Received: 24 September 2019 | Accepted: 16 December 2020

DOI: 10.1111/jvim.16021

STANDARD ARTICLE

Journal of Veterinary Internal Medicine Open Access

ACVIM American College of Veterinary Internal Medicine

Variation in biochemistry test results between annual wellness visits in apparently healthy Golden Retrievers

Unity Jeffery[1] | Nick D. Jeffery[2] | Kate E. Creevy[2] | Rod Page[3] | Melissa J. Simpson[4]

[1]Department of Veterinary Pathobiology, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, Texas

[2]Department of Small Animal Clinical Sciences, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, Texas

[3]Flint animal Cancer Center, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, Colorado

[4]Morris Animal Foundation, Denver, Colorado

Correspondence
Unity Jeffery, Department of Veterinary Pathobiology, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, TX 77843, USA.
Email: ujffery@cvm.tamu.edu

Funding information
Morris Animal Foundation, Grant/Award Number: D18CLP-802

Abstract

Background: Annual wellness testing is widely recommended for apparently healthy dogs, but there is little data to assist with distinguishing normal variation from clinically important changes.

Objectives: To define variability in biochemistry analytes between annual wellness tests in healthy Golden Retrievers.

Animals: Four hundred thirty-four Golden Retrievers undergoing annual health assessments by their primary care veterinarians as part of a prospective cohort study.

Methods: Changes in 23 biochemistry analytes were calculated between year 1 and year 2 health checks for 196 dogs classified as healthy for ≥3 consecutive years. Using a direct nonparametric method, annual change intervals were constructed to define normal variability. A validation cohort of 238 dogs without a diagnosis of systemic disease for ≥3 consecutive years were compared with the reference and annual change intervals, and the proportions of dogs outside annual change intervals and a population-based reference interval were compared by using a McNemar test.

Results: Annual change intervals were calculated based on 190 dogs after outlier removal. For all 23 analytes, >90% of dogs in the validation cohort were within the annual change interval. There were no significant differences in the classification by reference versus annual change intervals.

Conclusions and Clinical Importance: The annual change intervals met performance requirements for classification of dogs that did not develop systemic disease in the year following wellness testing as normal.

KEYWORDS
biologic variability, dog, health screening, reference interval

1 | INTRODUCTION

Routine biochemistry testing of apparently healthy dogs is often recommended but remains controversial.[1-3] In part, this reflects difficulty in interpreting the results because routine wellness testing frequently

Abbreviations: BUN, blood urea nitrogen; CI, confidence interval; RCV, reference change value.

detects values outside reference interval but there is little information regarding whether such changes are clinically important.[2,4,5] Failure to respond to abnormalities could be a missed opportunity for early diagnosis and initiation of treatments that might improve quality of life and longevity.[6,7] On the other hand, overreaction to unimportant alterations in laboratory data might trigger unnecessary diagnostics and interventions,[6] risking iatrogenic injury and increasing veterinary costs, which has been linked to owner dissatisfaction and reduced likelihood of seeking veterinary care.[8-10]

To assist with interpretation of wellness testing, veterinarians are advised to compare sequential annual results to identify trends away from an individual's baseline.[1,11] This approach could be helpful for determining the clinical importance of out-of-reference-interval results and for identification of alterations in analytes that vary more across the population than within a single individual.[12] For these analytes, a subject-based reference interval derived from changes within healthy individuals over time has greater sensitivity than a single-time point population-based reference interval for detecting an alteration in analyte concentration exceeding that expected in normal animals.[12] The most common form of subject-based reference interval is the reference change value (RCV), which defines the boundaries within which an analyte might be expected to vary in a healthy individual owing to analytical and biological variation alone.[12,13] Changes exceeding the RCV are therefore suggestive of a change in health or physiological status (eg, pregnancy).[14-16]

RCVs have been reported for canine biochemistry analytes.[12,17,18] RCVs are usually calculated based on estimates of between- and within-individual variation derived from statistical partitioning of results generated by repeated sampling of a relatively small number of healthy animals over weeks to months.[13,18] Such results are valuable for evaluating serial data from sick individuals generated over similar periods[19-22] and, for many analytes, appear to be generalizable to the longer intervals separating annual wellness samples.[22] However, it is unclear if this is true for all analytes. [22] Research studies might also underestimate the variability in a primary care setting, where veterinarians have little control over factors such as duration of fasting. The Golden Retriever Lifetime Study (GRLS) is a prospective cohort study that involves annual laboratory testing and health assessment of a large number of dogs and therefore offers a unique resource for establishing variability in biochemistry results between wellness assessments for healthy dogs.

The primary aim of this study was to determine reference intervals for variability in biochemistry analytes between annual wellness tests in healthy Golden Retrievers. To avoid confusion between our results and RCVs, the resulting reference intervals are referred to as annual change intervals (ACIs). Secondary aims were to (a) determine if a single time-point population-based reference interval or ACI produced a greater proportion of abnormal results for dogs that remained free from diagnosed systemic disease for 12 months after annual wellness testing, and (b) compare performance of ACIs and published RCVs for classification of dogs that remained free from diagnosed systemic disease for 12 months after annual wellness testing.

2 | MATERIALS AND METHODS

2.1 | Data generation

All data were generated as part of the GRLS with informed owner consent and ethical approval from the Morris Animal Foundation's Animal Welfare Advisory Board. Enrolled dogs underwent an annual health check by their primary care veterinarian, including a full physical examination and submission of blood, serum and urine to a commercial reference laboratory (ANTECH Diagnostics and Imaging) for complete blood count, biochemistry profile, and urinalysis. Methodology and instrumentation are provided in Supplementary Table 1. Owners were requested to fast dogs before blood sampling, and compliance was recorded. Physical exam findings and all prescribed medications and veterinary diagnoses in the previous 12 months were recorded by the veterinarian in an annual electronic questionnaire. Additionally, if the dog was diagnosed with malignant neoplasia, died, or was euthanized, this was reported. Owners also completed extensive annual health questionnaires including reproductive history. More detailed information regarding study methodology has been previously published.[23,24]

2.2 | Selection of dogs for inclusion in annual change interval construction

Initial search of the GRLS database identified 2458 dogs with at least 3 years of data available, to which strict criteria were applied to identify healthy dogs for inclusion in ACI construction. Because wellness testing aims to identify subclinical disease, for inclusion in the calculation of ACIs, dogs had to meet our criteria for classification as healthy in both of the 2 years used to calculate the annual change value and, also, in the following 12 months. Specifically, dogs were considered healthy if for 3 consecutive years:

1. No drugs were prescribed other than routine antiparasite treatment, vaccination and nonprescription supplements.
2. Physical exam was within normal limits. Physical exam findings not considered important are available as supplementary data (Supplementary Table 2).
3. No diagnosis was entered in the health questionnaire.
4. The dog remained alive and enrolled in the study.

To reduce effects of bone growth on analytes such as alkaline phosphatase (ALP) and phosphate, dogs were excluded if <1 year old at the start of the 3-year testing sequence.[25] Bitches were excluded if pregnant in years 1 or 2 of the testing sequence. Owing to requirements for future aspects of this project, dogs were excluded if they developed malignant neoplasia at any point in the study, regardless of whether the dog met inclusion criteria for other available years. For dogs with more than 3 consecutive years of data meeting these inclusion criteria, 1 block of 3 consecutive years for inclusion was selected by random number generator (Random.Org, www.random.org).

2.3 | Annual change and reference interval generation

For each dog included in the reference sample group, annual change values were calculated for each analyte by subtracting the second included year's result from the first included year's result. A percentage annual change was also calculated using the equation [(Year 2's result – Year 1's result)/Year 1's result)]*100. Positive annual change values therefore reflect an increase, and negative annual change values a decrease in the analyte between the year 1 and year 2 health check.

For both year 1 and 2 data, outliers were identified by the Reed method and dogs with 1 or more outlier result in either year were excluded from further analysis for all analytes.[26] Three reference intervals were constructed for each analyte: (a) a conventional population-based reference interval calculated using the year 1 data; (b) a reference interval for the annual change values in analyte-specific units (ie, analyte-specific unit ACI); and (c) a reference interval for the percentage annual change values (ie, percentage ACI). Reference intervals were calculated according to the American Society of Veterinary Clinical Pathology's guidelines using the Reference Value Advisor plugin for Excel.[27,28] Provided that at least 120 reference individuals were available, a nonparametric reference interval encompassing the central 95% of results was calculated, together with 90% confidence intervals (CIs) for the reference limits. For analytes with less than 120 reference individuals, normality of the data distribution was assessed by the Anderson Darling test, and if data were not normally distributed ($P < .05$), Box-Cox transformation was performed and the Anderson Darling test repeated to confirm normality. Reference intervals for analytes with between 40 and 119 reference individuals were then calculated using the robust method with 90% CIs determined by bootstrapping.

2.4 | Validation cohort of dogs without diagnosed systemic disease

ACIs were constructed using only dogs meeting very strict health inclusion criteria, potentially leading to excessively narrow intervals that would lead to out-of-interval results for dogs with minor health conditions but free from the serious systemic diseases that annual wellness testing aims to detect. To determine if ACIs would generate excessive false-positive results for the general population of Golden Retrievers undergoing wellness testing, a validation cohort of dogs were selected using less stringent health criteria. Dogs were eligible for inclusion in the systemically healthy cohort if they were not part of the ACI generation cohort due to 1 or more minor abnormalities but were free from systemic disease, as defined by the following criteria:

1. Not receiving drugs on the day of health check in year 1 or 2, other than routine antiparasite treatment, vaccination, or nonprescription supplements. Prescription drugs could have been prescribed at other times during years 1 and 2.
2. Unremarkable physical exam for 3 consecutive years. Physical exam findings not considered important are available as supplementary data (Supplementary Table 3).
3. No diagnoses entered for 3 consecutive years, other than otitis externa, hot spots, benign skin neoplasia, traumatic skin wounds, iris cyst, distichiasis, miscellaneous other diagnosis entered in the questionnaire but considered trivial (eg, full anal glands, missing dentition) and musculoskeletal disease other than panosteitis, immune-mediated arthritis, masticatory myositis, neoplastic lesions, or metabolic bone disease. Dogs with such diagnoses were included on the basis that they would be unlikely to influence biochemistry data or be a sign of serious systemic illness. Specific diagnoses in dogs included in the systemically healthy cohort are provided (Supplementary Table 4).
4. At least 12 months old at the time of the year 1 health check.
5. Not pregnant in year 1 or 2.
6. Did not develop malignant neoplasia, including in years outside the included 3-year testing block.

For dogs with more than 3 consecutive years of data meeting these inclusion criteria, 1 block of 3 consecutive years for inclusion was selected by a random number generator. Annual changes were calculated as described earlier and were compared to the upper and lower limits of the percentage and analyte-specific unit ACIs. Results of the second year's biochemistry testing were compared with the population-based reference interval. The proportion of dogs within reference interval or ACIs was calculated for each analyte.

We considered that an ACI was suitable for use for healthy Golden Retrievers if no more than 10% of the dogs without a diagnosis of systemic illness had results outside the ACI for that analyte (ie, ≤10% of healthy dogs had a false-positive result). For both the percentage and analyte-specific unit ACIs, the proportion of dogs classified as within the ACI was compared with the proportion of dogs within the population-based reference interval using the McNemar test for paired samples. Bonferroni correction was applied for the 46 comparisons performed, so $P < .0010$ was considered significant. A sample size calculation determined that at least 168 dogs were required to detect a 10% difference in the proportion of dogs outside reference interval vs ACI with 80% power and an alpha of 0.001.

2.5 | Comparison of annual change intervals and published reference change values

Classification of the validation cohort was compared between percentage ACIs and reference change intervals calculated using canine intraindividual coefficients of variations provided in 2 sources.[18,29] Reference change values were calculated using the equation: $RCV = 1.96 \times 2^{0.5} \times \left(CV_I^2 + CV_A^2\right)^{0.5}$, where CV_I is the within-individual coefficient of variation and the CV_A is the analytical variation provided in the original publication. These published CV_A estimates will reflect performance of the analytical methods used in the original studies, and as CV_A will vary between methods and laboratories,

Reference 30 might not be representative of analytical performance in the current study. Analytical imprecision data for the reference laboratory used in the current study are proprietary and thus a laboratory-specific RCV could not be calculated. To provide some estimate of the effect of methodology, a second RCV was calculated using the maximum imprecision reported for human serum or control materials in the datasheet for each of the assays used (Supplementary Table 1).

For each RCV, the proportion of dogs in the validation cohort with a percentage annual change exceeding RCV was compared to the proportion with an annual change exceeding the upper and lower limits of the percentage ACI, using the McNemar test for paired samples. Bonferroni correction was applied for the 25 comparisons performed, resulting in $P < .002$ being considered significant. Statistical tests and sample size calculation were performed using the MedCalc Software version 18.11.6.

3 | RESULTS

3.1 | Demographics

One hundred ninety-six dogs met the criteria for inclusion in reference interval and ACI generation. Physical exam findings that were considered trivial are available as supplementary data (Supplementary Table 2). At baseline, there were 54 intact females, 41 spayed females, 66 intact males, and 35 neutered males. Between the first and second blood draw, 9 females were spayed and 11 males were neutered. Median age at baseline was 2 years (range 1-5 years). Owners reported that dogs had been fasted in 181 of 196 (92%) of the included year 1 samples and 184 of 196 (94%) of the year 2 samples. For year 1 samples collected after fasting, duration of fasting was >12 hours for 86 of 181 dogs, 8 to 12 hours for 59 of 181 dogs, 4 to 8 hours for 20 of 181 dogs, and 2 to 4 hours for 12 of 181 dogs and not recorded for 4 of 181 dogs. For year 2 samples collected after fasting, duration of fasting was >12 hours for 86 of 184 dogs, 8 to 12 hours for 61/184 dogs, 4 to 8 hours for 20 of 184 dogs, and 2 to 4 hours for 17 of 184 dogs.

Two hundred thirty-eight dogs did not meet our strict inclusion criteria for reference interval and ACI generation but did meet our less rigorous inclusion criteria for classification as "free from diagnosed systemic disease for 3 consecutive years". For this validation cohort, diagnoses or physical exam abnormalities considered not to be evidence of systemic disease are available as supplementary data (Supplementary Tables 3 and 4). At baseline, there were 36 of 238 intact females, 65 of 238 spayed females, 67 of 238 intact males, and 70 of 238 neutered males. Neutering status did not change between year 1 and year 2 in any dog. Median age at baseline was 2 years (range 1-5 years). Owners reported that dogs had been fasted in 228 of 238 (96%) of the included year 1 samples and 223 of 238 (94%) of the year 2 samples. For year 1 samples collected after fasting, duration of fasting was >12 hours for 122/228 dogs, 8 to 12 hours for 67 of 228 dogs, 4 to 8 hours for 18 of 228 dogs, and 2 to 4 hours for 19 of 228 dogs and not recorded for 2 of 228 dogs. For year 2 samples collected after fasting, duration of fasting was >12 hours for 107 of 223 dogs, 8 to 12 hours for 75 of 223 dogs, 4 to 8 hours for 21 of 223 dogs, and 2 to 4 hours for 20 of 223 dogs. Statistical comparison of characteristics of the 2 cohorts is provided (Supplementary Table 5).

3.2 | Population and annual change reference intervals

Outlier detection identified 6 of 196 dogs with at least 1 extreme result by the Reed method, and these dogs were excluded from reference interval and ACI construction for all analytes. For the remaining 190 dogs, there was at least 1 dog with missing data in 1 or both years for 16 of 23 analytes. Reasons for missing data were not recoverable for individual dogs. For 22 of 23 analytes, the number of dogs with available data exceeded the minimum of 120 needed for nonparametric reference and annual change interval construction. For lipase, 83 dogs had results in year 1 and 82 in year 2. The reference interval and ACIs for lipase were therefore calculated using the robust method after Box-Cox transformation ($P \geq .672$ for symmetry tests after Box-Cox transformation). Population-based reference intervals (Table 1) and ACIs (Table 2) are provided. Histograms for each analyte are supplied (Supplementary Figures 1, 2, 3). Comparison of the Golden Retriever population-based reference interval calculated here and the reference laboratory's general population-based reference interval is provided (Supplementary Table 6).

3.3 | Classification of dogs in the validation cohort

Of the 238 dogs that were not included in reference or ACI generation but had at least 3 consecutive years of data without a diagnosis of systemic disease, 173 had at least 1 missing biochemistry value. Reasons for missing data for individual dogs were not recoverable. Additionally, 1 dog was excluded from analysis of calcium and potassium because of suspected data entry error or EDTA contamination (0.5 mg/dL for total calcium, 18.1 mmol for potassium). The sample size requirement for comparison of the proportion of dogs within the population-based reference interval vs ACI was exceeded for all analytes except lipase, for which data were available from only 65 dogs.

For all analytes, >90% of the dogs were within the reference interval and ACIs, meeting our pre-specified performance requirement for correctly classifying as normal at least 90% of dogs that did not develop diagnosed systemic diseases in the year following wellness testing. The proportion of dogs with year 2 results within the population-based interval was not different from the proportion within the analyte-specific unit or percentage ACI for any analyte using the multiplicity corrected alpha of 0.0010 (Table 3).

For 12 of 25 RCVs calculated using published biological and analytical variation estimates, there was no difference between the

TABLE 1 Population-based reference interval

Analyte	N	Median (range)	Lower limit (90% CI)	Upper limit (90% CI)
Albumin (g/dL)	189/190*	3.7 (3.2, 4.4)	3.2 (3.2, 3.3)	4.1 (4.0, 4.4)
ALP (U/L)	188/190*	24 (8, 199)	9 (8, 10)	65 (53, 199)
ALT (U/L)	190/190	30 (13, 242)	18 (16, 19)	100 (62, 126)
Amylase (U/L)	189/190*	529 (255, 1670)	321 (255, 338)	1195 (1076, 1670)
AST (U/L)	190/190	24 (8, 53)	15 (14, 17)	40 (38, 47)
BUN (mg/dL)	190/190	16 (8, 34)	10 (9,10)	25 (22,29)
Calcium (mg/dL)	189/190*	10.2 (8.1, 11.2)	9.4 (8.1, 9.6)	11.0 (10.7, 11.2)
Chloride (mmol)	189/190*	114 (108, 120)	110 (108, 111)	118 (117, 120)
Cholesterol (mg/dL)	189/190*	258 (133, 469)	164 (133, 173)	404 (382, 469)
Creatine kinase (U/L)	190/190	81 (13, 427)	35 (30, 44)	288 (201, 374)
Creatinine (mg/dL)	190/190	1.0 (0.0, 1.4)	0.7 (0.0, 0.8)	1.3 (1.3, 1.4)
GGT (U/L)	190/190	5 (1, 12)	2 (1, 3)	9 (9, 12)
Globulin (g/dL)	189/190*	2.7 (2, 3.6)	2.2 (2.0, 2.3)	3.4 (3.2, 3.6)
Glucose (mg/dL)	189/190*	90 (52, 123)	66 (52, 73)	110 (106, 123)
Lipase (U/L)	83/190*	230 (64, 707)	66 (54, 80)^	769 (635, 907)^
Magnesium (mg/dL)	189/190*	1.7 (1.3, 2.5)	1.5 (1.3, 1.5)	1.9 (1.9, 2.5)
Potassium (mmol)	189/190*	4.3 (3.7, 5.1)	3.8 (3.7, 3.9)	4.9 (4.8, 5.1)
Phosphate (mg/dL)	189/190*	3.5 (1.7, 6.5)	2.3 (1.7, 2.5)	4.8 (4.6, 6.5)
Sodium (mmol)	189/190*	148 (143, 154)	144 (143, 145)	152 (151, 154)
Total bilirubin (mg/dL)	189/190*	0.2 (0.1, 0.4)	0.1 (0.1, 0.1)	0.3 (0.3, 0.4)
Total protein (g/dL)	189/190*	6.3 (5.4, 7.4)	5.7 (5.4, 5.8)	6.9 (6.8, 7.4)
Total T4 (μg/dL)	190/190	1.7 (0.5, 3.7)	0.7 (0.7, 0.8)	2.8 (2.8, 3.2)
Triglycerides (mg/dL)	189/190*	47 (22, 174)	29 (22, 31)	131 (122, 174)

Note: * indicates data missing from at least one reference individual. Reasons for missing data could not be retrieved for individual dogs; ^indicates calculation using the robust method with 90% CI after Box-Cox transformation. All other reference intervals are nonparametric. CI, confidence interval.

proportion of the validation cohort with a change exceeding RCV and the proportion exceeding the ACI. For 1 of 25 RCVs, a higher proportion of the apparently healthy validation cohort were within the RCV than the ACI, and for 12 of 25, a higher proportion were within the ACI (Table 4). When the RCV was calculated using published biological variation estimates and the manufacturer reported analytical variation for the assays used in the current study, more apparently healthy validation dogs were outside RCV than the ACI for 14/25 RCVs. For the remaining 11 of 25 RCVs, there was no difference between the proportion of the validation cohort outside RCV and ACI (Table 5).

4 | DISCUSSION

This study defines annual variability in biochemistry testing results in a group of Golden Retrievers assessed as healthy by their owners and primary care veterinarians. To assist with interpretation of annual wellness testing, we have calculated ACIs using a direct nonparametric method. For an individual dog, the difference between 2 consecutive wellness testing results can be compared with the ACI for the analyte. Differences that fall above or below the ACI are greater than the changes observed in approximately 95% of the healthy Golden Retrievers included in this study, and thus could be evidence of a change in health status.

It should be emphasized that ACIs are a form of reference interval rather than a clinical decision limit, so an out-of-interval result does not definitively indicate disease or necessarily require diagnostic or therapeutic interventions. The direct nonparametric approach defines the central 95% of results for the population and so approximately 5% of healthy animals fall outside the ACIs, even when they are applied to results generated for dogs with very similar characteristics to our study population.[28] Therefore, as with single-time point population-based reference intervals, the clinical relevance of an out-of-interval result should be evaluated based on the magnitude of the difference between the dog's result and the reference limit, plus the overall clinical picture and information about husbandry and physiological alterations such as pregnancy.[31]

Misclassification of healthy dogs as abnormal or diseased dogs as normal is likely to increase if the ACIs are applied to results generated by another laboratory.[28,32] The statistical approach used for ACI generation does not attempt to separate preanalytical and analytical error from homeostatic variation. ACIs are therefore an estimate of the

TABLE 2 Analyte-specific unit and percentage annual change intervals

Analyte	N	Annual change interval (analyte-specific units) Median (range)	Lower limit (90% CI)	Upper limit (90% CI)	Annual change interval (%) Median (range)	Lower limit (90% CI)	Upper limit (90% CI)
Albumin (g/dL)	189/190*	0.0 (−0.8, 0.6)	−0.4 (−0.8, −0.3)	0.5 (0.3, 0.6)	0 (−20, 18)	−11 (−20, −8)	14 (10, 18)
ALP (U/L)	188/190*	−3.0 (−43, 90)	−30 (−43, −20)	19 (13, 90)	−13 (−61, 163)	−52 (−61, −48)	74 (43, 163)
ALT (U/L)	190/190	−1 (−119, 170)	−46 (−102, −18)	59 (22, 92)	−3 (−78, 271)	−60 (−70, −39)	149 (67, 264)
Amylase (U/L)	189/190*	−9 (−567, 1071)	−263 (−567, −217)	397 (269, 1071)	−1 (−52, 179)	−39 (−52, −34)	64 (44, 179)
AST (U/L)	190/190	0 (−31, 26)	−19 (−27, −12)	12 (10, 17)	0 (−70, 123)	−47 (−55, −36)	55 (50, 94)
BUN (mg/dL)	190/190	0 (−12, 12)	−8 (−9, −7)	6 (5, 8)	0 (−50, 64)	−40 (−48, −33)	49 (42, 58)
Calcium (mg/dL)	189/190*	0.0 (−1.8, 1.0)	−1.0 (−1.8, −0.8)	0.7 (0.5, 1.0)	0 (−18, 10)	−9 (−18, −7)	7 (5, 10)
Chloride (mmol)	189/190*	0 (−7, 7)	−4 (−7, −4)	5 (3, 7)	0 (−6, 6)	−4 (−6, −3)	5 (3, 6)
Cholesterol (mg/dL)	189/190*	7 (−157, 166)	−86 (−157, −75)	95 (88, 166)	3 (−36, 83)	−30 (−36, −25)	43 (33, 83)
Creatine kinase (U/L)	190/190	−1.5 (−438, 259)	−156 (−393, −92)	137 (97, 244)	−1.4 (−91, 323)	−73 (−82, −61)	179 (108, 213)
Creatinine (mg/dL)	190/190	0.0 (−1.3, 0.3)	−0.3 (−1.2, −0.2)	0.3 (0.2, 0.3)	0 (−100, 43)	−28 (−100, −21)	28 (22, 33)
GGT (U/L)	190/190	0 (−9, 8)	−5 (−7, −3)	6 (4, 7)	0 (−83, 600)	−66 (−80, −50)	323 (150, 500)
Globulin (g/dL)	189/190*	0.1 (−0.7, 0.9)	−0.5 (−0.7, −0.4)	0.8 (0.6, 0.9)	4 (−23, 45)	−16 (−23, −12)	34 (27, 45)
Glucose (mg/dL)	189/190*	−2 (−36, 41)	−26 (−36, −20)	25 (17, 40)	−2 (−38, 84)	−26 (−38, −20)	32 (21, 84)
Lipase (U/L)	82/190*	12 (−314, 349)	−190 (−225, −145)^	214 (170, 251)^	5 (−68, 306)	−46 (−52, −38)^	128 (87, 292)^
Magnesium (mg/dL)	189/190*	0.0 (−0.3, 1.0)	−0.2 (−0.3, −0.1)	0.3 (0.2, 1.0)	0 (−16, 67)	−11 (−16, −7)	22 (14, 67)
Potassium (mmol)	189/190*	0 (−0.8, 0.9)	−0.6 (−0.8, −0.5)	0.5 (0.4, 0.9)	0 (−16, 22)	−14 (−16, −11)	13 (9, 22)
Phosphate (mg/dL)	189/190*	−0.3 (−2.2, 2.8)	−1.9 (−2.2, −1.5)	1.3 (0.7, 2.8)	−7 (−38, 76)	−36 (−38, −33)	41 (23, 76)
Sodium (mmol)	189/190*	0 (−5, 9)	−4 (−5, −3)	4 (4, 9)	0 (−3, 6)	−3 (−3, −2)	3 (3, 6)
Total bilirubin (mg/dL)	189/190*	0 (−0.2, 0.2)	−0.1 (−0.2, −0.1)	0.1 (0.1, 0.2)	0 (−67, 200)	−50 (−67, −33)	100 (100, 200)
Total protein (g/dL)	189/190*	0.1 (−0.8, 1.2)	−0.6 (−0.8, −0.4)	0.9 (0.7, 1.2)	2 (−12, 21)	−9 (−12, −6)	16 (12, 21)
Total T4 (μg/dL)	190/190	0 (−1.6, 2.4)	−1.1 (−1.3, −0.9)	1.3 (1.0, 1.6)	0 (−88, 200)	−53 (−62, −47)	126 (82, 185)
Triglycerides (mg/dL)	189/190*	1 (−227, 126)	−60 (−227, −24)	80 (60, 126)	2 (−79, 286)	−50 (−79, −36)	183 (129, 286)

Note: Annual changes were calculated by subtracting the result for year 2 from the result for year 1, so positive changes are an increase and negative changes a decrease in analyte concentrations between annual health checks. * indicates data missing from at least one reference individual. Reasons for missing data could not be retrieved for individual dogs; ^indicates calculation using the robust method with 90% CI after Box-Cox transformation. CI, confidence interval.

TABLE 3 Classification of the validation cohort by the population-based reference interval and annual change intervals

Analyte (N)	Population-based reference interval		Annual change interval (analyte-specific units)			Annual change interval (%)		
	Median (range)	Proportion within RI (%)	Median (range)	Proportion within ACI (%)	*P* value	Median (range)	Proportion within ACI (%)	*P* value
Albumin (234)	3.6 (2.9, 4.3)	95.3	−0.1 (−0.6, 1.0)	97.0	.39	−3 (−17, 42)	94.9	1.0
ALP (231)	25 (7, 145)	94.4	−3 (−84, 97)	91.8	.21	−12 (−65, 363)	92.6	.56
ALT (238)	29 (6, 183)	96.6	−1 (−296, 146)	97.5	.76	−4 (−92, 395)	95.8	.79
Amylase (234)	541.5 (243, 2154)	94.0	−17 (−557, 1887)	95.7	.51	−3 (−78, 207)	97.4	.04
AST (238)	24 (6, 86)	95.0	−1 (−52, 58)	96.6	.49	−4 (−78, 207)	95.0	1.0
BUN (238)	15 (8, 35)	94.5	−1 (−14, 15)	91.2	.17	−5 (−57, 156)	90.8	.12
Calcium (233)	10.1 (9.4, 11.3)	99.1	−0.2 (1.0, −1.2)	96.6	.03	−2 (−11, 11)	96.1	.01
Chloride (234)	114 (104, 121)	96.6	0 (−9, 7)	96.6	1.0	0 (−7, 6)	96.6	1.0
Cholesterol (234)	272 (155, 573)	94.9	2 (−210, 165)	93.6	.65	1 (−50, 94)	95.3	1.0
Creatine kinase (238)	79 (4, 736)	95.8	−3 (−1014, 677)	95.0	.79	−3 (−95, 1147)	94.5	.63
Creatinine (238)	1 (0, 1.6)	95.4	0.0 (−1.2, 0.3)	97.9	.07	0 (−100, 43)	95.0	1.0
GGT (238)	5 (1, 10)	99.2	0 (−5, 7)	99.6	1.0	0 (−80, 400)	98.7	1.0
Globulin (234)	2.6 (1.9, 3.7)	96.6	0.1 (−0.7, 0.9)	98.7	.18	4 (−22, 44)	97.0	1.0
Glucose (234)	91 (53, 118)	96.6	0 (−36, 62)	97.0	1.00	0 (−40, 214)	96.6	1.00
Lipase (65)	233 (59, 676)	98.5	16 (−298, 499)	93.8	.38	8 (−39, 330)	95.4	.63
Magnesium (234)	1.7 (1.4, 3.1)	97.9	0.0 (−0.3, 1.5)	97.4	1.0	0 (17, 94)	95.3	.18
Potassium (233)	4.4 (3.5, 5.4)	97.0	0.0 (−0.9, 0.7)	97.4	1.0	0 (−18, 17)	96.6	1.00
Phosphate (234)	3.6 (1.6, 6.2)	94.4	−0.3 (−4.9, 3.4)	96.6	.27	−8 (−58, 121)	93.2	.58
Sodium (234)	148 (125, 155)	95.7	0 (−22, 9)	94.0	.39	0 (−15, 6)	94.0	.39
Total bilirubin (234)	0.2 (0.1, 0.5)	98.7	0.0 (−0.2, 0.2)	98.3	1.0	0 (−67, 200)	98.7	1.0
Total protein (234)	6.3 (5.3, 7.3)	92.3	0.1 (−0.7, 1.3)	98.7	.001	1 (−10, 25)	97.0	.04
Total T4 (238)	1.6 (0.49, 3.7)	96.2	−0.1 (−1.9, 1.3)	97.5	.61	−5 (−61, 120)	98.3	.23
Triglycerides (234)	46 (10, 151)	95.7	1.0 (−128, 106)	97	.59	2 (−76, 354)	96.2	1.0

Note: For the 238 dogs without a diagnosis of systemic disease for 3 consecutive years, year 2 results were compared with the population-based reference interval and the change between year 2 and year 1 compared with the annual change intervals. The proportion of dogs classified as within reference interval was compared with the proportion within each annual change interval by the McNemar test with $P < .001$ considered significant. ACI, annual change interval; RI, reference interval.

total, rather than physiologic, variation between annual wellness test results in healthy dogs. This is a major difference from most RCV studies, which typically provide separate estimates of physiologic within-individual variation and analytical variation introduced by imprecision in laboratory techniques.[12] Analytical performance varies between different laboratory methods and instruments, and even between laboratories using the same instrumentation and methodology.[30,33] Partitioning variation allows recalculation of the RCV using an estimate of analytical variation for the specific laboratory, reagent, and instrumentation used for clinical samples.[12] Our ACIs cannot be recalculated using laboratory-specific data, and instead it would be necessary to perform a transference study to determine if the ACIs can be used to interpret results from other laboratories.[32]

Although this is a weakness of our approach, it might be more practical for veterinarians to perform a transference study than to recalculate a laboratory-specific RCV. In its simplest form, a transference study can be performed using 20 normal animals, with a previously established interval considered suitable for use by another laboratory if results of no more than 2 animals fall outside the limits of the interval.[28] Veterinarians routinely performing wellness testing at a single laboratory could likely relatively easily identify 20 healthy animals with 2 serial results. In contrast, veterinarians might not be able to obtain proprietary information about analytical performance necessary for recalculation of the RCV for their reference laboratory. Neither transference studies nor RCV re-calculation would resolve difficulties in comparing results generated in multiple different laboratories, but this might be less relevant to wellness testing performed by a single primary care provider, compared to monitoring of sick dogs tested at their primary veterinarian, emergency and referral centers.

Transference studies are also needed to determine if our ACIs are suitable for use in other groups of dogs. As our study population

TABLE 4 Comparison of the proportion of the validation cohort with annual changes exceeding the percentage annual change value (ACV) and published reference change values (RCV)

Analyte	Sources	CVi	CVa (published)	RCV (%)	ACV (%)	Number of validation cohort with increase > RCV	Number of validation cohort with decrease > RCV	Proportion outside RCV (%)	Proportion outside ACV	*P* value
Glucose (234)	A	10.7	3.8	31.47372	−26 to 32	4	3	3.0	3.4	1.0
	B	9.5	3.7	28.25937		4	3	3.0	3.4	1.0
Phosphate (234)	A	12.7	4.4	37.25547	−36 to 41	4	12	6.8	6.8	1.0
BUN (238)	A	13.1	5.8	39.71117	−40 to 49	18	6	10.1	9.2	.5
	B	16.1	3.8	45.85311		17	2	8.0	9.2	.38
Calcium (233)	A	1.2	3.9	11.31041	−9 to 7	0	0	0	3.9	.003
Creatinine (238)	A	6.6	7.6	27.90093	−28 to 28	7	5	5.0	5	1.0
	B	14.6	2.9	41.25975		2	4	2.5	5	.03
Cholesterol (234)	A	5.7	4.3	19.79115	−30 to 43	25	8	14.1	4.7	**<.0001**
	B	7.3	3.0	21.87662		21	5	11.1	4.7	**.0001**
Albumin (234)	A	5.8	4.5	20.34816	−11 to 14	1	0	0.4	5.1	**.001**
	B	2.4	1.6	7.995263		23	26	20.9	5.1	**<.0001**
ALT (238)	A	19.5	5.8	56.39149	−60 to 149	15	8	9.7	4.2	**.0002**
	B	9.7	3.2	28.31233		26	30	23.5	4.2	**<.0001**
AST (238)	A	16.3	8.9	51.47753	−47 to 55	8	6	5.9	5.0	.5
	B	11.4	3.3	32.89648		21	17	16.0	5.0	**<.0001**
ALP (231)	A	13.4	7.6	42.70102	−52 to 74	16	22	16.5	7.4	**<.0001**
	B	8.6	1.7	24.29926		23	74	42.0	7.4	**<.0001**
GGT (238)	A	17.8	31.3	99.80731	−66 to 323	19	0	8.0	1.3	**.0001**
Total protein (234)	A	5.3	4.0	18.40522	−9 to 16	2	0	0.9	3.0	.07
	B	2.6	1.1	7.825286		28	6	14.5	3.0	**<.0001**
Potassium (233)	B	3.3	0.1	9.151332	−14 to 13	21	19	17.2	3.4	**<.0001**
Triglycerides (234)	A	35.8	4.0	99.85003	−50 to 183	9	0	3.8	3.8	1.0
Total bilirubin (234)	A	27.5	27.5	107.8	−50 to 100	1	0	0.4	1.3	.5
Total thyroxine (238)	B	17.0	4.0	48.40843	−53 to 183	25	8	13.9	1.7	**<.0001**

Note: Reference change values were calculated from intraindividual (CVi) and analytical (CV_A) coefficients of variation reported for dogs in two published sources (A: Ruaux et al, 2012[18]; B: reported in Jensen and Kjelgaard-Hansen, 2008[29]). The proportion of systemically healthy dogs in the validation cohort with changes exceeding RCV was compared with the proportion exceeding ACV using a McNemar test. After Bonferroni correction, $P < .002$ was considered significant (bold).

TABLE 5 Comparison of the proportion of the validation cohort with annual changes exceeding the percentage annual change value (ACV) and reference change values (RCV) calculated using assay manufacture-reported analytical variation

Analyte	Author	CVi	CVa (manufacturer)	RCV (%)	ACV (%)	Number of validation cohort with increase > RCV	Number of validation cohort with decrease > RCV	Proportion outside RCV (%)	Proportion outside ACV	Significance
Glucose (234)	A	10.7	1	29.78813	−26 to 32	4	3	3.0	3.4	1.0
	B	9.5	1	26.47814		5	4	3.8	3.4	1.0
Phosphate (234)	A	12.7	2.1	35.68061	−36 to 41	4	13	7.3	6.8	1.0
BUN (238)	A	13.1	2.5	36.96666	−40 to 49	20	8	11.8	9.2	.03
	B	16.1	2.5	45.16173		17	2	8.0	9.2	.37
Calcium (233)	A	1.2	1.3	4.903918	−9 to 7	13	32	19.3	3.9	**<.0001**
Creatinine (238)	A	6.6	3.7	20.97292	−28 to 28	19	12	13.0	5	**<.0001**
	B	14.6	3.7	41.74846		2	4	2.5	5	.03
Cholesterol (234)	A	5.7	1.1	16.09111	−30 to 43	34	18	22.2	4.7	**<.0001**
	B	7.3	1.1	20.463		24	8	13.7	4.7	**<.0001**
Albumin (234)	A	5.8	1.5	16.60572	−11 to 14	2	2	1.7	5.1	.007
	B	2.4	1.5	7.844898		24	28	22.2	5.1	**<.0001**
ALT (238)	A	19.5	3.8	55.06798	−60 to 149	15	8	9.7	4.2	**.0002**
	B	9.7	3.8	28.87659		24	29	22.3	4.2	**<.0001**
AST (238)	A	16.3	3.7	46.33069	−47 to 55	11	7	7.6	5.0	.03
	B	11.4	3.7	33.22186		21	17	16.0	5.0	**<.0001**
ALP (231)	A	13.4	1.5	37.37489	−52 to 74	19	29	20.8	7.4	**<.0001**
	B	8.6	1.5	24.19787		23	74	42.0	7.4	**<.0001**
GGT (238)	A	17.8	1.1	49.43321	−66 to 323	40	17	23.9	1.3	**<.0001**
Total protein (234)	A	5.3	0.8	14.85727	−9 to 16	4	0	1.7	3.0	.25
	B	2.6	0.8	7.540271		29	6	15.0	3.0	**<.0001**
Potassium (233)	B	3.3	1.1	9.641925	−14 to 13	15	17	13.7	3.4	**<.0001**
Triglycerides (234)	A	35.8	1.7	99.34436	−50 to 183	9	0	3.8	3.8	1.0
Total bilirubin (234)	A	27.5	3.3	76.77298	−50 to 100	16	0	6.8	1.3	.002
Total thyroxine (238)	B	17.0	7.4	51.39238	−53 to 183	19	6	10.5	1.7	**<.0001**

Note: Reference change values were calculated from intraindividual coefficient of variation (CVi) reported for dogs in two published sources (A: Ruaux et al, 2012[18]; B: reported in Jensen & Kjelgaard-Hansen, 2008[29]), and a coefficient of variation (CV_A) obtained from the manufacturers' datasheet for each assay used in the current study (Supplementary Table 1). The proportion of systemically healthy dogs in the validation cohort with changes exceeding RCV was compared with the proportion exceeding ACV using a McNemar test. After Bonferroni correction, $P < .002$ was considered significant (bold).

includes dogs from across the United States belonging to many different owners, there was likely considerable variation in husbandry and environmental exposures, so our results might not be directly applicable to dogs with specific exposures known to influence biochemistry parameters (eg, raw diets, high intensity exercise).[34-36] The current study is also restricted to Golden Retrievers, most of whom were young adults. Human data suggest that biological variation is similar between young and elderly subjects, but this has not been confirmed in dogs.[37,38] As the GRLS participants age, it will be possible to determine if these intervals appropriately reflect variability in geriatric dogs but, until these data are available, caution should be used if applying these intervals to older dogs. Similar transference studies will also be necessary before applying our ACIs to sequential results generated for other breeds and dogs <1 year old.[25,28,39] It will also be important to determine if the ACIs are suitable for use when different sample handling conditions apply. Our specimens were shipped to the commercial reference laboratory, potentially increasing variability of labile analytes compared with samples analyzed immediately after collection.[40,41]

Our ACIs are intended to be applied to apparently healthy dogs, rather than for monitoring animals with a previous diagnosis of disease. Individuals with stable disease show greater within-individual variation than healthy individuals for some but not all analytes and wider ACIs might therefore be required for monitoring disease progression.[42] Our inclusion requirements for defining dogs as healthy for ACI construction were rigorous, including 3 years of study data with no recorded diagnoses, no medications, or abnormal physical exam findings. One concern is that by using these restrictive inclusion criteria, we could have developed ACIs that were too narrow for a general population of apparently healthy dogs undergoing annual wellness testing.[43] Inclusion of only 1 pair of results from each dog could also have led to underestimation of within-individual variation. We therefore identified a validation cohort of dogs that were not receiving medication on the days of blood collection and were considered systemically healthy by their owners and primary care veterinarians at each health check in a 3-year period, but had been diagnosed with minor health problems such as otitis externa or traumatic injuries during this time. It should also be noted that for inclusion in the validation cohort, there was no requirement for a specific drug free interval, and it is possible that, for example, recent steroid administration might have influenced laboratory results for some included dogs. For all analytes, less than 10% of these dogs were outside the ACIs, which we considered evidence that our ACIs are not too narrow for application to a more general population of apparently systemically healthy Golden Retrievers undergoing wellness testing.

The use of published RCVs or RCVs recalculated using manufacturer-derived estimates of analytical precision frequently resulted in a higher proportion of the healthy validation cohort being classified as abnormal, compared with classification by the percentage ACIs. Poorer performance for RCVs likely reflects, at least in part, that RCVs could not be recalculated using laboratory-specific analytical coefficients of variation because confidential proprietary imprecision data were not available. The influence of analytical variation estimation on RCV is emphasized by the high proportion of the validation cohort outside many of the RCVs derived using imprecision reported by the assay manufacturer, probably reflecting that imprecision provided in assay datasheets underestimates long-term imprecision in a clinical setting. With the caveat that the RCVs used here are not specific for the testing laboratory, comparison of RCVs and the ACIs does identify 2 interesting observations. Firstly, there is consistency between our results and published RCVs for some analytes. For example, creatinine has an annual change interval that is almost identical to 2 published RCVs, suggesting that this analyte show similar short- and long-term variabilty.[18,44] Secondly, for some RCVs (eg, cholesterol, BUN, total thyroxine), there is a marked difference in the proportion of the validation cohort dogs with an increase vs a decrease that exceeds RCV. This could reflect that published canine studies typically use a nested analysis of variance to calculate RCVs for skewed datasets.[45]

Determining performance in the validation cohort will identify intervals that are too narrow and so misclassify an excessive number of healthy dogs as abnormal, but cannot reliably detect intervals that are too wide and so misclassify diseased dogs as normal. Excessively wide intervals can be generated when diseased animals are inadvertently included in the population used for interval construction.[46] Defining an animal as healthy is always a challenge, and this is particularly true for intervals that will be applied to wellness testing, where the aim is to detect subclinical or early disease rather than overt illness.[47] We tried to minimize the potential for including unhealthy animals in reference interval and ACI construction by requiring dogs to be free from disease both in the 2 years used to calculate the annual change value and in the following year. However, as this was based on assessment by a large number of owners and primary care veterinarians, there is likely to be variability in the detail entered in questionnaires or in the criteria used to make diagnoses. In a busy practice setting, it is also possible that suboptimal sample handling could have introduced artifacts for some samples, and self-reported data revealed variability in owner compliance with fasting.

To maximize the generalizability of our study, we wanted to incorporate this variability expected for samples collected in a practice setting but to minimize the likelihood of inadvertent inclusion of sick dogs or the effect of serious technical or data recording errors (eg, failure to separate serum before mailing).[41,46,48] We therefore employed statistical outlier detection to eliminate dogs with 1 or more extreme value from reference interval and annual change interval generation. This is sometimes discouraged when performing a tightly controlled reference interval generation study because it can lead to underestimation of the true variability in normal animals, but is generally accepted in settings such as this, where data were not specifically generated for reference interval construction.[28,46] However, it should be noted that statistical detection alone cannot completely eliminate the risk of including unhealthy animals, particularly as we used a conservative method for outlier detection that favors data retention.[26,28]

We used this method, rather than the more rigorous Horn's algorithm using Tukey's interquartile fences because it is relatively insensitive to data distribution and for several of our analytes, Gaussian distribution was not achieved even after data transformation.[26,49]

Future investigation of the ability of ACIs to distinguish healthy vs unhealthy dogs is required. It is interesting that we did not find significant differences in the proportions of dogs considered systemically healthy with results within the population-based reference interval compared with the ACIs. This suggests that there is not a clear advantage in using ACIs rather than single-time point population-based reference intervals for classifying as normal dogs that do not develop overt disease in the 12 months following wellness testing. Similarly, there does not appear to be an obvious advantage to using the percentage change rather than the more intuitive analyte-specific unit change. However, this information is of limited value until we determine the ability of the 3 intervals to correctly classify as abnormal dogs that develop systemic disease within 12 months. As advocates of annual wellness testing often emphasize detection of subclinical disease, a longer follow-up than 12 months will also be needed to determine the interval that performs best for early identification of slowly progressive disease. At the time points included here, too few GRLS participants had developed diseases of interest for wellness testing programs (eg, renal or hepatic insufficiency or malignant neoplasia) to reliably determine the sensitivity and specificity of our ACIs for early disease detection. As the cohort ages and the prevalence of disease increases, we are planning to analyze serial testing results to assess the clinical usefulness of ACIs compared with single-time point population-based reference intervals for different analytes. Until such data are available, we cannot draw any conclusions about the clinical usefulness of ACIs or of annual wellness testing in general. This is an important question that requires further investigation, especially as a recent systematic review of general health checks concluded that screening of apparently healthy adult people was unlikely to be beneficial in reducing mortality or morbidities such as nonfatal stroke or ischemic heart disease.[6]

In conclusion, we have calculated ACIs that define the variability in biochemistry results between annual wellness tests for healthy Golden Retrievers. This is the first veterinary study to address biological variability in the context of wellness testing, and so might better reflect long-term homeostatic and analytical variability than previously reported RCVs calculated using a small number of animals repeatedly sampled over weeks to months.[17,18,22] Our data are likely also more representative of variation for samples collected as part of routine visits to primary care practices than RCVs generated from tightly controlled research studies. Additionally, access to data from a large group of dogs allowed us to use the direct nonparametric approach developed by Kairisto et al (1995) for ACI construction.[50] This avoids controversies regarding the most appropriate means of estimating reference change values for nonnormal data.[45,51,52] However, it should be stressed that transference studies are needed before application of our ACIs to other populations or laboratories and further data are needed before the clinical value of the ACIs can be evaluated.

ACKNOWLEDGMENT

Funding provided by Morris Animal Foundation, grant D18CLP-802, through financial support provided by the Morris Family Foundation, Blue Buffalo Cancer Research Foundation, Petco Foundation, Zoetis, Antech Inc., the Golden Retriever Foundation, the Hadley and Marion Stuart Foundation, Mars veterinary, generous private donors, and the Flint Animal Cancer Center, Colorado State University. This study was presented at the American College of Veterinary Pathology 2019 meeting. We thank the owners, breeders, and study veterinarians for their invaluable contribution to the Golden Retriever Lifetime Study.

CONFLICT OF INTEREST DECLARATION

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors declare no conflict of interest.

OFF-LABEL ANTIMICROBIAL DECLARATION

Authors declare no off-label use of antimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) OR OTHER APPROVAL DECLARATION

The study was reviewed and approved by Morris Animal Foundation's Animal Welfare Advisory Board in 2012 and 2018.

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

Unity Jeffery https://orcid.org/0000-0002-0495-2741
Kate E. Creevy https://orcid.org/0000-0003-4169-374X

REFERENCES

1. Bartges J, Boynton B, Vogt AH, et al. AAHA canine life stage guidelines. *J Am Anim Hosp Assoc*. 2012;48:1-11.
2. McKenzie B. Letter to the editor. *J Vet Intern Med*. 2017;31:1369.
3. Australian Small Animal Veterinarians Group. *Standards of Care-Regular Health Check Standards for Dogs and Cats*. St Leonards, NSW: Australian Veterinary Association; 2013.
4. Willems A, Paepe D, Marynissen S, et al. Results of screening of apparently healthy senior and geriatric dogs. *J Vet Intern Med*. 2017; 31:81-92.
5. Webb JA, Kirby GM, Nykamp SG, Gauthier MJ. Ultrasonographic and laboratory screening in clinically normal mature golden retriever dogs. *Can Vet J*. 2012;53:626-630.
6. Krogsboll LT, Jorgensen KJ, Gotzsche PC. General health checks in adults for reducing morbidity and mortality from disease. *Cochrane Database Syst Rev*. 2019;1:Cd009009.
7. Hall JA, Fritsch DA, Yerramilli M, Obare E, Yerramilli M, Jewell DE. A longitudinal study on the acceptance and effects of a therapeutic renal food in pet dogs with IRIS-stage 1 chronic kidney disease. *J Anim Physiol Anim Nutr (Berl)*. 2018;102:297-307.
8. Volk JO, Felsted KE, Thomas JG, Siren CW. Executive summary of the Bayer veterinary care usage study. *J Am Vet Med Assoc*. 2011;238: 1275-1282.

9. Leveille R, Partington BP, Biller DS, et al. Complications after ultrasound-guided biopsy of abdominal structures in dogs and cats: 246 cases (1984-1991). *J Am Vet Med Assoc*. 1993;203:413-415.
10. Vaden SL, Levine JF, Lees GE, Groman RP, Grauer GF, Forrester SD. Renal biopsy: a retrospective study of methods and complications in 283 dogs and 65 cats. *J Vet Intern Med*. 2005;19:794-801.
11. Fortney WD. Implementing a successful senior/geriatric health care program for veterinarians, veterinary technicians, and office managers. *Vet Clin North Am Small Anim Pract*. 2012;42:823-834.
12. Walton RM. Subject-based reference values: biological variation, individuality, and reference change values. *Vet Clin Pathol*. 2012;41:175-181.
13. Freeman KP, Baral RM, Dhand NK, Nielsen SS, Jensen AL. Recommendations for designing and conducting veterinary clinical pathology biologic variation studies. *Vet Clin Pathol*. 2017;46:211-220.
14. Jeffery ND, Bate ST, Safayi S, Howard MA III, Moon L, Jeffery U. When neuroscience met clinical pathology: partitioning experimental variation to aid data interpretation in neuroscience. *Eur J Neurosci*. 2018;47:371-379.
15. Ozturk OG, Paydas S, Balal M, et al. Biological variations of some analytes in renal posttransplant patients: a different way to assess routine parameters. *J Clin Lab Anal*. 2013;27:438-443.
16. Ricos C, Cava F, Garcia-Lario JV, et al. The reference change value: a proposal to interpret laboratory reports in serial testing based on biological variation. *Scand J Clin Lab Invest*. 2004;64:175-184.
17. Jensen AL, Aaes H. Critical differences of clinical chemical parameters in blood from dogs. *Res Vet Sci*. 1993;54:10-14.
18. Ruaux CG, Carney PC, Suchodolski JS, Steiner JM. Estimates of biological variation in routinely measured biochemical analytes in clinically healthy dogs. *Vet Clin Pathol*. 2012;41:541-547.
19. Frankenstein L, Wu AH, Hallermayer K, et al. Biological variation and reference change value of high-sensitivity troponin T in healthy individuals during short and intermediate follow-up periods. *Clin Chem*. 2011;57:1068-1071.
20. Wu AH, Lu QA, Todd J, et al. Short- and long-term biological variation in cardiac troponin I measured with a high-sensitivity assay: implications for clinical practice. *Clin Chem*. 2009;55:52-58.
21. Choudhury N, Wall PM, Truswell AS. Effect of time between measurements on within-subject variability for total cholesterol and high-density lipoprotein cholesterol in women. *Clin Chem*. 1994;40:710-715.
22. Jones GRD. Estimates of within-subject biological variation derived from pathology databases: an approach to allow assessment of the effects of age, sex, time between sample collections, and analyte concentration on reference change values. *Clin Chem*. 2019;65:579-588.
23. Simpson M, Searfoss E, Albright S, et al. Population characteristics of golden retriever lifetime study enrollees. *Canine Genet Epidemiol*. 2017;4:14.
24. Guy MK, Page RL, Jensen WA, et al. The Golden Retriever Lifetime Study: establishing an observational cohort study with translational relevance for human health. *Philos Trans R Soc Lond B Biol Sci*. 2015;370:20140230.
25. Harper EJ, Hackett RM, Wilkinson J, Heaton PR. Age-related variations in hematologic and plasma biochemical test results in Beagles and Labrador Retrievers. *J Am Vet Med Assoc*. 2003;223:1436-1442.
26. Reed AH, Henry RJ, Mason WB. Influence of statistical method used on the resulting estimate of normal range. *Clin Chem*. 1971;17:275-284.
27. Geffre A, Concordet D, Braun JP, et al. Reference value advisor: a new freeware set of macroinstructions to calculate reference intervals with Microsoft excel. *Vet Clin Pathol*. 2011;40:107-112.
28. Friedrichs KR, Harr KE, Freeman KP, et al. ASVCP reference interval guidelines: determination of de novo reference intervals in veterinary species and other related topics. *Vet Clin Pathol*. 2012;41:441-453.
29. Jensen AL, Kjelgaard-Hansen M. Method comparison in the clinical laboratory. *Vet Clin Pathol*. 2006;35:276-286.
30. Harr KE, Flatland B, Nabity M, Freeman KP, ASVCP. ASVCP guidelines: allowable total error guidelines for biochemistry. *Vet Clin Pathol*. 2013;42:424-436.
31. Whyte MB, Kelly P. The normal range: it is not normal and it is not a range. *Postgrad Med J*. 2018;94:613-616.
32. Tate JR, Yen T, Jones GR. Transference and validation of reference intervals. *Clin Chem*. 2015;61:1012-1015.
33. Koerbin G, Sikaris KA, Jones GR, et al. Evidence-based approach to harmonised reference intervals. *Clin Chim Acta*. 2014;432:99-107.
34. McKenzie EC, Jose-Cunilleras E, Hinchcliff KW, et al. Serum chemistry alterations in Alaskan sled dogs during five successive days of prolonged endurance exercise. *J Am Vet Med Assoc*. 2007;230:1486-1492.
35. Evason MD, Carr AP, Taylor SM, Waldner CL. Alterations in thyroid hormone concentrations in healthy sled dogs before and after athletic conditioning. *Am J Vet Res*. 2004;65:333-337.
36. Freeman LM, Chandler ML, Hamper BA, Weeth LP. Current knowledge about the risks and benefits of raw meat-based diets for dogs and cats. *J Am Vet Med Assoc*. 2013;243:1549-1558.
37. Fraser CG, Cummings ST, Wilkinson SP, et al. Biological variability of 26 clinical chemistry analytes in elderly people. *Clin Chem*. 1989;35:783-786.
38. Pineda-Tenor D, Laserna-Mendieta EJ, Timon-Zapata J, et al. Biological variation and reference change values of common clinical chemistry and haematologic laboratory analytes in the elderly population. *Clin Chem Lab Med*. 2013;51:851-862.
39. Dunlop MM, Sanchez-Vazquez MJ, Freeman KP, Gibson G, Sacchini F, Lewis F. Determination of serum biochemistry reference intervals in a large sample of adult greyhounds. *J Small Anim Pract*. 2011;52:4-10.
40. Henriksen LO, Faber NR, Moller MF, Nexo E, Hansen AB. Stability of 35 biochemical and immunological routine tests after 10 hours storage and transport of human whole blood at 21 degrees C. *Scand J Clin Lab Invest*. 2014;74:603-610.
41. Thoresen SI, Havre GN, Morberg H, Mowinckel P. Effects of storage time on chemistry results from canine whole blood, heparinized whole blood, serum and heparinized plasma. *Vet Clin Pathol*. 1992;21:88-94.
42. Ricos C, Iglesias N, Garcia-Lario JV, et al. Within-subject biological variation in disease: collated data and clinical consequences. *Ann Clin Biochem*. 2007;44:343-352.
43. Ritchie RF, Palomaki G. Selecting clinically relevant populations for reference intervals. *Clin Chem Lab Med*. 2004;42:702-709.
44. Pagitz M, Frommlet F, Schwendenwein I. Evaluation of biological variance of cystatin C in comparison with other endogenous markers of glomerular filtration rate in healthy dogs. *J Vet Intern Med*. 2007;21:936-942.
45. Lund F, Petersen PH, Fraser CG, Sölétormos G. Different percentages of false-positive results obtained using five methods for the calculation of reference change values based on simulated normal and Innormal distributions of data. *Ann Clin Biochem*. 2016;53:692-698.
46. Horn PS, Feng L, Li Y, Pesce AJ. Effect of outliers and nonhealthy individuals on reference interval estimation. *Clin Chem*. 2001;47:2137-2145.
47. Katayev A, Balciza C, Seccombe DW. Establishing reference intervals for clinical laboratory test results: is there a better way? *Am J Clin Pathol*. 2010;133:180-186.
48. Collicutt NB, Garner B, Berghaus RD, Camus MS, Hart K. Effect of delayed serum separation and storage temperature on serum glucose

concentration in horse, dog, alpaca, and sturgeon. *Vet Clin Pathol*. 2015;44:120-127.
49. Solberg HE, Lahti A. Detection of outliers in reference distributions: performance of Horn's algorithm. *Clin Chem*. 2005;51:2326-2332.
50. Kairisto V, Kouri T, Virtanen A, Uusipaikka E, Koivula T, Näntö V. Estimation of reference change limits using patient data. *Scand J Clin Lab Invest Suppl*. 1995;222:37-41.
51. Braga F, Panteghini M. Generation of data on within-subject biological variation in laboratory medicine: an update. *Crit Rev Clin Lab Sci*. 2016;53:313-325.
52. Roraas T, Stove B, Petersen PH, et al. Biological variation: the effect of different distributions on estimated within-person variation and reference change values. *Clin Chem*. 2016;62:725-736.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: 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;35:912–924. https://doi.org/10.1111/jvim.16021