Eurasia Specialized Veterinary Publication

International Journal of Veterinary Research and Allied Science

ISSN:3062-357X

2025, Volume 5, Issue 2, Page No: 22-31

Available online at: www.esvpub.com/

The Reference Interval Is Not the Patient: A Clinical Reasoning Framework for Reconciling Serial Biomarkers, Biological Variation, Imaging Evidence, Pretest Probability, and Discordant Findings in Veterinary Diagnosis

Sofía Martínez[1]*, Carlos Gómez[1], Lucia Navarro[2]

[1]Department of Veterinary Clinical Pathology and Diagnostic Reasoning, Faculty of Veterinary Medicine, Polytechnic University of Madrid, Madrid, Spain.
[2]Department of Veterinary Internal Medicine and Diagnostic Imaging, Faculty of Veterinary Medicine, University of Barcelona, Barcelona, Spain.

***E-mail** ✉ sofia.martinez@upm.es

ABSTRACT

Reference intervals are indispensable for veterinary laboratory interpretation, but they describe distributions in selected reference populations rather than the evolving biological state of an individual animal. Diagnostic error may arise when population-normality is treated as equivalent to patient stability, when serial change is interpreted without accounting for analytical or biological variation, or when laboratory, imaging, and clinical findings are expected to converge synchronously. This article develops an original non-empirical diagnostic reasoning framework for reconciling population reference values, individual biological baselines, preanalytical and analytical variation, serial biomarker trajectories, pretest probability, imaging evidence, and discordant findings. The synthesis distinguishes whether a new result represents expected fluctuation, measurement-related change, biologically meaningful deviation, context-dependent evidence, or unresolved discordance requiring reassessment. It further argues that the diagnostic meaning of any finding depends not only on whether it crosses a reference limit, but also on its direction and magnitude of change, the reliability of the measurement process, the patient's previous trajectory, disease probability before testing, and agreement or disagreement with independent evidence streams. The proposed framework is intended to organize reasoning rather than provide a numerical diagnostic score or validated decision rule. Its applicability will depend on species, analyte, disease, care setting, temporal stage, availability of prior measurements, and quality of the underlying biological-variation and diagnostic-performance evidence. Prospective validation is therefore required before claims of improved diagnostic accuracy, clinical utility, or patient outcomes can be made.

Keywords: Biological variation, Reference intervals, Serial biomarkers, Pretest probability, Diagnostic discordance, Veterinary imaging

Received: 21 June 2025
Revised: 18 August 2025
Accepted: 23 August 2025

How to Cite This Article: Martínez S, Gómez C, Navarro L. The Reference Interval Is Not the Patient: A Clinical Reasoning Framework for Reconciling Serial Biomarkers, Biological Variation, Imaging Evidence, Pretest Probability, and Discordant Findings in Veterinary Diagnosis. Int J Vet Res Allied Sci. 2025;5(2):22-31.
https://doi.org/10.51847/8HEQtsIidH

Introduction

Veterinary clinicians commonly interpret laboratory results by asking whether a value lies inside or outside a population reference interval. That comparison is useful, but it cannot by itself determine whether an individual patient's physiology is stable. Observed laboratory values reflect both biological and analytical variation, and the magnitude of expected within-patient fluctuation can differ substantially among analytes [1]. A value that remains within a conventional reference interval may therefore represent meaningful movement away from an animal's previous state, whereas an isolated value outside the interval may reflect expected individuality rather than new

disease. Contemporary veterinary biological-variation scholarship similarly supports separating population reference distributions from within-individual change and using serial information where sufficiently robust data exist [2].

The same reasoning problem extends beyond laboratory medicine. Diagnostic interpretation is conditional on clinical context: in canine thoracic radiography, providing relevant history increased diagnostic accuracy even though inter-reader agreement did not necessarily improve [3]. Imaging is therefore not context-free evidence, and a visually apparent lesion cannot be interpreted independently of the question being asked, the prior probability of disease, or alternative causes.

Canine pancreatitis illustrates the broader problem particularly well because clinical signs, pancreatic lipase measurements, ultrasonography, and other diagnostic modalities may contribute different and imperfect information [4]. The central argument developed here is that diagnostic reasoning should focus on reconciling evidence around the patient rather than forcing each observation into a binary normal/abnormal classification. The proposed framework separates population reference status, measurement validity, individual baseline, serial trajectory, contextual probability, multimodal concordance, competing explanations, diagnostic confidence, and reassessment. These distinctions are intended as an evidence-bounded reasoning architecture, not as a validated diagnostic algorithm.

Reference intervals and their clinical boundaries

A reference interval answers a population question: where do measurements from an appropriately defined reference population usually fall under specified analytical conditions? It does not directly answer whether a particular animal is healthy, whether disease is progressing, or whether a change should trigger intervention. Veterinary biological-variation literature emphasizes that analytical variation, within-individual variation, and between-individual variation are separable components of observed laboratory variability [1]. A reference limit can therefore be crossed for reasons other than clinically important deterioration, and clinically important change can occur without crossing it.

Population dispersion and individual biological fluctuation must consequently be treated as related but distinct interpretive layers [1, 2]. When within-individual variation is narrower than variation across the reference population, a substantial patient-specific change can remain numerically “normal.” Conversely, an animal whose stable homeostatic set point lies near a population boundary may intermittently fall outside that interval without having undergone a meaningful change in physiological state.

This distinction also separates reference intervals from clinical decision limits. The latter may incorporate disease risk, treatment consequences, or diagnostic performance, whereas a population reference boundary need not. The proposed framework therefore treats reference status as one evidence descriptor rather than a verdict. Its relevance increases or decreases according to analyte-specific biological variation, measurement quality, previous results, disease context, and the consequences of acting on an abnormal classification.

Population reference values versus individual biological baselines

Longitudinal information can reveal distinctions that a population interval cannot. In healthy cats, repeated measurement of urinary protein:creatinine ratio and urine specific gravity demonstrated appreciable biological variation, and some initially abnormal protein measurements were not persistently abnormal on subsequent sampling [5]. The clinical implication is not that repeated testing automatically resolves uncertainty, but that a single comparison with a population boundary may inadequately characterize an individual animal's usual state.

Comparable canine work demonstrates that the amount of serial change considered unusual is analyte specific. In healthy dogs, weekly biological variation differed markedly between urinary protein:creatinine ratio and urine specific gravity, producing substantially different reference-change behavior [6]. This argues against generic rules such as assuming that the same percentage increase has equivalent meaning across measurements.

An individual biological baseline should therefore be conceptualized as a trajectory rather than a single historical number. Its evidentiary value depends on whether earlier measurements were obtained under comparable clinical and analytical conditions, whether the animal was plausibly in a stable state, and whether the analyte has sufficiently characterized within-patient variation. A prior result should not be privileged merely because it is earlier. Instead, repeated compatible observations can establish a more credible patient-specific context against which subsequent change is judged. The proposed framework uses this baseline as an additional reference frame, not as a replacement for population intervals or disease-specific decision thresholds.

Analytical and preanalytical variation

Before biological meaning is assigned to a serial difference, clinicians must consider whether the measurement process itself changed. Feeding is one example. In healthy dogs evaluated before and after food intake, selected biochemical analytes changed after feeding, demonstrating that preanalytical state can produce apparent biological movement even in the absence of newly developing disease [7]. Comparable samples are therefore preferable when a serial trend is being interpreted clinically.

Storage and analytical platform can also alter apparent trajectories. Studies of symmetric dimethylarginine in canine and feline serum show that measurement stability and agreement depend on storage conditions and analyzer characteristics [8]. A newly discordant result can consequently originate from the interval between patient and analyzer rather than from the patient alone.

Preanalytical interference may also act selectively rather than globally. In dogs studied after a hypercaloric meal, postprandial lipemia had limited effects on several direct blood-gas measurements but altered some derived biochemical or acid-base interpretations [9]. This matters diagnostically because apparently coherent calculated abnormalities can inherit error from affected component measurements even when other values remain stable.

The proposed reasoning sequence therefore requires a measurement-validity check before a serial difference is classified as physiological change. Relevant questions include sampling state, collection technique, processing delay, storage, analyzer platform, and known interference. This is not an invitation to dismiss unexpected abnormalities as laboratory artifacts. Rather, analytical and preanalytical explanations should remain active competing hypotheses until the result has been shown sufficiently comparable with the measurements against which it is being judged.

Biological variation within the same patient

Even under stable measurement conditions, biological systems fluctuate. Within-individual variation reflects ordinary movement around a patient's homeostatic state and is conceptually distinct from analytical imprecision and between-individual variation [1]. This makes “different from last time” an incomplete clinical conclusion: the important question is whether the observed difference is larger, more directional, or more persistent than would reasonably be expected for that analyte and patient context.

Repeated feline urinary measurements illustrate why this matters. Variability in urine protein:creatinine ratio and urine specific gravity occurred among healthy animals, and transient departures from conventional expectations did not necessarily persist [5]. Serial interpretation can therefore reduce overreaction to isolated values, but only when the available biological-variation evidence applies to a sufficiently similar population.

The converse problem is equally important. Healthy-dog data show that biologically meaningful reference-change behavior differs sharply between urinary measures [6]. Consequently, a modest numerical shift in one analyte may be unusual while a larger proportional change in another remains compatible with expected fluctuation. Biological variation cannot be reduced to a universal percentage threshold.

The framework proposed here treats within-patient variation as an uncertainty envelope around a trajectory. Movement inside that envelope does not prove health, and movement outside it does not prove disease. Instead, the magnitude and direction of change become evidence that must be reconciled with measurement validity, clinical context, other biomarkers, imaging, and subsequent observations.

Serial change and longitudinal biomarker interpretation

Serial data become clinically informative when change is interpreted against an expected temporal pattern rather than against the previous result alone. Longitudinal evaluation of apparently healthy Golden Retrievers showed that year-to-year biochemical changes could be characterized and validated within that cohort, but the work did not establish that those change intervals detect incipient disease with known diagnostic accuracy [10]. The distinction is fundamental: describing unusual change is not equivalent to diagnosing its cause.

Disease-associated biomarkers also follow different kinetic trajectories. In dogs with pneumonia, septic peritonitis, or pyometra, inflammatory biomarkers changed over time at different rates, and some normalized before clinician-directed antimicrobial treatment ended [11]. Serial biomarker improvement can therefore be biologically meaningful without constituting a validated treatment-stopping rule.

Similarly, longitudinal cardiac variables in dogs with atrial fibrillation evolved during follow-up, yet baseline measurements did not reliably predict subsequent rate-control status [12]. A variable can thus be useful for describing a trajectory while remaining insufficient as a prospective predictor.

For this article, serial interpretation is separated into four questions: whether the measurement is comparable with prior measurements, whether the magnitude of change exceeds expected variation, whether its direction is biologically coherent with the suspected process, and whether other evidence streams evolve on a compatible timescale. These distinctions prevent "change" from being treated as a single diagnostic state.

Table 1 orders the serial-comparison questions so that measurement comparability, expected variation, biological coherence, and multimodal timing are resolved before diagnostic weight is assigned.

Table 1. Sequential questions that determine whether a serial biomarker difference can be interpreted biologically

Question in the serial comparison	Finding that permits progression	Finding that diverts interpretation	Conclusion permitted at this step
Were collection and measurement conditions comparable?	Sampling state, handling, storage, and analyzer are sufficiently aligned	Feeding, handling, storage, interference, or analyzer conditions changed	A nonbiological contribution is less likely, or comparability must be restored before inference
Does the difference exceed expected analytical and within-patient variation?	Directional or persistent movement is unusual for the analyte and context	Small, non-directional movement remains compatible with expected fluctuation	The change is unusual or compatible with fluctuation; neither conclusion identifies disease
Is the direction coherent with the suspected biological process?	Trajectory and disease context agree	Direction lacks clinical or mechanistic fit	Increase evidential weight or retain competing explanations
Are independent evidence streams moving on compatible timescales?	Clinical, laboratory, or imaging evidence converges	One stream improves while another persists	Concordance strengthens the explanation; asynchrony requires modality-specific timing
Does conflict remain after technical and temporal causes are considered?	A credible source of discordance is identified	Persistent conflict remains unexplained	Condition confidence on the explanation or reopen the differential diagnosis

Note: The sequence organizes inference from serial change; it does not supply analyte-independent thresholds, diagnostic classes, or action scores.

Pretest probability and clinical context

The diagnostic meaning of a finding depends on what was plausible before the finding was obtained. In canine leishmaniasis, analysis of clinical findings using likelihood ratios demonstrated that the same observation can produce different post-test implications depending on pretest probability and local disease prevalence [13]. A test result therefore modifies an existing probability; it does not create diagnostic meaning independently of exposure history, geography, clinical phenotype, and competing diagnoses.

Context may also affect the apparent performance characteristics of the test itself. Bayesian evaluation of *Mycoplasma bovis* serological testing in dairy herds showed that sensitivity could differ with population or infection context rather than remaining a fixed property across all circumstances [14]. This spectrum dependence is especially important when published diagnostic performance is transferred between production systems, disease stages, or populations with materially different case mixes.

Canine *Brucella suis* serology provides a related example. Bayesian latent-class evaluation demonstrated that predictive values changed with expected prevalence and with whether tests were interpreted singly or in combination [15]. Thus, even technically identical results can have different clinical implications when prior risk differs.

The proposed framework consequently places pretest probability upstream of laboratory and imaging interpretation but not outside them. Clinical history, species, breed predisposition, age, epidemiological exposure, disease prevalence, treatment status, and temporal stage establish an initial evidentiary context. Each subsequent result updates that context to a degree justified by its reliability and diagnostic properties. No universal numerical probability threshold is proposed here; the essential requirement is that context remain explicit when apparently discordant findings are reconciled.

Imaging findings as conditional diagnostic evidence

Imaging may appear to provide a direct window on pathology, yet its meaning remains conditional on the clinical question. In canine thoracic radiography, diagnostic accuracy improved when readers received relevant clinical history, demonstrating that identical images can yield different diagnostic conclusions when contextual evidence changes [3]. Imaging therefore contributes conditional evidence rather than an autonomous diagnostic verdict.

Canine pancreatitis illustrates these limitations: ultrasonography contributes useful structural information, but no single imaging feature or modality resolves the diagnosis in every patient [4]. Pretest probability consequently remains relevant before an imaging abnormality is weighted. The same general probability principle demonstrated for clinical findings in canine leishmaniasis—that prior probability modifies the post-test meaning of evidence—can be applied conceptually to imaging without transferring disease-specific likelihood ratios [13].

The proposed framework therefore asks three questions of an imaging finding: what anatomical or physiological phenomenon it actually represents, how reliably that feature can be detected in this patient and setting, and how strongly it modifies the current diagnostic explanation. Incidental, residual, technically uncertain, or context-incongruent findings should remain distinguishable from findings that coherently reinforce the broader evidence pattern.

Reconciling laboratory and imaging evidence

Laboratory and imaging findings should not be expected to agree simply because both are related to the same disease. In dogs evaluated for pancreatitis, abdominal ultrasonographic findings and specific canine pancreatic lipase results showed incomplete correspondence, and changes in one did not consistently mirror changes in the other [16]. Discordance can therefore reflect different biological targets, thresholds, or temporal sensitivities rather than the failure of one modality.

Cardiac biomarkers and point-of-care ultrasonography contributed discriminative information, but biomarker overlap and incomplete sensitivity prevented a single result from replacing the broader cardiac assessment [17]. In canine neurology, serum neurofilament light chain distinguished selected structural brain diseases from comparison groups with useful but imperfect discrimination, supporting structural suspicion without replacing anatomical evaluation [18].

Across these examples, laboratory, biomarker, and imaging evidence provide partially noninterchangeable information [16–18]. The proposed reconciliation principle is therefore functional rather than hierarchical: each modality should be weighted according to what it measures, its reliability, its temporal behavior, and its fit with the current patient context. Concordance can strengthen an explanation, but discordance should first prompt examination of non-equivalent targets and timing before one result is labeled erroneous.

Discordant findings and competing explanations

Discordance is diagnostically informative only if competing explanations remain open. In dogs investigated for cerebrovascular accidents, D-dimer concentration and thromboelastography did not independently discriminate the target diagnosis adequately despite biological plausibility [19]. A plausible mechanistic relationship is therefore insufficient grounds for privileging a laboratory result over stronger disease-defining evidence.

Alternative explanations also require testing rather than assumption. In canine epilepsy, serum neurofilament light chain differed between structural and idiopathic disease groups, while an examined seizure-to-sampling timing contrast within idiopathic epilepsy did not explain the observed pattern [20]. Conversely, disagreement may originate within the imaging process itself. A canine and feline lung-ultrasound reliability study found excellent overall score reliability but weaker agreement for some individual features, showing that global reproducibility does not eliminate feature-level observer uncertainty [21].

The proposed discordance audit separates measurement or handling effects, biological variation, non-equivalent modalities, temporal mismatch, and competing disease. When evidence remains incompatible after plausible technical and temporal causes are considered, diagnostic confidence should decrease or become conditional rather than being rescued by selectively privileging the most convenient result.

A patient-centered evidence-reconciliation framework

The proposed Patient-Centered Evidence-Reconciliation Framework organizes diagnosis around the evolving evidentiary state of an individual animal. Its first boundary is measurement validity: a result should be interpreted as biological evidence only after relevant analytical and preanalytical threats have been considered [1]. The second is patient trajectory, which distinguishes population reference status from change relative to a credible baseline.

The third is contextual probability; clinical history can alter the interpretation of otherwise identical diagnostic material [3].

Beyond these boundaries, laboratory and imaging evidence are assigned roles according to what they measure rather than forced into a fixed hierarchy. The incomplete correspondence between ultrasonography and pancreatic lipase in canine pancreatitis illustrates why multimodal disagreement need not identify a "wrong" test [16]. Negative diagnostic evidence likewise matters: poorly discriminating but biologically plausible biomarkers show why discordance should reopen alternatives instead of being rationalized away [19].

The framework therefore proposes an iterative sequence of validity checking, baseline comparison, probability conditioning, modality-specific interpretation, discordance analysis, confidence assignment, and reassessment. These operations are not a numerical score; they make explicit which inference changes when new evidence arrives and what uncertainty remains.

The architecture is compressed visually because its evidence streams, interpretation boundary, and feedback relationships are difficult to recover from a linear list alone **(Figure 1)**.

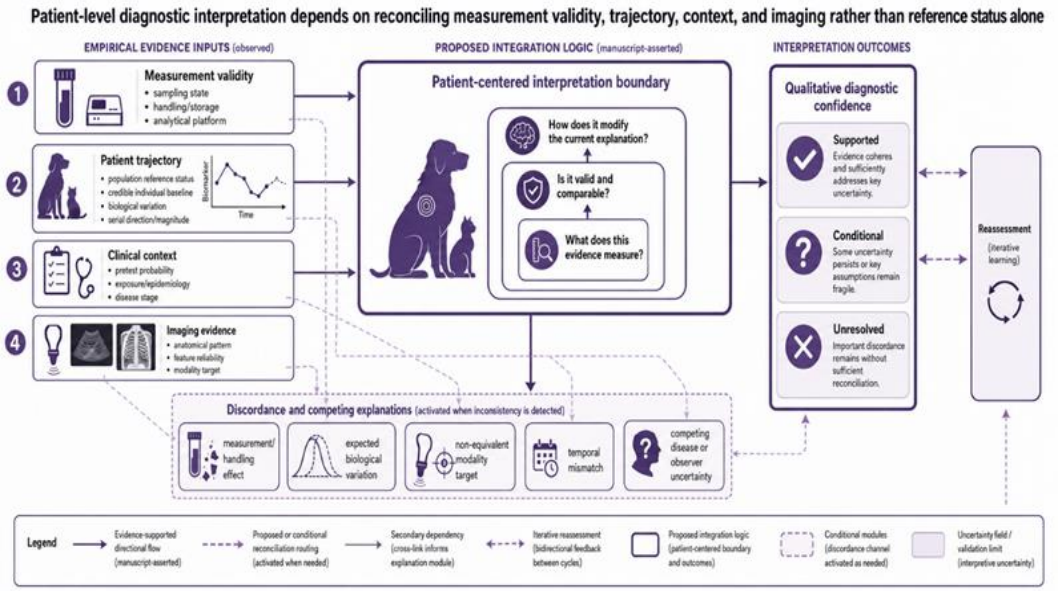

Figure 1. Discordant veterinary evidence crosses distinct validity, context, and reconciliation boundaries before patient-level confidence changes

Table 2 crosses the population reference interval with the patient-relative trajectory, revealing four distinct interpretations that a normal-versus-abnormal label cannot represent.

Table 2. Four interpretations created by crossing population reference status with the patient's own trajectory

Population reference status	Patient trajectory remains compatible with expected variation	Patient trajectory changes beyond expected variation
Inside the reference interval	Population-normal and patient-stable are aligned, although disease is not excluded	A baseline-discordant but population-normal result deserves attention to persistence and corroboration
Outside the reference interval	Individuality, a chronic set point, or stable abnormality may be plausible; clinical context determines significance	Valid, directional change plus population abnormality increases concern, but the cause still requires diagnostic reconciliation

Diagnostic confidence states

Diagnostic confidence should represent the adequacy and coherence of evidence, not the number of abnormal tests. The ACVIM consensus approach to immune thrombocytopenia illustrates a broader veterinary reality: some diagnoses remain dependent on exclusion, and evidentiary support for individual diagnostic steps can vary substantially [22]. The framework therefore proposes qualitative confidence states rather than a binary confirmed/not-confirmed label.

Simulation work comparing latent-class approaches shows that estimated test characteristics can shift with model structure and the handling of individual or cluster-level covariates when no perfect gold standard exists [23].

Threshold choice introduces another layer. In dairy calves assessed for bronchopneumonia, the sensitivity and specificity of an auscultation score changed with the selected cutoff and assumptions of an imperfect-reference model [24].

The proposed states are "supported," "conditional," and "unresolved" evidentiary positions rather than probability bands. Supported evidence is coherent across sufficiently valid streams; conditional evidence is plausible but materially dependent on an unresolved assumption, threshold, or conflicting modality; unresolved evidence leaves competing explanations insufficiently discriminated. These labels require prospective testing for reproducibility and clinical usefulness and should not be interpreted as validated diagnostic classes.

Reassessment and evidence updating

Reassessment is not simple repetition. Its purpose is to determine whether new evidence behaves as expected under the current explanation. In dogs with suspected pancreatitis, serial pancreatic lipase, C-reactive protein, ultrasonographic findings, and clinical severity did not resolve synchronously; sonographic abnormalities could persist after biochemical or clinical improvement [25]. Persistence therefore cannot automatically be equated with persistent active disease, but neither can it be dismissed without context.

The prior probability of disease can also change over time. Repeated health screening in mature and senior cats identified animals who did not remain classifiable as apparently healthy across longitudinal follow-up [26]. In canine cardiogenic pulmonary edema, serial lung ultrasonography and thoracic radiography likewise showed differing resolution patterns during treatment [27].

Together, these studies support time- and modality-aware evidence updating rather than an expectation of simultaneous normalization [25–27]. The proposed framework asks whether each subsequent observation strengthens the existing explanation, weakens it, or reveals a new competing account. Reassessment therefore changes both the evidentiary weight of the new result and the context in which earlier results are understood.

The temporal logic is summarized visually because delayed resolution, measurement drift, new disease, and true progression can generate superficially similar serial discordance **(Figure 2)**.

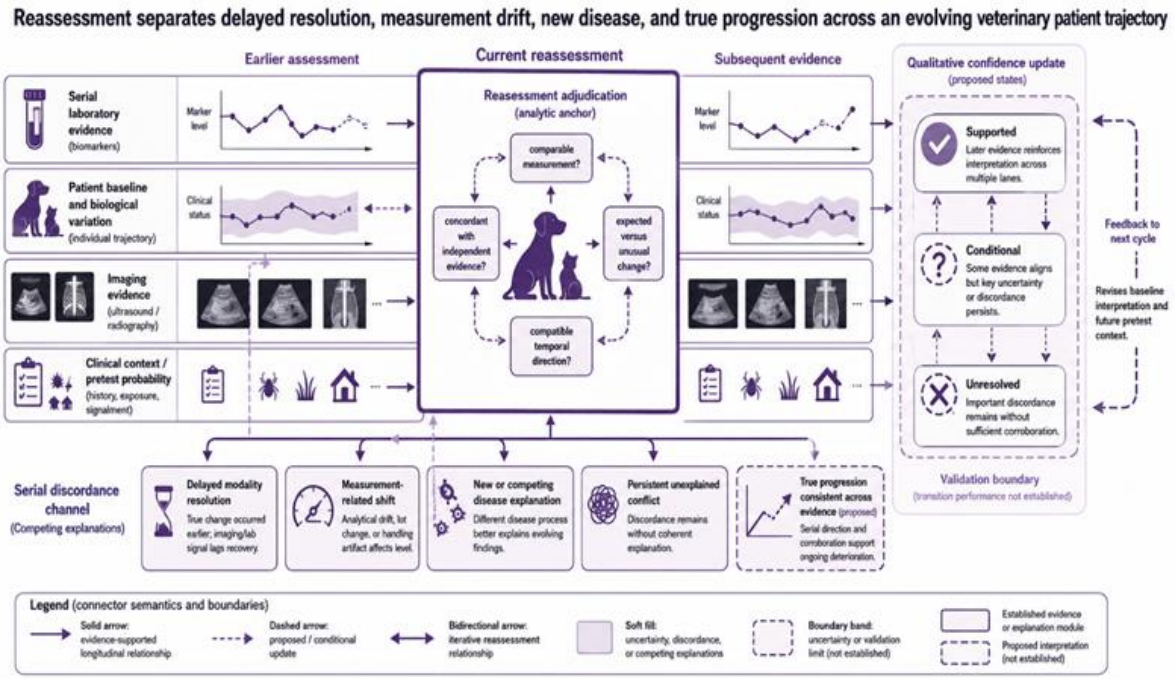

Figure 2. Reassessment separates delayed resolution, measurement drift, new disease, and true progression across an evolving veterinary patient trajectory

Applications to common veterinary diagnostic problems

Proteinuria illustrates the distinction between population abnormality and patient trajectory: serial feline urinary measurements show that an isolated abnormal value may not persist [5]. A clinician can therefore ask whether the current result represents sustained change while still respecting disease-specific thresholds and clinical context.

Pancreatitis illustrates multimodal and temporal discordance. Ultrasonography and pancreatic lipase provide non-equivalent information [16], so disagreement should prompt consideration of disease stage, test characteristics, and alternative explanations rather than automatic selection of one modality as definitive. Cardiogenic pulmonary edema provides a monitoring analogue because lung ultrasound and radiography can change on different timelines [27].

Neurologic disease illustrates a different boundary. Serum neurofilament light chain can increase suspicion of structural disease, but its imperfect discrimination means that a biomarker result cannot substitute automatically for anatomical diagnosis [18]. Across these examples, the framework specifies the reasoning question rather than dictating a single action. Disease-specific evidence, urgency, available expertise, patient stability, owner constraints, and consequences of delayed or unnecessary testing remain clinically decisive.

Limitations and validation requirements

Biological-variation estimates are not automatically reliable; the Biological Variation Data Critical Appraisal Checklist demonstrates that design, sampling, analytical, and statistical quality materially affect whether an estimate is suitable for interpretation [28]. Veterinary application therefore requires analyte- and species-specific appraisal rather than indiscriminate use of published reference-change values.

Personalized reference approaches also require credible historical data. Statistical work on personalized reference intervals emphasizes assumptions concerning steady state, previous observations, and model specification [29]. They cannot be transferred into veterinary decisions without validation, particularly with sparse records, changing physiology, treatment exposure, or intermittent disease.

Transportability is another major boundary. External validation of a canine Cushing's prediction tool showed that discrimination can remain useful while calibration changes after movement from one clinical setting to another [30]. The proposed framework similarly requires prospective testing across species, diseases, primary and referral care, analytical platforms, and prevalence settings. Needed studies include blinded multimodal adjudication, interobserver assessment, longitudinal calibration of updating rules, external replication, and comparison with usual diagnostic reasoning. Clinical utility, welfare consequences, resource use, and downstream outcomes must be evaluated separately rather than inferred from conceptual coherence.

Conclusion

A reference interval is an important population descriptor, but it is not a complete representation of an individual veterinary patient. Diagnostic interpretation becomes more defensible when population reference status is separated from measurement validity, biological variation, patient-specific trajectory, pretest probability, modality-specific evidence, discordance, and time.

The framework developed here proposes that unexpected findings should be reconciled rather than automatically normalized, escalated, or forced into agreement. Serial change should be judged against plausible variation and kinetics; imaging should be interpreted within context; discordance should preserve competing explanations; and reassessment should update both current confidence and the interpretation of earlier evidence. These principles constitute a reasoning architecture, not a validated score or clinical decision rule. It also makes explicit when uncertainty should be preserved, when discordance warrants explanation, and when additional evidence may be more informative than immediate reclassification.

Its principal value is therefore conceptual: making the sources and boundaries of diagnostic inference visible. Whether this approach improves diagnostic accuracy, efficiency, welfare, or outcomes remains unknown and requires disease-, species-, setting-, and modality-specific prospective validation.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Flatland B, Baral RM, Freeman KP. Current and emerging concepts in biological and analytical variation applied in clinical practice. J Vet Intern Med. 2020;34(6):2691-700. doi:10.1111/jvim.15929
2. Fernandes T, Robin M, Freeman KP. Review of biological variation and its applications in interpretation of equine clinical pathology results. Equine Vet J. 2025;57(4):862-9. doi:10.1111/evj.14454
3. Arruda Bergamaschi N, Huber L, Ludewig E, Böhler A, Gumpenberger M, Hittmair KM, et al. Association between clinical history in the radiographic request and diagnostic accuracy of thorax radiographs in dogs: A retrospective case-control study. J Vet Intern Med. 2023;37(6):2453-9. doi:10.1111/jvim.16899
4. Cridge H, Twedt DC, Marolf AJ, Sharkey LC, Steiner JM. Advances in the diagnosis of acute pancreatitis in dogs. J Vet Intern Med. 2021;35(6):2572-87. doi:10.1111/jvim.16292
5. Mortier F, Daminet S, Duchateau L, Biscop A, Paepe D. Biological variation of urinary protein: Creatinine ratio and urine specific gravity in cats. J Vet Intern Med. 2023;37(6):2261-8. doi:10.1111/jvim.16881
6. Couture Y, Keys D, Summers S. Weekly biological variation of urine protein creatinine ratio and urine specific gravity in healthy dogs. J Vet Intern Med. 2025;39(2). doi:10.1111/jvim.70052
7. Yi KC, Heseltine JC, Jeffery ND, Cook AK, Nabity MB. Effect of withholding food versus feeding on creatinine, symmetric dimethylarginine, cholesterol, triglycerides, and other biochemical analytes in 100 healthy dogs. J Vet Intern Med. 2023;37(2):626-34. doi:10.1111/jvim.16630
8. Brans M, Marynissen S, Mortier F, Duchateau L, Daminet S, Paepe D. Effect of storage temperature and time on measurement of serum symmetric dimethylarginine concentration using point-of-care and commercial laboratory analyzers in cats and dogs. J Vet Intern Med. 2023;37(5):1794-805. doi:10.1111/jvim.16811
9. Bonatto NCM, Alves LSS, Silva LE, Milhorine CA, de Barros LD, Santos JA, et al. Does postprandial lipemia interfere with blood gas analysis and assessment of acid-base status in dogs? Res Vet Sci. 2023;154:52-8. doi:10.1016/j.rvsc.2022.11.008
10. Jeffery U, Jeffery ND, Creevy KE, Page R, Simpson MJ. Variation in biochemistry test results between annual wellness visits in apparently healthy Golden Retrievers. J Vet Intern Med. 2021;35(2):912-24. doi:10.1111/jvim.16021
11. Goggs R, Robbins SN, LaLonde-Paul DM, Menard JM. Serial analysis of blood biomarker concentrations in dogs with pneumonia, septic peritonitis, and pyometra. J Vet Intern Med. 2022;36(2):549-64. doi:10.1111/jvim.16374
12. Pedro B, Mavropoulou A, Oyama MA, Linney C, Neves J, Dukes-McEwan J, et al. Longitudinal analysis of echocardiographic and cardiac biomarker variables in dogs with atrial fibrillation: The optimal rate control in dogs with atrial fibrillation II study. J Vet Intern Med. 2024;38(4):2076-88. doi:10.1111/jvim.17120
13. Galán-Relaño Á, Maldonado A, Gómez-Gascón L, Tarradas C, Astorga RJ, Luque I, et al. Pre-test probability and likelihood ratios for clinical findings in canine leishmaniasis. Transbound Emerg Dis. 2022;69(6):3540-7. doi:10.1111/tbed.14717
14. Veldhuis A, Aalberts M, Penterman P, Wever P, van Schaik G. Bayesian diagnostic test evaluation and true prevalence estimation of *Mycoplasma bovis* in dairy herds. Prev Vet Med. 2023;216:105946. doi:10.1016/j.prevetmed.2023.105946
15. Kneipp CC, Coilparampil R, Westman M, Suann M, Robson J, Firestone SM, et al. Evaluation of three serological tests for the diagnosis of *Brucella suis* in dogs using Bayesian latent class analysis. Prev Vet Med. 2024;233:106345. doi:10.1016/j.prevetmed.2024.106345
16. Cridge H, Sullivant AM, Wills RW, Lee AM. Association between abdominal ultrasound findings, the specific canine pancreatic lipase assay, clinical severity indices, and clinical diagnosis in dogs with pancreatitis. J Vet Intern Med. 2020;34(2):636-43. doi:10.1111/jvim.15693
17. Morey AGN, Lamb KE, Karnia JJ, Wiggen KE, Lyons BM, Nafe LA, et al. N-terminal brain natriuretic peptide, cardiac troponin-I, and point-of-care ultrasound in dogs with cardiac and noncardiac causes of nonhemorrhagic ascites. J Vet Intern Med. 2023;37(3):900-9. doi:10.1111/jvim.16702
18. Sung J, Chae Y, Yun T, Koo Y, Lee D, Kim H, et al. Use of neurofilament light chain to identify structural brain diseases in dogs. J Vet Intern Med. 2024;38(4):2196-203. doi:10.1111/jvim.17110
19. DiPaola E, Cameron S, Rylander H, Zidan N, Hetzel S. Comparison of D-dimer concentration and thromboelastography for diagnosis of cerebrovascular accidents in dogs: A retrospective study. J Vet Intern Med. 2024;38(2):1083-91. doi:10.1111/jvim.17000

20. Fowler KM, Shinn RL, Rossmeisl JH, Parker RL. Evaluation of neurofilament light chain as a biomarker in dogs with structural and idiopathic epilepsy. J Vet Intern Med. 2024;38(3):1577-82. doi:10.1111/jvim.17033
21. Oricco S, Medico D, Tommasi I, Bini RM, Rabozzi R. Lung ultrasound score in dogs and cats: A reliability study. J Vet Intern Med. 2024;38(1):336-45. doi:10.1111/jvim.16956
22. LeVine DN, Kidd L, Garden OA, Brooks MB, Goggs R, Kohn B, et al. ACVIM consensus statement on the diagnosis of immune thrombocytopenia in dogs and cats. J Vet Intern Med. 2024;38(4):1958-81. doi:10.1111/jvim.16996
23. Ni H, Koop G, Klugkist I, Nielen M. Evaluation of Bayesian Hui-Walter and logistic regression latent class models to estimate diagnostic test characteristics with simulated data. Prev Vet Med. 2023;217:105972. doi:10.1016/j.prevetmed.2023.105972
24. Boccardo A, Ferraro S, Sala G, Ferrulli V, Pravettoni D, Buczinski S. Bayesian evaluation of the accuracy of a thoracic auscultation scoring system in dairy calves with bronchopneumonia using a standard lung sound nomenclature. J Vet Intern Med. 2023;37(4):1603-13. doi:10.1111/jvim.16798
25. Mitchell L, Wang S, Lawver J, Cridge H. Serial monitoring of pancreatic lipase immunoreactivity, C-reactive protein, abdominal ultrasonography, and clinical severity in dogs with suspected pancreatitis. J Vet Intern Med. 2024;38(2):987-94. doi:10.1111/jvim.17020
26. Mortier F, Daminet S, Marynissen S, Smets P, Paepe D. Value of repeated health screening in 259 apparently healthy mature adult and senior cats followed for 2 years. J Vet Intern Med. 2024;38(4):2089-98. doi:10.1111/jvim.17138
27. Murphy SD, Ward JL, Viall AK, Tropf MA, Walton RL, Fowler JL, et al. Utility of point-of-care lung ultrasound for monitoring cardiogenic pulmonary edema in dogs. J Vet Intern Med. 2021;35(1):68-77. doi:10.1111/jvim.15990
28. Aarsand AK, Røraas T, Fernandez-Calle P, Ricos C, Díaz-Garzón J, Jonker N, et al. The Biological Variation Data Critical Appraisal Checklist: A standard for evaluating studies on biological variation. Clin Chem. 2018;64(3):501-14. doi:10.1373/clinchem.2017.281808
29. Coskun A, Sandberg S, Unsal I, Yavuz FG, Cavusoglu C, Serteser M, et al. Personalized reference intervals—statistical approaches and considerations. Clin Chem Lab Med. 2022;60(4):629-35. doi:10.1515/cclm-2021-1066
30. Ruijter BEW, Bik CA, Schofield I, Niessen SJM. External validation of a United Kingdom primary-care Cushing's prediction tool in a population of referred Dutch dogs. J Vet Intern Med. 2023;37(6):2052-63. doi:10.1111/jvim.16848