Study to determine clinical decision thresholds in small animal veterinary practice

Natalie Toulme Guevara,[1] Erik Hofmeister,[2] Mark Ebell,[3] Isabella Locatelli[4]

Abstract

This study aimed to determine clinical decision thresholds for six common conditions in small animal veterinary practice. Participants were provided with an online survey. Five questions described scenarios of canine patients with suspected panosteitis, hypothyroidism, urinary tract infection (UTI), mechanical gastrointestinal obstruction (GIO) and idiopathic epilepsy, and one question described a feline patient with suspected chronic kidney disease. A range of probabilities was applied to each scenario. Test and treatment threshold levels were computed for each scenario from 297 usable responses. The test and treatment thresholds were determined for UTI (test=12.8 per cent; 95 per centCI=1.1 to 20.7; treatment=82.0per cent; 95 per centCI=66.3 to 100) and GIO (test=3.2 per cent; 95 per cent CI=0 to 10.4; treatment=87.3 per cent; 95 per centCI=82.6 to 93.5). All other scenarios did not provide data that allowed interpretable test and treatment thresholds. This pilot study has used a new approach in determining clinical thresholds in small animal medicine. Thresholds were successfully determined for two common conditions—canine mechanical GIO and canine UTI. Future research should broaden investigation of methods to determine group clinical threshold levels among veterinarians, which may be used as the basis for clinical decision rules.

Introduction

Like all medical professionals, veterinarians must often make clinical decisions based on imperfect or incomplete information. Veterinarians must rely on clinical experience and training as well as an understanding of the likelihood of various diagnoses when it comes to deciding on when to rule a disease out, pursue diagnostic tests or treat a disease. Balancing one's expertise and experience with current scientific data and evidence-based recommendations can be incredibly challenging.[1]

The movement for evidence-based practice places greater emphasis on using current best evidence rather than expert or intuitive reasoning.[2] The emergence of evidence-based practice took hold during the early 1990s and began to enter into the veterinary profession approximately a decade later. While it is ideal that every clinician use current best evidence conjointly with clinical expertise, clinical decisions are often dictated by one's experience which can lead to biased decision making. Current barriers to evidence-based veterinary practice include lack of patient-centred research and epidemiological data, as well as veterinarians having the access, time and training to appropriately use such data.[3] However, even with these obstacles, there is growing enthusiasm within the veterinary profession to adopt evidence-based care.

In an effort to minimise bias, evidence-based medicine often employs clinical decision-making models. The threshold model of diagnosis is a theoretical framework that explains an implicit process through an explicit formula. This model was originally described by Pauker and Kassirer and has two thresholds: a 'testing threshold' and a 'test-treatment threshold'.[4] The testing threshold (hereon referred to as the test threshold) is the probability of disease above which a clinician decides to pursue testing in order to reach a diagnosis. The test-treatment threshold (hereon referred to as the treatment threshold) is the probability of disease above which a clinician is certain enough that the disease or

Veterinary Record (2019) doi: 10.1136/vr.104596

[1]Banfield Pet Hospital, Vancouver, Washington, USA
[2]Department of Surgery, Midwestern University, Glendale, Arizona, USA
[3]Department of Epidemiology and Biostatistics, University of Georgia, Athens, Georgia, USA
[4]Department of Ambulatory Care and Community Medicine, Universite de Lausanne, Lausanne, Switzerland

E-mail for correspondence: natalie.guevara@banfield.com

Provenance and peer review Not commissioned; externally peer reviewed.

Received July 26, 2017
Revised January 5, 2019
Accepted April 22, 2019

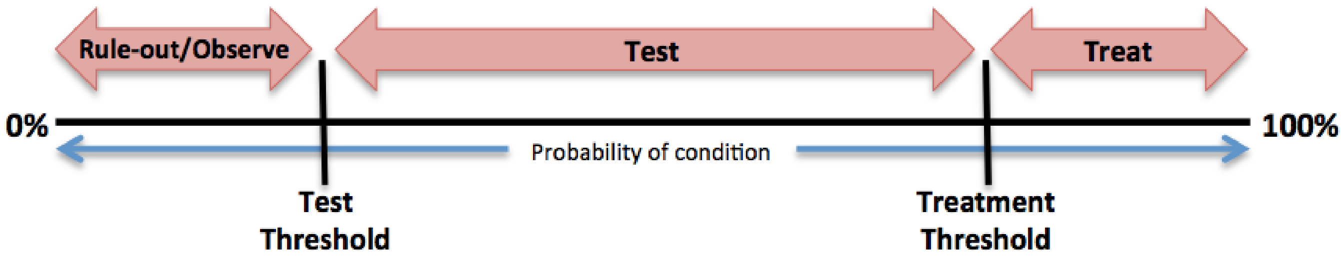

Figure 1 Illustrated depiction of the threshold model for clinical decision making.

condition in question is likely and treatment should be initiated (figure 1).

Previous studies have used several approaches to investigate the test and treatment thresholds for conditions among both individual physicians and groups of physicians.[5–9] Ebell and others recently developed a novel approach to determining physician threshold levels.[8] Primary care and emergency physicians were presented with clinical scenarios in which the probabilities of common conditions were randomised for each respondent. These scenarios were all structured in the same style providing a patient signalment, presenting complaint and options for next steps (table 1). Group threshold levels were determined using an adaptation of the logistic regression model described by Plasencia and others.[9] Recently, this same model was successfully used to determine thresholds for diagnosis of tuberculosis in HIV-positive patients.[10]

This approach can be applied to decision making in small animal veterinary medicine.[11] Establishing these thresholds may provide insight into rationale for clinical decisions made by small animal veterinarians. Clinical decision thresholds also provide the potential establishment of clinical decision rules, which are significant in human medicine yet remain to be investigated in veterinary medicine.[8] The purpose of this

Table 1 Clinical vignettes, possible probabilities and response options

Suspected condition	Clinical scenario	Range of probabilities	Decision options
Panosteitis	An eight-month-old male neutered German shepherd dog presents to your clinic for evaluation of acute forelimb lameness over the past several days. Based on the dog's history and physical exam, you believe there is an X per cent probability that he has panosteitis. Radiographs of the long bones are approximately US$100 per limb. The owner has moderate financial restrictions. You decide to...	X=(2, 5, 10, 15, 18, 20, 22, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90)	A. Recommend radiographs of affected long bone to evaluate for injury vs panosteitis. B. Recommend a non-steroidal anti-inflammatory to help with the discomfort and explain that panosteitis should resolve over time.
Hypothyroidism	A nine-year-old female spayed Golden retriever presents to your clinic for weight gain and mild lethargy over the past several months. Based on the history and physical exam, you believe there is an X per cent probability that this dog has hypothyroidism. The owner has moderate financial limitations and further blood work is approximately US$140. You decide to...	X=(2, 5, 10, 15, 18, 20, 25, 30, 35, 40, 45, 48, 50, 55, 58, 60, 65, 68, 70, 72, 75)	A. Recommend a prescription for a metabolic (weight maintenance) diet to aid in weight loss and have the dog follow up in 6 months. B. Recommend a systemic workup including complete blood count (CBC), chemistry and total T4.
Urinary tract infection (UTI)	A seven-year-old female spayed Schaunzer presents to your clinic for a first time episode of stranguria, polyuria and hematuria. Based on the history, physical exam and in-house lab work (CBC, chemistry and urinalysis), you feel there is an X per cent probability that this dog has a lower UTI. The owner has moderate financial restrictions. Urine culture is approximately US$50. You decide to...	X=(2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98)	A. Recommend pursuing other diagnostics. B. Recommend a cystocentesis and culture submission. C. Recommend a two to three weeks course of prophylactic antibiotics and recheck at that time.
Mechanical gastrointestinal obstruction	A five-year-old male neutered labrador mix presents to your clinic for vomiting and diarrhoea over 12–24 hours duration. After completing the history and physical exam, the owner agrees to two-view abdominal radiographs. Based on your history, physical exam and images, you feel there is an X per cent probability that this dog has a mechanical obstruction. A barium study or abdominal ultrasound is approximately US$250 each. Abdominal exploratory surgery will be an estimated US$2500. The owner has moderate financial limitations. You decide to...	X=(2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98)	A. Send the dog home with a prokinetic and antibiotic. Have the owner monitor at home and advise to return for a recheck if vomiting and diarrhoea continue. B. Recommend further diagnostics, such as a barium study or abdominal ultrasound. C. Recommend emergency abdominal exploratory surgery.
Idiopathic epilepsy	A two-year-old female spayed Beagle presents to your clinic for evaluation of an initial episode of a generalised seizure. Based on the history, physical and neurological exam and normal findings on blood work, you are X per cent certain that the dog has idiopathic epilepsy. MRI is US$1800 (assume you are working at a facility with an MRI). The owner has moderate financial restrictions. You decide to...	X=(2, 5, 8, 10, 12, 15, 18, 20, 25, 28, 30, 35, 38, 40, 45, 48, 50, 55, 58, 60, 65)	A. Treat this as an isolated incident or potential syncope and discharge the patient. B. Recommend an MRI. C. Recommend an antiseizure medication.
Chronic kidney disease (CKD)	An eight-year-old male neutered domestic short hair cat presents to your clinic for an annual wellness exam. Based on the history, physical exam and PCV (packed cell volume)/TS (total solids), you feel that there is an X per cent chance that this cat has early stage CKD. The owner has moderate financial restrictions and further lab work is approximately US$100. You decide to...	X=(10, 12, 15, 18, 20, 22, 25, 28, 30, 35, 38, 40, 45, 48, 50, 55, 58, 60, 65, 70, 75)	A. Tell the owner that CKD is unlikely, and recommend returning in 1 year unless signs of CKD develop such as increased thirst or inappropriate elimination. B. Recommend a blood chemistry and urinalysis to further evaluate kidney function. C. Recommend a prescription for a renal diet and have him return in 6 months for follow-up exam.

study is to determine test and treatment thresholds for suspected canine panosteitis, hypothyroidism, urinary tract infection (UTI), mechanical gastrointestinal obstruction (GIO), idiopathic epilepsy and feline chronic kidney disease (CKD). The hypothesis is that clinical decision thresholds can be determined for common conditions in small animal practice using the methods developed by Ebell and others.

Materials and methods

Study design

This was a cross-sectional survey design. Participants were asked to provide informed consent, and the online survey was both anonymous and voluntary. Qualtrics software was used to design and launch the online survey.

The clinical scenarios were developed by the primary investigator and were based on disorders emphasised during the University of Georgia (UGA) veterinary curriculum as commonly encountered conditions within companion animal practice. The clinical scenarios included suspected canine panosteitis, hypothyroidism, UTI, GIO, idiopathic epilepsy and feline CKD. Typical signalments and histories were applied to each scenario (table 1).

Demographic and background questions were included at the beginning of the survey and were mandatory to complete. These included gender, age, year of graduation from veterinary school, area of practice and postgraduate training. Clinical scenario questions (table 1) were not mandatory to complete. All clinical scenarios were presented to each participant, and participants could choose to skip these questions if preferred.

Before survey distribution, the clinical scenarios were pilot tested on five small animal clinicians within the UGA faculty and were altered according to their suggestions and critiques. Wide ranges of probability were applied to conditions in which a clinician may realistically reach extreme levels of certainty based simply on the patient's history, signalment, physical exam and/or minimal diagnostics. These scenarios included panosteitis (probabilities ranged from 2 per cent to 90 per cent), UTI (probabilities ranged from 2 per cent to 98 per cent) and GIO (probabilities ranged from 2 per cent to 98 per cent). Narrower ranges of probability were applied to conditions in which it would be unlikely for clinicians to reach as extreme levels of certainty without further examination and/or diagnostics. These scenarios included hypothyroidism (probabilities ranged from 2 per cent to 75 per cent), idiopathic epilepsy (probabilities ranged from 2 per cent to 65 per cent) and CKD (probabilities ranged from 10 per cent to 75 per cent). Respondents received randomly assigned probabilities for each of the six clinical scenario questions and selected responses based on this randomly assigned probability. Because established prevalence data for these diseases are lacking, the ranges of probabilities were developed from pretest probabilities based on intuitive reasoning and expert opinion.

The panosteitis and hypothyroidism scenarios provided two possible options. For the panosteitis scenario, the options included either 'testing' by performing radiographs or treating empirically for the disease. The hypothyroidism scenario allowed one to 'rule-out' hypothyroidism by sending the patient home or 'test' by performing bloodwork. The UTI, GIO, idiopathic epilepsy and CKD scenarios allowed one to choose from the three following options: rule the condition out, pursue further diagnostics or treat for the suspected condition. Ruling a condition out in this context involves deciding that the patient does not have the disease of interest or feeling that the condition is unlikely enough that the best course is to observe or consider other aetiologies.

Study participants

A convenience sample of small animal general practitioners, internists, criticalists and a handful of other specialties was collected through various recruitment methods. Emails with an invitation to participate and the survey link were sent out to the UGA Alumni Association listserv, the American College of Veterinary Internal Medicine (ACVIM) listserv (n>1000) and the American College of Veterinary Emergency and Critical Care (ACVECC) listserv (n>1000). The survey was included in a monthly UGA rDVM e-brief with an invitation to participate. The survey link and invitation were also posted to the Georgia Veterinary Medical Association (GVMA) Facebook page (n=285). Flyers were distributed at the UGA Alumni Association weekend with the survey link and an invitation to participate (n=120). Finally, several small animal clinics (Ansley Animal Clinic, Ark Animal Hospital, Paces Ferry Veterinary Clinic, Peachtree Hills Animal Hospital, Treehouse Animal Clinic, Briarcliff Animal Clinic, Intown Animal Hospital, Clairmont Animal Hospital, Ormewood Animal Hospital, and Avondale Veterinary Hospital) in the metro-Atlanta area were sent invitation emails with the survey link. Because it is likely that many small animal clinicians may have experienced overlapping recruitment, it was requested that each participant only complete the survey once. Altogether, over 2500 veterinarians were contacted to participate in the study. There were 314 total responses received, 297 of which were included in the final data analysis. Responses were excluded if more than 50 per cent of the questions were incomplete.

Data analysis

Descriptive statistics were carried out using percentages, medians and IQRs where needed. The distribution of clinical decision (rule out/test/treat) on sample subgroups according to clinician gender, age, internship completion (yes/no) and doctor practice (general practice (GP), internal medicine (IM), emergency and critical care (ECC) and other) were represented using bar plots.

Threshold analysis was performed using the same statistical model developed by Ebell and others.[8] This logistic regression method was adapted from the methods by Plascencia and others.[9] The *test threshold* was determined through the following logistic function:

$$In\left[p_1/(1-p_1)\right] = a_1 + b_1 x \quad (1)$$

where p_1 is the probability of testing or treating (not ruling out) for the suspected condition; a_1 is regression constant; b_1 is regression coefficient; x is disease probability randomly provided in the scenario.

The threshold is defined as the probability of disease where there is equipoise between two decisions. For the testing threshold, this is the probability of disease where it is equally likely that the clinician would choose to rule out disease on the one hand, or test or treat on the other. Thus, the test threshold is defined as the disease probability x_{test} such that the corresponding probability of not ruling out, as estimated by the logistic model, is equal to 0.5: $p_1(x_{test})$=0.5.

Solving equation (1) with respect to x and replacing p_1 by 0.5, one obtains:

where $\hat{a}_1$ and $\hat{b}_1$ are model (1) estimated coefficients.

The *treatment threshold* was estimated in a similar fashion using a second logistic function:

$$In\left[p_2/(1-p_2)\right] = a_2 + b_2 x \quad (2)$$

where p_2 is the probability of treating (not ruling out nor testing) the suspected condition. Thus, the treatment threshold is defined as the disease probability x_{treat} such that the corresponding probability of treating, as estimated by the logistic model (2), is equal to 0.5:

$$\hat{P}_1(\chi test) = 0.5$$

Solving equation (2) with respect to x and replacing by 0.5, one obtains:

$$x_{treat} = -\hat{a}_2/\hat{b}_2$$

where $\hat{a}_2$ and $\hat{b}_2$ are model (2) estimated coefficients. Confidence intervals were determined through the adapted model described by Ebell and colleagues.[8]

Models (1) and (2) were adjusted in order to account for clinician characteristics: gender, age, practice type (GP/IM/ECC/other) and internship completion (yes/no). Interactions between gender and age and between gender and internship completion were also tested in the initial multivariate model. A backward selection based on Akaike Information Criterion (AIC) was carried out in order to select the final model. Thresholds (with 95 per cent CIs) were estimated for each level of the retained covariates.

Table 2 Characteristics of survey respondents

Characteristic	Variable	n (%)
Gender	Male	91 (30.64%)
	Female	206 (69.36%)
Area of practice	General practice	47 (15.82%)
	Internal medicine	141 (47.47%)
	Emergency and critical care	89 (29.97%)
	Other*	20 (6.73%)
Completed internship/ postgraduate training	Yes	246 (82.83%)
	No	51 (17.17%)
Age (years)	Median	41±36–51.5 (25–68)
Graduation year	Median	2002±1992–2007 (1956–2015)

For continuous data, values are reported as median±IQR (range).
*Written responses for respondents that reported 'other' for area of practice include military service, cardiology, soft tissue surgery, academia, neurology, industry, oncology, exotic/zoo animal and preventive medicine.

Statistical analyses were performed using software R (R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/). Logistic regressions were performed using the function glm(). CIs around thresholds were obtained within the model using the function vcov() giving the covariance matrix of the parameter estimates. Graphs of predicted probability of the logistic regression were obtained using graphical capabilities of R.

Results

There were a total of 314 responses of which 297 were usable. Demographics and characteristics of respondents are summarised in table 2.

The test and treatment thresholds were determined for UTI and GIO. For UTI, the test threshold was identified as 12.8 per cent (95 per cent CI=1.1 to 20.7) and the treatment threshold as 82 per cent (95 per cent CI=66.3 to 100). For GIO, the test threshold was identified as 3.2 per cent (95 per cent CI=0 to 10.4) and the treatment threshold as 87.3 per cent (95 per cent CI=82.6 to 93.5) (figure 2).

All other scenarios did not allow us to estimate thresholds in the meaningful probability range of 0–1. A threshold smaller than zero can be interpreted as a case where doctors *always* take a given decision; a threshold larger than one can be understood as describing a situation where doctors *never* take the decision.

For the panosteitis scenario (figure 3a), randomly assigned probabilities ranged from 5 per cent to 90 per cent. Even at the high extremes of probability (≥75 per cent), 58 per cent of respondents elected to perform radiographs before treating. Based on these data, one would conclude that clinicians will *always* pursue further diagnostics before treating for suspected canine panosteitis.

Similar outcomes for threshold analysis were found for hypothyroidism (figure 3b) and feline CKD (figure 3c). For the hypothyroidism scenario,

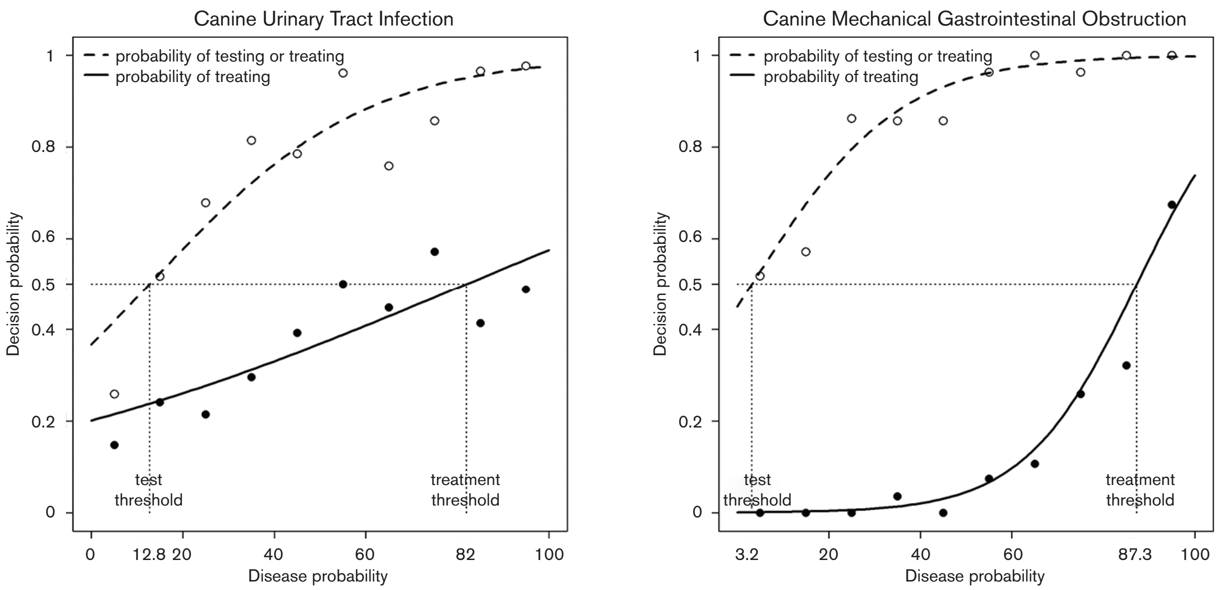

Figure 2 Decision thresholds for canine urinary tract infection (A) and mechanical gastrointestinal obstruction (B) (thresholds results within the interval 0–1). Open and filled circles represent the empirical proportions of doctors not ruling out (respectively treating) for each 10 points disease probability interval and they are drawn at the midpoint of each interval (5 per cent for the interval 0 per cent–10 per cent , ...). Broken and solid lines represent probabilities of not ruling out (respectively treating) estimated by a logistic regression models as a function of disease probabilities.

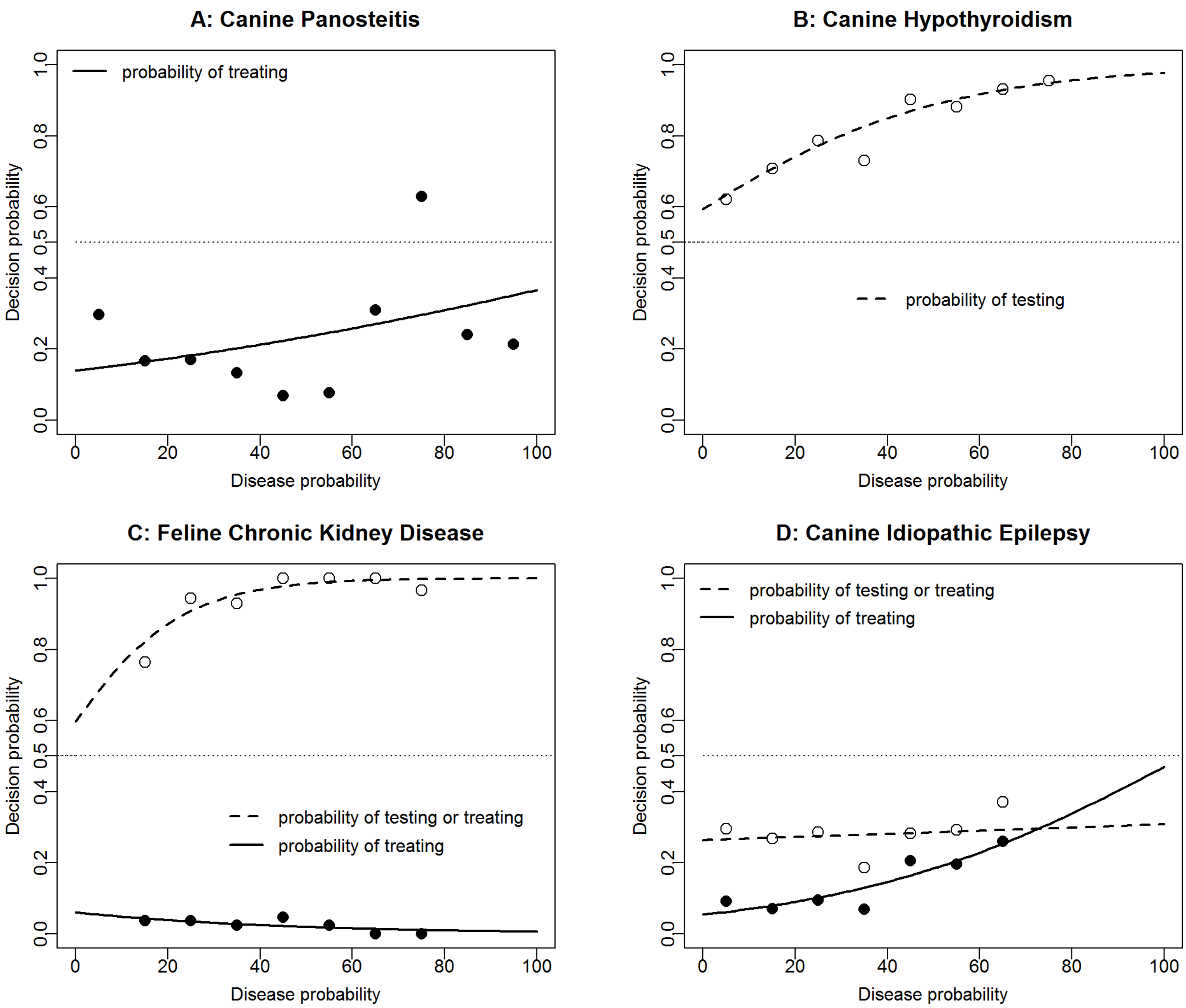

Figure 3 Decision probabilities for canine panosteitis (a), hypothyroidism (b), idiopathic epilepsy (d) and feline chronic kidney disease (c) (thresholds results beyond the interval 0–1). Open and filled circles represent the empirical proportions of doctors not ruling out (respectively treating) for each 10 points disease probability interval and they are drawn at the midpoint of each interval (5 per cent for the interval 0 per cent–10 per cent, ...). Broken and solid lines represent probabilities of not ruling out (respectively treating) estimated by a logistic regression models as a function of disease probabilities.

probabilities ranged from 2 per cent to 75 per cent. Even at low levels of probability (≤30 per cent), 71 per cent of respondents chose to perform blood work instead of sending the patient home. For the Feline CKD scenario, where probabilities ranged from 10 per cent to 75 per cent, at both low and high levels of probability 90 per cent of respondents chose to recommend further lab work as opposed to monitoring at home or prescribing a renal diet. Based on these data, one would conclude that clinicians will *always* pursue further diagnostics before ruling out suspected canine hypothyroidism. These data also indicate that clinicians will *always* pursue further diagnostics before ruling out or treating for feline CKD.

For the idiopathic epilepsy scenario (figure 3d), randomly assigned probabilities ranged from 2 per cent to 65 per cent. In this case, at all levels of probability, 71% of respondents elected to send the patient home as opposed to performing advanced imaging or prescribing an antiseizure medication. Based on these data, one would conclude that clinicians will *never* test nor treat for canine idiopathic epilepsy.

In the case of UTI and GIO, the authors stratified the clinical decisions into subgroups of the sample population according to respondent gender, age, practice and internship. In the UTI case (figure 4), clinician decisions distributed quite fairly among the three options (rule out, test and treat), while in the GIO case (figure 5) the test option was the most likely to be chosen for each population subgroup. Empirical discrepancies observed among groups are the following: male clinicians, clinicians who have not completed internship training and general practitioners decided to treat for UTI more often than their colleagues; clinicians who have not completed internship training and clinicians which identified their practice as ‘other’ decided to rule out for GIO more often than their colleagues, while general practitioners decided more often to treat in this scenario.

In order to account for confounding, the authors performed multivariate threshold analysis including respondent gender, age, practice and internship into models (1) and (2), also testing for gender-age and gender-internship interactions. Results of the final

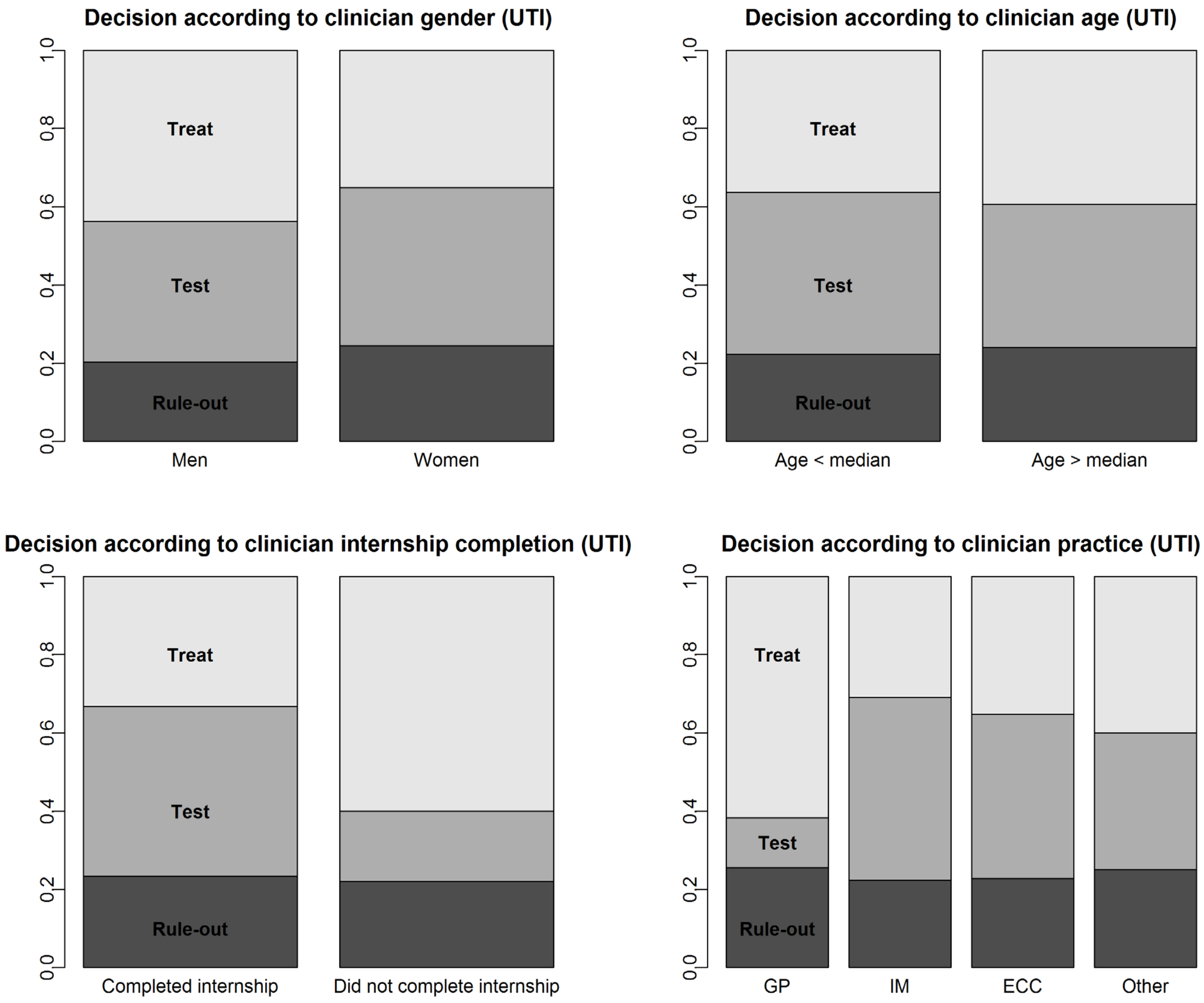

Figure 4 Distribution of clinical decisions (rule out/test/treat) on sample subgroups according to doctor gender, age (</>median age of 41 years), internship completion (yes/no) and doctor practice (GP/IM/ECC/‘other’). Urinary tract infection (UTI) scenario. GP, general practice; ECC, emergency and critical care; IM, internal medicine.

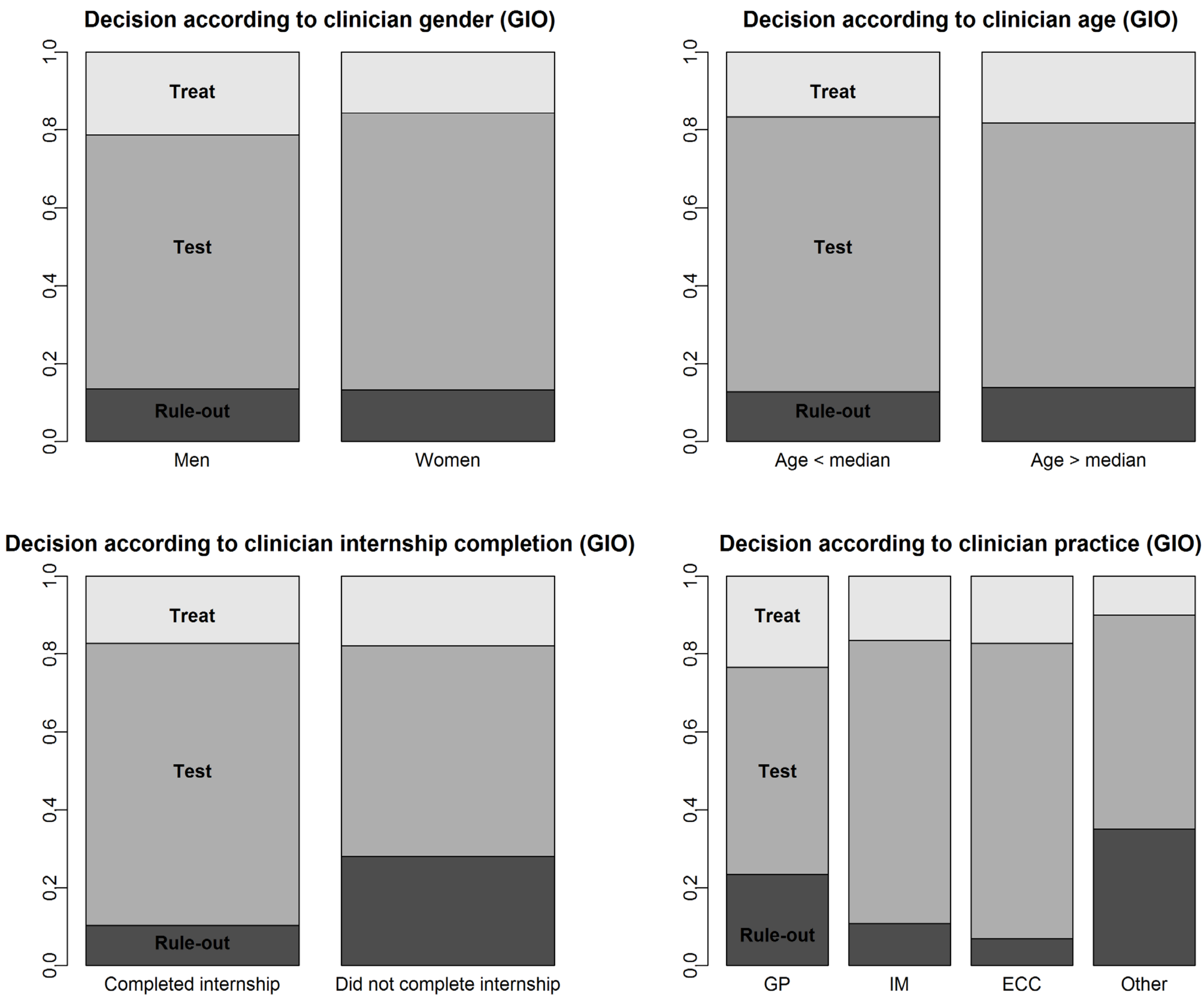

Figure 5 Distribution of clinical decisions (rule out/test/ treat) on sample subgroups according to doctor gender, age (</>median age of 41 years), internship completion (yes/no) and doctor practice (GP/IM/ECC/'other'). Gastrointestinal obstruction (GIO) scenario. GP, general practice; ECC, emergency and critical care; IM, internal medicine.

models obtained via backward selection based on AIC are presented in table 3. In the case of UTI, the selected test threshold model only included the disease probability (model 1): none of the considered individual characteristics showed an effect on probability of testing for UTI in an adjusted model. On the contrary, the probability of treating for UTI was significantly affected by internship and marginally affected by gender (without interaction between the two) in the selected model (table 3). Women showed to be less predisposed to treat than men (OR=0.62, P=0.09), and clinicians who had not completed internship training were estimated to be more inclined to treat (OR=3, P<0.001). In the case of GIO, both the probability of testing and the probability of treating were influenced by individual characteristics. Clinicians who had not completed internship training were less inclined to test (OR=0.29, P=0.008) as were respondents that identified their area of practice as 'other' (OR=0.17, P=0.007). General practitioners showed to be more predisposed to treat (OR=2.96, P=0.04). All described adjusted covariate effects confirmed empirical discrepancies observed in figures 4 and 5. Thresholds estimates for each combination of values of the selected variables are given in table 3 along with their CIs.

Discussion

This study succeeded in determining test and treatment thresholds for two commonly encountered conditions in small animal practice: UTI and GIO. This sample of veterinarians is comfortable simply observing when the likelihood of a UTI is less than 13 per cent and is comfortable treating when the likelihood of a UTI is greater than 82 per cent. Because UTIs were considered a relatively low-risk condition in this study, the authors would expect this test threshold to be somewhat higher.[12] The high treatment threshold determined through this study may be a result of veterinarians taking greater caution in prescribing antibiotics without a definitive diagnosis.[13]

Multivariate analysis (table 3) revealed that female clinicians were more likely to reach for further

Table 3 Adjusted threshold models taking into account characteristics (gender, age, practice and internship) also testing for gender-age and gender-internship interactions Results of final models obtained via backward selection based on AIC (Akaike Information Criterion)

	OR	P value	Threshold	
UTI: test threshold				
Intercept	0.58	0.036	Overall	12.8 (1.05 to 20.7)
Disease probability	1.53	<0.001		
	Estimate	P value	Thresholds	
UTI: treatment threshold				
Intercept	−1.25	<0.001	Gender: male Internship: yes	75.0 (48.1 to 100)
Disease probability	1.18	<0.001	Gender: male Internship: no	8.7 (0.0 to 51.4)
Gender: female	0.63	0.090	Gender: female Internship: yes	100 (79.5 to 100)
Internship: no	3.03	<0.001	Gender: female Internship: no	36.4 (0.0 to 73.7)
	OR	P value	Thresholds	
GIO: test threshold				
Intercept	1.25	0.509	Practice: GP, IM or ECC Internship: yes	0.00 (0.00 to 5.63)
Disease probability	1.90	<0.001	Practice: GP, IM or ECC Internship: no	15.7 (1.14 to 27.9)
Practice: 'other'	0.17	0.007	Practice: other Internship: yes	24.1 (3.14 to 43.6)
Internship: no	0.29	0.008	Practice: other Internship: no	43.3 (20.8 to 68.3)
	Estimate	P value	Thresholds	
GIO: treatment threshold				
Intercept	0.00	<0.001	Practice: IM, ECC or other	89.0 (84.0 to 95.7)
Disease probability	2.33	<0.001	Practice: GP	76.2 (65.0 to 88.3)
Practice: GP	2.96	0.042		

Results of final models obtained via backward selection based on Akaike Information Criterion.
ECC, emergency and critical care; GIO, gastrointestinal obstruction; GP, general practice; IM, internal medicine; UTI, urinary tract infection.

diagnostics before initiating treatment for suspected UTI as compared with male colleagues. Future studies may consider investigating potential impacts of gender on clinical decision making. The empirical analysis of clinical decision making (figure 4) revealed that gender, internship completion and area of practice may impact a clinician's decision to provide empirical therapy for a UTI, as respondents that were male, had completed an internship or were general practitioners were more likely to do so. For the discrepancy between general practitioners and other areas of practice, this finding may be related to the socioeconomic status of typical clients.

This sample of veterinarians is comfortable simply observing the patient when the likelihood of a GIO is less than 3 per cent and is comfortable performing an abdominal exploratory surgery when the likelihood of a GIO is greater than 87 per cent. At face value these thresholds are reasonable for a high-risk condition such as GIO. This is a condition in which missing the diagnosis is dangerous and potentially fatal, yet at the same time the treatment option is also an invasive surgical procedure that requires general anaesthesia and also carries its own risks of mortality.[14] The thresholds for high-risk conditions, such as acute coronary syndrome, determined by Ebell and others similarly had a relatively low test threshold and high treatment threshold.[8]

Multivariate analysis (table 3) showed that respondents with internship training were more likely to test before treating for suspected GIO, and general practitioners were more likely to treat without testing. This finding could be linked to different barriers that general practitioners face as compared with their colleagues in specialty or referral practice. Those respondents in general practice may not have ready access to more advanced imaging such as ultrasound, and may hesitate to use upper GI studies when there is the possibility of surgery. It is also a valid assumption that those respondents that completed internship training are more likely to have gone on to specialise.

There are several possibilities as to why thresholds were not producible for panosteitis, hypothyroidism, idiopathic epilepsy and CKD. For the conditions in which respondents overwhelmingly chose to pursue further diagnostics no matter the level of probability (panosteitis, hypothyroidism and CKD), the diagnostics are generally accessible and non-invasive. For panosteitis, respondents may have been concerned about ruling out a more high-risk condition (such as traumatic injury, immune-mediated joint disease, septic physitis, etc) before ruling in a self-limiting condition such as panosteitis.

In the case of hypothyroidism, a similar scenario may have taken place in which clinicians were focused on both ruling out more serious conditions with a similar presentation (such as neoplasia) as well as potentially ruling in a medically treatable condition like hypothyroidism. Similarly, respondents may have approached the CKD scenario with the mindset that a responsible workup should include blood and urine analysis no matter the probability of the suspected condition.

Respondents overwhelmingly chose to send the patient home for further observation for the idiopathic epilepsy scenario, no matter the level of probability. This may be because idiopathic epilepsy is an expensive and difficult condition to diagnose.[15] A single seizure episode also may not have been reason enough for most of the respondents to suggest imaging or prescribe an anticonvulsant medication. Because of realistic restrictions and the limited nature of the scenario questions, respondents may have been tempted to choose the option of sending the patient home for further observation no matter the probability of disease.

Survey respondents may have approached these scenarios without fully considering the probability of the condition. The signalment, clinical presentation and the respondent's own experience may have had greater influence on the decision than the provided probability. Previous studies have found that physicians very rarely use numerical probabilities to make decisions, and often prefer communication of uncertainty in qualitative terms (such as 'unlikely' or 'probable') rather than in a numerical context.[16] There is limited evidence as to how veterinary practitioners interpret numerical uncertainty, however Christopher and others found that clinicians showed wide variation in their interpretation of terms expressing probability of cytological diagnoses.[17] These variations in clinical decision making are consistent with authors' findings, and further research should examine both quantitative and qualitative approaches in establishing clinical thresholds within veterinary practice.

There are several limitations to this study design and the application of this model to small animal veterinary practice. Future studies should aim to increase the sample size as this will provide more precise estimates of the threshold levels. A larger sample would likely capture a more even distribution of respondents' training and backgrounds. The pretest probability is the probability of disease in a patient before conducting any tests and should be the starting point at which a clinician weighs the likelihood of certain diagnoses. This probability may be based on one's clinical experience, but ideally should be based on sound epidemiological data, a resource that veterinary medicine is deficient in compared with human medicine.[3] An important difference between using the threshold model in human and veterinary medicine is that veterinary medicine lacks pretest probabilities that are based on surveillance data or studies of disease prevalence. The survey design was impacted by the lack of prevalence data for the conditions of interest. Ebell and others applied specific ranges of probability to each clinical scenario centred around the data based pretest probability of each condition.[8] For this study, ranges of probability were intentionally very broad in an effort to capture the true probabilities of disease. This limitation may have impacted the study's capacity to determine threshold levels for those conditions that did not produce usable data.

Because veterinarians have access to only limited levels of epidemiological data, it may be that many veterinarians do not think in terms of probability of disease based on breed and signalment or simply think in terms of probability based on one's own clinical expertise. For the conditions in which thresholds could not be determined, respondents may not have not understood the question or may have simply disregarded the level of probability altogether—it is possible that respondents answered the questions purely based on the provided signalment and history. Survey design, choice of conditions and limited answer options also may have led to inadequate data.

It is also important to note that arriving at a proper clinical diagnosis requires synthesis and analysis of multiple levels of information—this is a complex process, one that cannot be accurately represented through multiple-choice responses. There are a range of factors that influence a veterinarian's decisions about an individual patient including the relationship with the client, the client's financial resources, the client's relationship with the patient and various other circumstances.[18] The information provided in the survey was purposely limited with the objective that respondents would aim their focus on the probability when making the decision. Future studies should aim to craft clinical scenarios in a simple manner, yet make the choices as realistic as possible.

Price sensitivities of owners play a large role in diagnostic and treatment choices; therefore, fabricated monetary values were applied to each diagnostic and treatment option. Because prices of services can vary drastically between clinics, the applied monetary values in this study may have impacted respondents' choices in an inconsistent fashion. Future studies should consider allowing respondents to apply their own practice's price points for each scenario in order to reduce this potential effect, or possibly avoid cost completely through a scenario in which the client has pet insurance with no financial constraints whatsoever. While there are several limitations associated with this specific study, the study design has the potential to increase understanding in how veterinarians use evidence to make decisions. Continued improvements in study questionnaire design may produce more clinical thresholds for small animal practice.

This study has identified clinical decision thresholds for two common conditions in small animal medicine—canine mechanical GOI and canine UTIs. It is possible to identify test and treatment thresholds in small animal practice. Future research may broaden investigation of methods to determine group clinical threshold levels among veterinarians, which may be used as the basis for clinical decision rules that use low-risk, moderate-risk and high-risk groups, as well as examine other factors, outside of probability, that affect clinician diagnostic and treatment decisions.

Acknowledgements The authors would like to thank the participating clinicians at the University of Georgia College of Veterinary Medicine for their insight into developing the online survey as well as all the clinicians who took the time to complete our survey. A link to the online Qualtrics survey: https://ugeorgia.qualtrics.com/SE/?SID=SV_5BWpLFLFFF2yWEt.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Ethics approval The University of Georgia Institutional Review Board approved the study (IRB ID: STUDY00003188).

References

1 May SA. Clinical reasoning and case-based decision making: the fundamental challenge to veterinary educators. *J Vet Med Educ* 2013;40:200–9.
2 Schmidt PL. Evidence-based veterinary medicine: evolution, revolution, or repackaging of veterinary practice? . Veterinary Clinics of North America: Small Animal Practice, 2007:37. 409–17.
3 Vandeweerd JM, Kirschvink N, Clegg P, *et al*. Is evidence-based medicine so evident in veterinary research and practice? History, obstacles and perspectives. *Vet J* 2012;191:28–34.
4 Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980;302:1109–17.
5 Boland MV, Lehmann HP. A new method for determining physician decision thresholds using empiric, uncertain recommendations. *BMC Med Inform Decis Mak* 2010;10:20.
6 Eraker SA, Eeckhoudt LR, Vanbutsele RJM, *et al*. To test or not to test — To treat or not to treat. *J Gen Intern Med* 1986;1:177–82.
7 Eisenberg JM, Hershey JC. Derived thresholds. Determining the diagnostic probabilities at which clinicians initiate testing and treatment. *Med Decis Making* 1983;3:155–68.
8 Ebell MH, Locatelli I, Senn N. A novel approach to the determination of clinical decision thresholds. *Evid Based Med* 2015;20:41–7.
9 Plasencia CM, Alderman BW, Barón AE, *et al*. A method to describe physician decision thresholds and its application in examining the diagnosis of coronary artery disease based on exercise treadmill testing. *Med Decis Making* 1992;12:204–12.
10 Boyles T, Locatelli I, Senn N, *et al*. Determining clinical decision thresholds for HIV-positive patients suspected of having tuberculosis. *Evid Based Med* 2017;22:132–8.
11 O'Neill DG, Church DB, McGreevy PD, *et al*. Approaches to canine health surveillance. *Canine Genet Epidemiol* 2014;1:2.
12 Smee N, Loyd K, Grauer GF. UTIs in small animal patients: part 2: diagnosis, treatment, and complications. *J Am Anim Hosp Assoc* 2013;49:83–94.
13 Weese JS, Giguère S, Guardabassi L, *et al*. ACVIM consensus statement on therapeutic antimicrobial use in animals and antimicrobial resistance. *J Vet Intern Med* 2015;29:487–98.
14 Ettinger SJ, Feldman EC. Textbook of Veterinary Internal Medicine-eBook: Elsevier health sciences, 2009.
15 De Risio L, Bhatti S, Muñana K, *et al*. International veterinary epilepsy task force consensus proposal: diagnostic approach to epilepsy in dogs. *BMC Vet Res* 2015;11:148.
16 Renooij S, Witteman C. Talking probabilities: communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning* 1999;22:169–94.
17 Christopher MM, Hotz CS, Shelly SM, *et al*. Interpretation by clinicians of probability expressions in cytology reports and effect on clinical decision-making. *J Vet Intern Med* 2010;24:496–503.
18 Everitt S. Clinical decision making in veterinary practice: Doctoral dissertation, University of Nottingham), 2011.