

TECHNICAL NOTE

Reference Value Advisor: a new freeware set of macroinstructions to calculate reference intervals with Microsoft Excel

Anne Geffré¹, Didier Concordet², Jean-Pierre Braun^{1,2}, Catherine Trumel¹

¹Department of Clinical Sciences and ²UMR181 Physiopathologie and Toxicologie Expérimentales INRA, ENVT, Ecole Nationale Vétérinaire, Toulouse, France

Key Words

Box–Cox, confidence intervals, Gaussian, Q–Q plot, reference values, robust method

Correspondence

A. Geffré, Department of Clinical Sciences, Ecole Nationale Vétérinaire, 23 Chemin des Capelles, 31076 Toulouse Cedex 3, France
E-mail: a.geffre@envt.fr

DOI:10.1111/j.1939-165X.2011.00287.x

Abstract: International recommendations for determination of reference intervals have been recently updated, especially for small reference sample groups, and use of the robust method and Box–Cox transformation is now recommended. Unfortunately, these methods are not included in most software programs used for data analysis by clinical laboratories. We have created a set of macroinstructions, named Reference Value Advisor, for use in Microsoft Excel to calculate reference limits applying different methods. For any series of data, Reference Value Advisor calculates reference limits (with 90% confidence intervals [CI]) using a nonparametric method when $n \geq 40$ and by parametric and robust methods from native and Box–Cox transformed values; tests normality of distributions using the Anderson–Darling test and outliers using Tukey and Dixon–Reed tests; displays the distribution of values in dot plots and histograms and constructs Q–Q plots for visual inspection of normality; and provides minimal guidelines in the form of comments based on international recommendations. The critical steps in determination of reference intervals are correct selection of as many reference individuals as possible and analysis of specimens in controlled preanalytical and analytical conditions. Computing tools cannot compensate for flaws in selection and size of the reference sample group and handling and analysis of samples. However, if those steps are performed properly, Reference Value Advisor, available as freeware at <http://www.biostat.envt.fr/spip/spip.php?article63>, permits rapid assessment and comparison of results calculated using different methods, including currently unavailable methods. This allows for selection of the most appropriate method, especially as the program provides the CI of limits. It should be useful in veterinary clinical pathology when only small reference sample groups are available.

Determination of reference intervals is a long, difficult, and expensive task for all laboratories, mainly owing to the difficulty in selecting sufficient numbers of well-characterized healthy subjects and in ensuring perfectly controlled preanalytical and analytical conditions.^{1–3} Computations comprise standard statistical procedures,^{4–6} but are often perceived as the most difficult part of the task and are sometimes performed using simplistic assumptions about the data. International recommendations have been recently updated by the International Federation of Clinical Chemistry (IFCC) and the Clinical and Laboratory Standards Institute (CLSI).⁷ These recommendations confirm that a

nonparametric method is preferred when the number of reference individuals within 1 group is at least equal to 120. The revised CLSI guideline introduces determination of reference intervals from smaller reference samples based on a robust method,⁸ preferably after transformation of the data to a distribution that is closer to Gaussian or normal. Box–Cox transformation is often used to transform data to normality, but is only available on statistical software, such as R (<http://www.r-project.org>), S-PLUS (Tibco Software Inc., Palo Alto, CA, USA), and SAS (SAS Institute Inc., Cary, NC, USA). Moreover, the robust method is not currently widely available. Thus, the aim of this study was to

create a set of macroinstructions for Microsoft Excel (Microsoft, Redmond, WA, USA) that would permit evaluation and transformation of data distributions and computation of reference intervals with the corresponding 90% confidence intervals (CI) with a single click. In addition, we aimed to provide minimal guidance in the form of comments based on international recommendations. Thus, the program was named Reference Value Advisor.

Reference Value Advisor

The software is a set of visual basic macros developed in Microsoft Excel for Windows v. 2003, chosen as the development tool because of its wide availability. The selection of computations performed was guided by the IFCC–CLSI recommendations and included:

- Common descriptive statistics: sample size, mean, median, standard deviation, minimum, and maximum values
- Test of normality according to Anderson–Darling, with histograms and Q – Q plots: regardless of the results of this test on raw data, the generalized Box–Cox transformation is performed. Its λ_1 and λ_2 parameters are determined by the maximum likelihood estimator that provides the highest precision.
- Tests of “outliers”: Dixon–Reed and Tukey’s tests are used.⁷ The former detects a single potential outlier based on the ratio of its distance to the nearest value divided by the whole range of values. The latter is based on the median and interquartile range. The number of potential outliers is reported on the main results sheet with a list of outliers produced on a separate sheet. A box and whiskers diagram (median, 25th and 75th percentiles, 95% CI of mean), with a dot plot showing all values, can also be used to visually identify the outliers.
- Computations of reference intervals: for any series of data, Reference Value Advisor calculates and reports 5 reference intervals based on assumptions about data distribution. The first 4 intervals are obtained by using standard and robust methods on both nontransformed and transformed data, thus combining the need for data transformation with assumptions regarding Gaussian or symmetrical distribution of data (Table 1). The last (fifth) interval is obtained without any assumption about data distribution. Thus, it is distribution free or equivalently nonparametric, but is only computable when the sample size is large enough ($n \geq 40$). The standard intervals require data distribution to be Gaussian before or after Box–Cox transfor-

Table 1. Assumptions regarding data distribution.

	Native Data	Box–Cox Transformed Data
Standard interval	Gaussian	Gaussian after Box–Cox transformation
Robust Horn interval	Symmetrical	Symmetrical after Box–Cox transformation

mation, whereas the robust method only requires symmetry of the data.

- CIs of the reference limits computed for each method: for the standard method, the 90% CI is obtained using a parametric bootstrap when $n \leq 20$. In all other cases a nonparametric bootstrap method is used. For the nonparametric method, the CI is determined according to tables for $120 < n < 370$ ^{7,9}; if $n < 120$, a bootstrap method is used.

It was decided to consistently provide all computation results to permit selection of the most relevant results by expert users. However, it was also decided that, for inexperienced users, guidance should be provided based on IFCC–CLSI recommendations. Thus, a color code is used when reporting reference limits: green indicates *in agreement with recommendations*, orange indicates *use with caution or avoid using because possible outliers are detected*, and red indicates that *the distribution is not Gaussian*. In a comments section, the following explanations are added and are activated according to the series of data analyzed, information in the IFCC–CLSI recommendations, and the calculation methods used:

- Possible outliers detected according to Tukey or Dixon: IFCC–CLSI C28-A3 recommends that unless outliers are known to be aberrant observations, emphasis should be on retaining rather than deleting them.
- Suspect data detected according to Tukey: IFCC–CLSI C28-A3 recommends that unless these data are known to be aberrant observations, emphasis should be on retaining rather than deleting them.
- The sample size is large enough to compute a nonparametric reference interval.
- The sample size is too small ($n < 40$) to compute a nonparametric reference interval.
- The sample size is too small ($n < 80$) to compute precise CIs for the limits of the nonparametric reference interval.
- The CIs of the limits of the nonparametric reference interval were determined using a bootstrap method.
- Data should be analyzed with a nonparametric method. As an alternative, the robust method with Box–Cox transformation may be used after checking the symmetry of the distribution.

Table 2. Example of 83 randomly selected plasma creatinine values ($\mu\text{mol/L}$) in healthy dogs.

50	53	58	62	62	62	62	62	64	64	64	67
67	68	69	71	71	71	72	72	75	77	79	80
80	80	80	80	81	81	82	82	82	83	84	84
85	87	88	88	89	90	91	91	92	93	94	96
96	97	97	97	97	99	99	100	100	101	106	106
106	107	107	107	109	112	114	115	115	115	117	119
120	120	120	124	127	127	141	146	147	168	177	

- The robust method with Box–Cox transformation of the data gives the following reference interval (x ; y) (this message is provided when the nonparametric reference interval cannot be determined in accordance with the IFCC–CLSI C28-A3 recommendation and when the performance of the robust method+Box–Cox is satisfactory).
- The sample size is too small ($n < 20$) to compute CI for the limits of the reference interval obtained with the robust method with Box–Cox transformation.
- The 90% CI of 1 (or more) limit is wider than is recommended in IFCC–CLSI C28-A3.

Example of Use

Reference Value Advisor was utilized to evaluate a set of 83 canine plasma creatinine values randomly selected using the Excel RAND function from a large sample of values obtained from healthy dogs used in a previous study¹⁰ (Table 2). Descriptive statistics, detection of possible outliers, and estimations of reference intervals were performed (Figure 1), and a histogram and Q – Q plot were generated (Figure 2). Results obtained using Analyse-It (Analyse-It, Leeds, UK) were identical to those obtained using Reference Value Advisor (data not shown).

Discussion

Reference Value Advisor met the established objectives as it automatically performs a series of computations on any set of data and provides more information than is currently available in other software programs. Provision of these calculations is also potentially risky if they are used in the absence of a certain degree of expertise in determining reference intervals. For this reason, we provided comments that may aid less experienced users and encourage them to refer to IFCC–CLSI recommendations.

The main strengths of the Reference Value Advisor are the:

- ease of use and availability of computations, such as the robust method and Box–Cox transformation,

which are either rarely performed or not performed at all by commonly used mathematical programs;

- reporting of all values obtained in a graphical format, either as a dot plot or histogram;
- detection of “extreme” values, ie, possible outliers, by the 2 methods recommended by IFCC–CLSI and their clear identification on a separate report sheet. These outliers have little influence on determination of limits by nonparametric methods but greatly influence the width of the CI for these limits, especially for the upper limit. However, emphasis should be on retaining rather than rejecting outliers unless it is certain they represent aberrant observations.⁷ In the case of aberrant values, computations must be performed again after the outliers are deleted;
- systematic determination of the 90% CI of the reference limits. It is recommended that the width of the 90% CI of reference limits should not exceed the width of the reference interval by 20% or more⁶;
- possible comparison of reference intervals estimated by different methods in small samples. When different methods do not give similar results, the validity of the data may be questioned.

Reference Value Advisor also has weaknesses:

- In some cases, Gaussian distribution cannot be achieved by Box–Cox transformation, and other transformations, not included in this software, might be better.
- There is no test for distribution symmetry, which is required for the robust method. The current version includes only a test for normality, which requires more than symmetrical distribution.
- Reference Value Advisor currently lacks some procedures, such as Harris and Boyd’s test for partitioning¹¹ and a method for regression-based determination of reference intervals, eg, for age-dependent variables.¹² However, these methods can be implemented in the future.
- The set of macroinstructions currently works on Microsoft Excel for Windows up to Excel XP and 2003; updated versions are currently being developed for Excel for Mac and Excel 2010.

From a practical point of view, when $n > 120$ the recommended nonparametric method can be

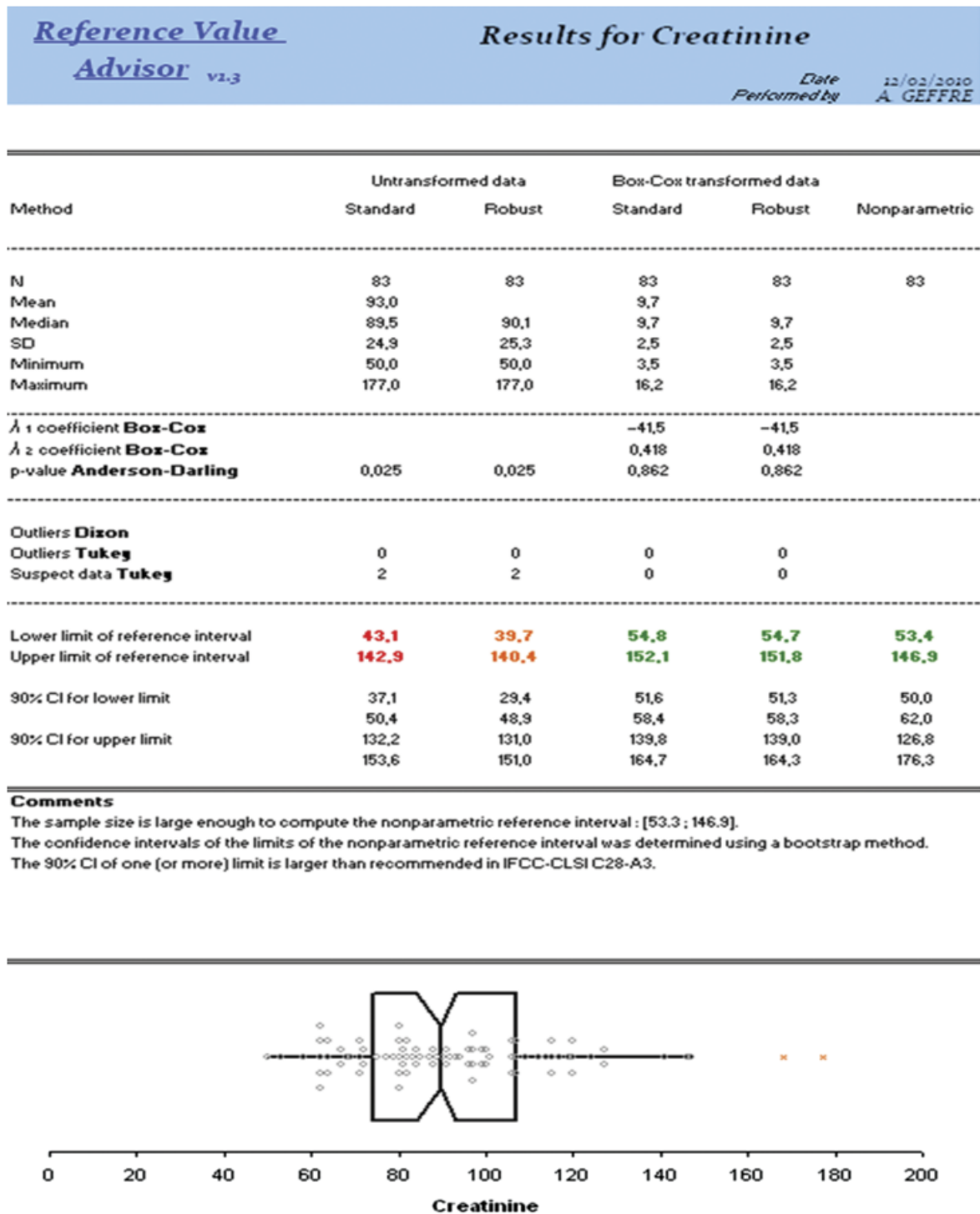


Figure 1. Printout of computations made with a series of 83 canine plasma creatinine concentrations in Table 2 using Reference Value Advisor. The untransformed distribution is not Gaussian (P [Anderson–Darling] = 0.025), but Box–Cox-transformed values fit a Gaussian distribution (P [Anderson–Darling] = 0.862). This can be visually inspected using a Q–Q plot and histogram (Figure 2). The 2 suspect outliers detected by Tukey's test appear as orange crosses on the box and whiskers diagram; they are close to the rest of the data and there is no valid reason to reject them. The color code indicates that the results calculated from the untransformed data (red and orange) should preferably not be used as the distribution is not Gaussian. The green color indicates that the reference limits are in agreement with recommendations.

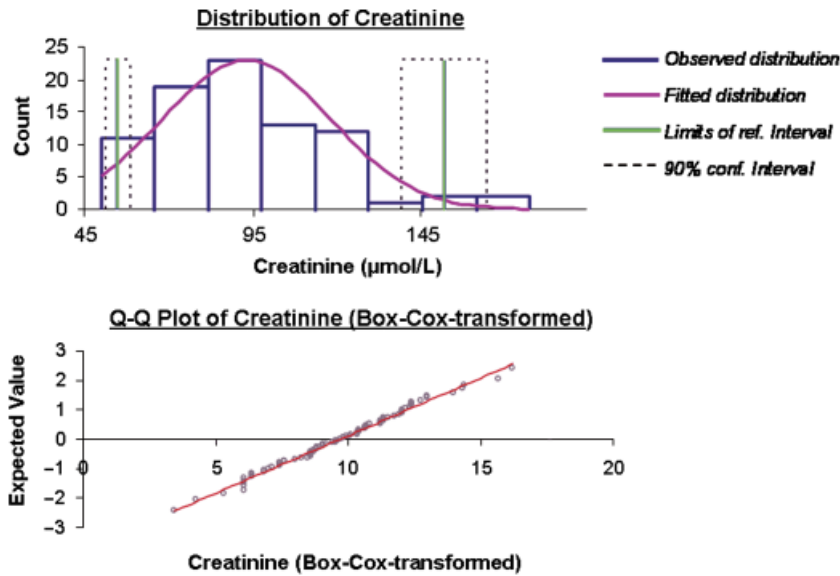


Figure 2. Histogram of untransformed and Q–Q plot of Box–Cox-transformed canine plasma creatinine results obtained from 83 healthy dogs. The histogram shows the asymmetric distribution of creatinine values; the linear arrangement of the Q–Q plot of Box–Cox-transformed values shows that the transformed distribution was very close to Gaussian.

satisfactorily used, except when visual inspection of data indicates that distribution is not unimodal. For smaller samples, the robust method on Box–Cox transformed data usually gives the narrowest CIs for the reference limits.

Conclusion

Reference Value Advisor computes reference intervals, regardless of data type or species. However, it must be emphasized that computations represent only the last step in determination of reference intervals and that the most critical steps in this process are selection of reference individuals and preanalytical and analytical conditions. Computations cannot compensate for lack of quality in any of these areas.

Availability of Reference Value Advisor

The freeware can be downloaded at <http://www.bio-stat.envt.fr/spip/spip.php?article63>

Acknowledgments

The authors greatly appreciate the input of Dr. Kristen Friedrichs, University of Wisconsin–Madison, who evaluated preliminary versions of this software and provided a series of detailed observations, comments, corrections, and suggestions.

Disclosure: The authors have indicated that they have no affiliations or financial involvement with any organization or entity with a financial interest in, or in financial competition with, the subject matter or materials discussed in this article.

References

1. Geffré A, Friedrichs K, Harr K, et al. Reference values: a review. *Vet Clin Pathol.* 2009;38:288–298.
2. Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, eds. *Tietz Textbook of Clinical Chemistry.* 3rd ed. Philadelphia, PA: Saunders; 1999:336–356.
3. Horn PS, Pesce AJ. *Reference Intervals. A User's Guide.* Washington, DC: AACC Press; 2005.
4. Solberg HE. The theory of reference values Part 5. Statistical treatment of collected reference values. Determination of reference limits. *J Clin Chem Clin Biochem.* 1983;21:749–760.
5. Solberg HE. The IFCC recommendation on estimation of reference intervals. The RefVal program. *Clin Chem Lab Med.* 2004;42:710–714.
6. Harris EK, Boyd JC. *Statistical Basis of Reference Values in Laboratory Medicine.* New York, NY: Marcel Dekker; 1995.
7. CLSI. *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory: Approved Guideline.* 3rd ed. Wayne, PA: CLSI; 2008.

8. Horn PS, Pesce AJ, Copeland BE. A robust approach to reference interval estimation and evaluation. *Clin Chem*. 1998;44:622–631.
9. Reed AH, Henry RJ, Mason WB. Influence of statistical method used on the resulting estimate of normal range. *Clin Chem*. 1971;17:275–284.
10. Concordet D, Geffré A, Braun JP, et al. A new approach for the determination of reference intervals from hospital-based data. *Clin Chim Acta*. 2009;405:43–48.
11. Harris EK, Boyd JC. On dividing reference data into subgroups to produce separate reference ranges. *Clin Chem*. 1990;36:265–270.
12. Virtanen A, Kairisto V, Uusipaikka E. Regression-based reference limits: determination of sufficient sample size. *Clin Chem*. 1998;44:2353–2358.