increased performance of the LFIs, as ultimately there would be less disagreement between modalities. In the absence of a perfect reference standard, the latent class modelling was unable to differentiate instances when modalities disagreed. However, this purposive subject selection would not reflect the use of these tests by veterinarians in practice and would reduce the generalisability of our results, and regardless also fits more in the testing context of screening an apparently healthy individual for infection (rule-out context).

We wish to acknowledge the observed breach of the Hui and Walter (1980) requirement of constant test performance between populations that occurred in Model 1 (Run i)) by separating the sampled population by clinical presentation. This was explained in the substantial difference to microfilarial load between the two sub-populations. However, alternative splitting of the study population did not impact final estimates, and we ultimately reported test performance using Run ii), i.e., study population split by dog sex. The width of the credibility intervals of our results from Model 1 precluded the ranking of one test modality as clearly superior to others. We do not believe increasing the sample size would have addressed this (median values similar) and consider the interval width to result from genuine variability to test performance across the spectrum of the target condition in the population.

CRediT authorship contribution statement

P.J. Atkinson: Conceptualisation, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualisation, Writing – original draft, Writing – review & editing. **C. Quimby:** Project administration, Data curation, Writing – review & editing. **A. Datt:** Data curation, Writing – review & editing. **T.D. Nielsen:** Investigation, Methodology, Visualisation, Writing – original draft, Writing – review & editing. **C.G.B. Caraguel:** Conceptualisation, Formal analysis, Investigation, Methodology, Project administration, Visualisation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

Bionote Co. provided their tests for initial fresh blood testing in-kind. Bionote Co., IDEXX Laboratories and Heska provided their tests for serum testing at cost-price, discounted from the recommended wholesale price. Zoetis Inc. were not able to offer discounted tests, and their tests were purchased at the wholesale price.

Acknowledgements

We would like to acknowledge The University of Adelaide's School of Animal and Veterinary Sciences for providing the internal funding to support this project. We would like to thank Andrew Falkiner and Jacinta Finlayson for their assistance in the laboratory phase of this work. Finally, we would like to thank the two reviewers who ultimately

provided suggestions to improve this work.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.prevetmed.2026.106783.

References

AHS, 2024. Prevention, diagnosis, and management of heartworm (Dirofilaria immitis) infection in dogs. ⟨https://www.heartwormsociety.org/images/pdf/2018-AHS-Canine-Guidelines.pdf⟩ (accessed 16 April 2025).

Ames, M.K., Atkins, C.E., 2020. Treatment of dogs with severe heartworm disease. Vet. Parasitol. 283, 109131.

Atkinson, P.J., O'Handley, R., Nielsen, T., Caraguel, C.G.B., 2023. Relative diagnostic accuracy of point-of-care tests to rule-in *Dirofilaria immitis* infection in clinically suspect dogs: A systematic review and meta-analysis. Prev. Vet. Med. 217, 105970.

Australian Government, 2025. Minimum Wages. ⟨https://www.fairwork.gov.au/pay-and-wages/minimum-wages#:~:text=From%201%20July%202024%2C%20the,or%20after%201%20July%202024⟩. (accessed 13th March 2025).

Bowman, D.D., Atkins, C.E., 2009. Heartworm biology, treatment, and control. Vet. Clin. North Am. Small Anim. Pr. 39, 1127–1158.

Bowman, D.D., Mannella, C., 2011. Macrocyclic lactones and *Dirofilaria immitis* microfilariae. Top. Companion Anim. Med. 26, 160–172.

CAPC, 2020. Heartworm. ⟨https://capcvet.org/guidelines/heartworm/⟩ (Accessed 27 January 2022).

Caraguel, C.G.B., Colling, A., 2021. Diagnostic likelihood ratio - the next-generation of diagnostic test accuracy measurement. Rev. Sci. Tech. 40, 299–309.

Cheung, A., Dufour, S., Jones, G., Kostoulas, P., Stevenson, M.A., Singanallur, N.B., Firestone, S.M., 2021. Bayesian latent class analysis when the reference test is imperfect. Rev. Sci. Tech. 40, 271–286.

Constantinoiu, C., Croton, C., Paterson, M.B.A., Knott, L., Henning, J., Mallyon, J., Coleman, G.T., 2023. Prevalence of canine heartworm infection in Queensland, Australia: comparison of diagnostic methods and investigation of factors associated with reduction in antigen detection. Parasit. Vectors 16.

Courtney, C.H., Zeng, Q., 1993. Comparison of two antigen tests and the modified Knott test for detection of canine heartworm at different worm burdens. Canine Pr. 18, 5–7.

CPEP, 2019. Guidelines for the managment of parasites in dogs and cats. ⟨https://research-groups.usask.ca/cpep/index.php⟩ (Accessed 24th February 2025).

Curtis, S., 2018. mcmcplots: Creat Plots from MCMC Output.

Dendukuri, N., Joseph, L., 2001. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics 57, 158–167.

Denwood, M.J., 2016. runjags: An R Package Providing Interface Utilities, Model templates, parallel computing methods and additional distributions for MCMC models in JAGS. J. Stat. Softw. 71, 1–25.

Denwood, M., Nielsen, S.S., Olsen, A., Jones, H.E., Coffeng, L.E., Landfried, G., Nielsen, M.K., Levecke, B., Thamsborg, S.M., Eusebi, P., Meletis, E., Kostoulas, P., Hartnack, S., Erkosar, B., Toft, N., 2024. All that glitters is not gold: an interpretive framework for diagnostic test evaluation using *Ascaris lumbricoides* as a conceptual example. PLoS Negl. Trop. Dis. 18, e0012481.

Dujardin, B., Van den Ende, J., Van Gompel, A., Unger, J.P., Van der Stuyft, P., 1994. Likelihood ratios: a real improvement for clinical decision making? Eur. J. Epidemiol. 10, 29–36.

ESDA, 2017. Guidelines for clinical management of canine heartworm disease. Eur. Soc. Dirofilariosis Angiostrongylosis.

ESSCAP, 2022. Parasitological Diagnosis in Cats, Dogs and Equines. ⟨https://www.esccap.org/uploads/docs/ebjqliec_1335_ESCCAP_GL4_English_v4_1p.pdf⟩ (Accessed 13th February 2025).

Evans, C.C., Burkman, E.J., Dzimianski, M.T., Moorhead, A.R., Savadelis, M.D., Angenendt, C., Zymny, S., Kulke, D., 2017. Periodicity of *Dirofilaria immitis* in longterm infections. Parasitol. Res 116, 75–80.

Fijian Government, 2015. Fijian Government Increases National Minimum Wage. ⟨https://www.fiji.gov.fj/Media-Centre/News/FIJIAN-GOVERNMENT-INCREASES-NATIONAL-MINIMUM-WAGE?feed=news⟩ (Accessed 13th March 2025).

Gardner, I.A., Stryhn, H., Lind, P., Collins, M.T., 2000. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. Prev. Vet. Med 45, 107–122.

Georgiadis, M.P., Johnson, W.O., Gardner, I.A., Singh, R., 2003. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. J. R. Stat. Soc. Ser. C. Appl. Stat. 52, 63–76.

Goodwin, J.-K., 1998. The serologic diagnosis of heartworm infection in dogs and cats. Clin. Tech. Small Anim. Pr. 13, 83–87.

Hayden, S.R., Brown, M.D., 1999. Likelihood ratio: A powerful tool for incorporating the results of a diagnostic test into clinical decisionmaking. Ann. Emerg. Med. 33, 575–580.

Hui, S.L., Walter, S.D., 1980. Estimating the error rates of diagnostic tests. Biometrics 36, 167–171.

Knight, D.H., 1987. Heartworm infection. Vet. Clin. North Am. Small Anim. Pr. 17, 1463–1518.

Knott, J., 1939. A method for making microfilarial surveys on day blood. Trans. R. Soc. Trop. Med. Hyg. 33, 191–196.

Korman, R., Labuc, R., Traub, R., Atwell, R.B., 2017. Australian Guidelines for Heartworm Diagnosis in Dogs. Vets Australia Australia.

Kostoulas, P., Nielsen, S.S., Branscum, A.J., Johnson, W.O., Dendukuri, N., Dhand, N.K., Toft, N., Gardner, I.A., 2017. STARD-BLCM: standards for the reporting of diagnostic accuracy studies that use bayesian latent class models. Prev. Vet. Med. 138, 37–47.

Little, S., Saleh, M., Wohltjen, M., Nagamori, Y., 2018. Prime detection of *Dirofilaria immitis*: understanding the influence of blocked antigen on heartworm test performance. Parasit. Vectors 11.

Martini, M., Poglayen, G., Capelli, G., Roda, R., 1991. Diagnosis of canine filariosis: relative sensitivity and specificity of some haematological techniques. Angew. Parasitol. 32, 133–136.

Mataika, J.U., Dando, B.C., Spears, G.F., Macnamara, F.N., 1971. Mosquito-borne infections in Fiji. III. Filariasis in northern Fiji: epidemiological evidence regarding the mechanisms of pathogenesis. J. Hyg. 69, 297–306.

McCall, J.W., Genchi, C., Kramer, L.H., Guerrero, J., Venco, L., 2008. Heartworm disease in animals and humans. Adv. Parasitol. 66, 193–285.

McCrea, A.R., Edgerton, E.B., Oliver, G.T., O'Neill, F.M., Nolan, T.J., Lok, J.B., Povelones, M., 2021. A novel assay to isolate and quantify third-stage *Dirofilaria immitis* and *Brugia malayi* larvae emerging from individual *Aedes aegypti*. Parasit. Vectors 14, 30.

Meng, X.-L., 1994. Posterior predictive *p*-values. Ann. Stat. 22, 1142–1160, 1119.

Mylonakis, M.E., Papadopoulos, E., Koutinas, A.F., Paitaki, C., Leontides, L., 2004. Comparative methodology for the detection and differentiation of circulating microfilariae of *Dirofilaria immitis* in the dog. J. Helminthol. 78, 137–140.

Naaktgeboren, C.A., Bertens, L.C.M., Smeden, Mv, Groot, J.A.Hd, Moons, K.G.M., Reitsma, J.B., 2013. Value of composite reference standards in diagnostic research. BMJ 347, f5605.

Nérette, P., Stryhn, H., Dohoo, I., Hammell, L., 2008. Using pseudogold standards and latent-class analysis in combination to evaluate the accuracy of three diagnostic tests. Prev. Vet. Med. 85, 207–225.

Olver, J., 2022. Heartworm diseases. ⟨https://www.fijitimes.com/heartworm-diseases/⟩ (Accessed 16th June 2023).

Pauker, S.G., Kassirer, J.P., 1980. The threshold approach to clinical decision making. N. Engl. J. Med. 302, 1109–1117.

Plummer, M., 2023. rjags: Bayesian Graphical Models using MCMC.

Polizopoulou, Z.S., Koutinas, A.F., Saridomichelakis, M.N., Patsikas, M.N., Desiris, A.K., Roubies, N.A., Leontidis, L.S., 2000. Clinical and laboratory observations in 91 dogs infected with *Dirofilaria immitis* in northern Greece. Vet. Rec. 146, 466–469.

R Core Team, 2025. R: A language and environment for statistical computing. ⟨https://www.R-project.org/⟩ (Accessed 14 June 2024).

Rohrbach, B.W., Patton, S., 2013. Effects of diagnostic test accuracy and treatment efficacy on the occurrence of suspected failure of heartworm prophylaxis in dogs. J. Vet. Intern Med. 27, 791–797.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B Stat. Method. 64, 583–639.

Su, Y., Yajima, M., 2024. R2jags: Using R to Run 'JAGS'.

Symes, C.B., 1960. A note on *Dirofilaria immitis* and its vectors in Fiji. J. Helminthol. 34, 39–42.

TroCCAP, 2019. Guidelines for the diagnosis, treatment and control of canine endoparasites in the tropics. ⟨https://www.troccap.com/wp-content/uploads/2019/05/TroCCAP_Canine_Endo_Guidelines_English_Ver2.pdf⟩ (Accessed 23rd February 2025).

Weil, G.J., 1987. *Dirofilaria immitis*: identification and partial characterization of parasite antigens in the serum of infected dogs. Exp. Parasitol. 64, 244–251.

Wilks, C., 2001. Gold standards as fool's gold. Aust. Vet. J. 79, 115.

Zajac, A.M., Conboy, G.A., 2012. Detection of parasites in blood. Veterinary Clinical Parasitology. Wiley-Blackwell.