

## Reference intervals (2) - some practical considerations

July 2009



**Chris Higgins**

Little Acre, Main Road  
Shurdington  
Nr Cheltenham  
Gloucester  
GL51 4XF, UK  
E-mail: [cjhiggins@hotmail.co.uk](mailto:cjhiggins@hotmail.co.uk)

This is the second of two articles focusing on the reference interval, the most widely used tool for interpretation of patient test results.

The first [1] was an introduction to the theoretical concepts that underpin the significance, construction and use of reference intervals. Here consideration is given to the more practical issues of transferring an established reference interval and validating the adoption of an established reference interval. Establishing reference intervals from scratch is an enormous undertaking that could not reasonably be expected to be within the remit of most clinical laboratories.

As will become clear, transference and validation of established reference intervals are by contrast relatively simple procedures that can and arguably should be undertaken by all clinical laboratories when making analytical and sampling procedural changes, and when introducing an established test to the laboratory repertoire.

This article is based largely on consensus expert opinion contained in recently published guidelines (C28-A3) prepared jointly by the Clinical and Laboratory Standard Institute (CLSI) and the International Federation of Clinical Chemistry (IFCC); these guidelines [2] represent the single most authoritative source for laboratorians seeking to establish or validate reference intervals.

There is general acceptance that only a minority of clinical laboratories are sufficiently resourced in terms of time, finance and expertise to establish reference intervals for all, or indeed any of the tests they routinely perform.

Establishment of a reference interval, which is of course mandatory before a new test can be introduced to routine clinical use, requires extensive knowledge (or study) of the pathophysiological significance of the analyte in question; considerable laboratory work, including analyzing a minimum of 120 reference

samples; knowledge and application of appropriate statistical tools; and copious documentation [2].

However, all laboratories are responsible for the reference intervals that they are mandated to publish alongside patient test results, and in recent years there has been increased emphasis of the notion that all laboratories, no matter what their size or level of resource, should at least reflect on, and justify at some level, the reference intervals they adopt.

Echoes of this position can be found in advice/directives from laboratory regulatory authorities.

For example ISO 15189:2007 [3], the international standard that defines quality and competency in clinical laboratories and provides the basis for laboratory accreditation across Europe and beyond, states *"reference intervals shall be periodically reviewed"* and verified every time a variation in analytical and/or preanalytical procedure takes place.

In the US, Clinical Laboratory Improvements Amendments (CLIA) from 2003 states that when FDA-approved test systems are adopted unmodified, laboratories should *"verify that the manufacturers' reference intervals are appropriate for the laboratory's patient population"* [4].

Despite moves to raise the profile of reference intervals within the laboratory community at large, there remains a paucity of knowledge about just how in practice the generality of laboratories select and/or validate their published reference intervals.

A recently published US survey of 163 clinical laboratories [5], conducted under the auspices of the Q-probe study program of the College of American Pathologists [6], provides some answers and raises some concerns.

### The q-probe study of reference intervals

The 163 clinical laboratories participating in this survey of reference interval policy and practice are representative of laboratories throughout the US hospital system

(including large and small, teaching and non-teaching and city, suburban, rural location).

Participating laboratories were asked to supply their adult and pediatric reference intervals (low and high limits) for four common clinical chemistry parameters (potassium, calcium, magnesium and TSH) and three equally common hematological parameters (hemoglobin, platelet count and activated partial thromboplastin time).

They were also asked when and how these reference intervals were arrived at, how long since they were last reviewed and the measuring platform for each analyte.

Survey results revealed that a range of approaches were used to arrive at selected reference intervals. Only a half of the laboratories reported analyzing samples from healthy individuals in preparation of adult reference intervals.

Even fewer (25 %) reported analyzing samples in preparation of pediatric reference intervals. The remaining laboratories adopted reference intervals from external sources without any internal study.

The most frequent external source was manufacturers' recommendations/package inserts, but text books/medical journals and non-laboratory medical staff recommendation were the source for some laboratories.

Among those laboratories that conducted any sort of internal study, the number of samples analyzed ranged from as few as 20 to >100. The results of sample analysis were used to establish reference intervals in around a half of these laboratories. For the remaining laboratories, results of the internal study were used to validate externally sourced reference intervals.

The survey revealed that 26 % of the participating laboratories do not have a written policy for establishing, revising or updating reference intervals.

Approximately two thirds of the laboratories reported that they had revalidated their reference intervals in

the year that a new analyzer was purchased, but some laboratories reported no validation of reference intervals in the previous 10 years, and in one case there had been no validation for at least 22 years.

A “*number of institutions*” reported that they did not know the year their reference intervals were established or when they were last revalidated.

Analysis of the submitted reference intervals (Table 1) revealed that for most (80 %) laboratories there was “*only slight*” variation in reference interval limits. However, among the remaining 20 % of laboratories, for which more substantial variation was evident, there were some with “*surprisingly low and high limits*” for their reference intervals.

For example, in the case of potassium the majority of laboratories had a lower limit close to the median value of 3.5 mmol/L and a high limit close to median value 5.1 mmol/L, but one laboratory quoted a lower limit of 3.0 mmol/L, and another a lower limit of 4.0 mmol/L. The minimum and maximum high limits were 4.5 and 5.7 mmol/L, respectively.

Statistical analysis of the whole data set (reference intervals from all laboratories for all seven analytes) revealed that of 1271 adult reference intervals 40 (3.1 %) contained at least one limit that was a statistical outlier.

For some of the analytes (magnesium, TSH and APPT) a certain amount of the observed variation in reference intervals between laboratories could be accounted for by differences in analytical methodology, but it certainly did not account for all of the variation.

Despite the author-acknowledged limitations of this self-reporting study, it provides the best available evidence surrounding reference interval policy and practice in the generality of hospital laboratories and suggests that in some laboratories - albeit a small minority - inaccurate reference intervals may be being used to interpret patient test results.

How then can laboratories ensure that the reference intervals they adopt are fit for purpose?

## CLSI/IFCC guidelines

Recently updated CLSI/IFCC guidelines [2] for the establishment of reference intervals acknowledge that it is neither feasible nor necessary for most laboratories to establish their own reference intervals.

Two far simpler approaches are proposed: transference of an established reference interval and validation of an established reference interval.

It is of course assumed that the reference interval to be transferred or validated has been established in accord with guidelines, and the first step for all laboratories, if it has not already been done, is to document and review all that is known about how the reference interval to be transferred/adopted was established.

Aspects to be considered include:

- Reference population demographics (age, gender, ethnicity)
- Inclusion/exclusion criteria for selection of reference sample group
- Size of reference sample group
- Preanalytical and analytical procedures for generation of reference values
- Method of estimating reference interval from reference values

## Transferring existing reference intervals

The guideline advice relating to transferring existing reference intervals is applicable in the situation where a laboratory is changing the analytical method for a particular analyte. The laboratory has an acceptable reference interval for the old method and needs to know if that reference interval is applicable to the new method.

The guideline for transferring reference intervals is based on the notion that the two most important variables

All institution Percentiles								
Analyte	n	Minimum	10th	25th	50th	75th	90th	Maximum
<b>Potassium mmol/L</b>								
low limit:	162	3.0	3.3	3.5	<b>3.5</b>	3.6	3.6	4.1
high limit:	162	4.5	5.0	5.0	<b>5.1</b>	5.2	5.3	5.3
<b>Calcium mg/dL</b>								
low limit:	162	7.6	8.3	8.4	<b>8.5</b>	8.5	8.8	9.0
high limit:	161	9.6	10.1	10.1	<b>10.2</b>	10.5	10.5	10.8
<b>Magnesium mg/dL</b>								
low limit:	157	1.3	1.5	1.6	<b>1.6</b>	1.8	1.8	2.1
high limit:	157	2.0	2.2	2.3	<b>2.4</b>	2.6	2.7	3.2
<b>TSH mIU/L</b>								
low limit:	154	0.10	0.30	0.34	<b>0.35</b>	0.40	0.47	0.50
high limit:	155	3.00	4.20	4.68	<b>4.94</b>	5.50	5.60	6.00
<b>Hemoglobin g/dL</b>								
male low limit:	161	11.5	13.0	13.0	<b>13.6</b>	14.0	14.0	14.8
high limit:	161	15.0	16.5	17.0	<b>17.5</b>	18.0	18.0	18.1
female low limit:	159	10.4	11.5	11.7	<b>12.0</b>	12.0	12.1	14.0
high limit:	159	13.5	15.0	15.5	<b>16.0</b>	16.0	16.0	18.0
<b>Platelets × 10/L</b>								
low limit:	161	100	130	130	<b>142</b>	150	150	174
high limit:	161	316	400	400	<b>400</b>	440	450	500

TABLE 1: Analysis of the submitted reference intervals

that influence a reference interval are the method of analysis and the population from which the reference individual samples are taken.

Since the test population is unchanged in the scenario outlined above, the only consideration for transference of the reference interval is comparability of the two analytical methodologies. When implementing a new method, it is normal laboratory practice to perform a method comparison study in which the same fresh patient samples are measured by both methods.

If the study shows that the two assays are completely comparable across the measuring range (good correlation and no bias), then the reference interval can be adopted unchanged.

Alternatively, if the study shows good correlation but a proportional negative or positive bias between the two methods, it may be acceptable to use the regression equation generated by the study to “correct” the reference interval to take account of this systematic bias.

The guidelines provide the following example of the way this is applied:

The results of a comparison study of methods x (old method) and y (new method to be adopted) across a concentration range of 50-250 give the best-fit linear regression line:

$$y = 1.57x - 0.832 \text{ correlation coefficient } r^2 = 0.990$$

The established reference interval for method x is 50-150.

Since there is excellent correlation but proportional bias between the two methods, the “corrected” reference interval for method y can be calculated thus:

For the lower limit 50

$$y = (1.57 \times 50) - 0.832 = 77.72 \text{ (which rounds up to 78)}$$

For the high limit 150

$$y = (1.57 \times 150) - 0.832 = 234.82 \text{ (which rounds up to 235)}$$

The reference interval to be adopted for the new method y is 78-235.

Essentially it is acceptable to simply transfer an existing reference interval so long as the population being tested is the same, preanalytical procedures are unchanged and comparability of the two methods has been demonstrated by an acceptably conducted method comparison study.

A minimum of 40 patient samples should be tested and they should be selected so that full concentration range in health and disease is represented. The detail of conducting an acceptable method comparison study is contained in a separate CLSI document EP09 [7].

The obvious advantage of the transferring protocol is that it does not require analysis of samples from reference individuals. However, it has limited application because it only applies if the reference interval in question has been in use at that particular institution.

Furthermore, a level of judgment is required to make the decision about whether or not the two methods agree sufficiently for them to share the same reference interval. In cases where there is some doubt, the guidelines suggest that validation of the reference interval is indicated.

## Validating an established reference interval

Validation of an established reference interval is appropriate when a laboratory wishes to adopt an established reference interval supplied by a manufacturer or another laboratory for the same or similar analytical system.

The preanalytical protocol used in the adopting laboratory for processing patient samples should not be significantly different from that used for determining reference values when establishing the reference interval.

The validation study is designed to confirm that the established reference interval is appropriate for the population served by the adopting laboratory. It involves

determining reference values for at least 20 individuals who are judged to be representative of the adopting laboratory's healthy population.

The exclusion criteria used to select these individuals should reflect those originally used in selection of reference individuals for the establishment of the reference interval.

The procedure used for determination of the 20+ reference values must also accord with preanalytical/analytical procedures defined by the original reference interval study and therefore with protocols adopted in the laboratory for measuring patient samples.

The test results from the reference individuals are first examined for the presence of outliers. The guidelines recommend that either the Dixon/Reed method [8, 9] or Tukey method [10] be used to test for outliers.

The Dixon/Reed method for identifying outliers is based on the ratio  $D:R$  where  $D$  is the absolute difference between the most extreme value of a data set (i.e. the possible outlier) and the next most extreme value, and  $R$  is the range of all values. If  $D$  is equal to or greater than one third of the range  $R$ , then the most extreme value is an outlier.

The Tukey approach, which is described in detail in the guidelines, is more complicated but statistically more robust than the Dixon/Reed method.

Any outliers identified must be eliminated and replacement samples obtained, so that a statistically homogeneous group of at least 20 reference values are available for comparison with the established reference interval.

The guidelines stipulate that so long as no more than two of the 20 (10 %) reference values fall outside the limits of the established reference interval, it is appropriate for the laboratory to adopt the reference interval. If three or more values fall outside the reference interval, the whole procedure should be repeated with samples from a different set of 20 reference individuals.

As before, if no more than two of 20 reference values fall outside the reference interval, it is appropriate for the laboratory to adopt the reference interval.

However, if once again three or more values fall outside the reference interval, it is an indication that the population served by the laboratory differs significantly from that used to prepare the reference interval, and it might therefore be inappropriate to adopt the reference interval.

The lack of agreement might alternatively be due to unrecognized differences in preanalytical/analytical procedures and this possibility should be reviewed and, if confirmed, corrected.

If after full investigation and further validation study the problem remains unresolved, guidelines suggest that the laboratory should consider establishing its own reference interval.

## Summary

The topic of reference intervals seemed for many years to be the sole preserve of an expert clique(s) within the laboratory community, and the generality of clinical laboratory staff, whilst appreciating its importance, viewed it as a rather arcane subject, perhaps best left to the experts.

Regulatory authorities now demand that more laboratory staff engage with the topic in a proactive way. It is no longer acceptable laboratory practice, if indeed it ever was, to simply adopt a published reference interval without careful consideration (due diligence in modern, post-credit crunch, parlance).

These two articles were intended to introduce the topic of reference intervals to those laboratory staff and other interested parties who have no particular knowledge of, or expertise in the field.

This second article highlights the lack of conformity surrounding reference interval policy and describes an expert-devised approach to the validation of reference intervals that could be applied in all laboratories, no matter what their level of resource.

## References

1. Higgins C. An introduction to reference intervals (1) - some theoretical considerations. [www.bloodgas.org](http://www.bloodgas.org) 2009.
2. Clinical and Laboratory Standards Institute (CLSI). Defining, establishing and verifying reference intervals in the clinical laboratory. Approved guideline. 3rd ed. CLSI document C28-A3. Clinical and Laboratory Standards Institute Pennsylvania USA 2008.
3. ISO 15189:2007 Medical laboratories - particular requirements for quality and competence. International Organization for Standardization. Geneva Switzerland 2007.
4. Clinical Laboratory Improvement Amendments of 1988 (CLIA). Final Rule § 493.1253 (b)(1)(ii) (2003) available at: [www.cdc.gov/clia/pdf/CMS-2226-F.pdf](http://www.cdc.gov/clia/pdf/CMS-2226-F.pdf) (accessed March 2009).
5. Freidberg R, Souers R, Wager E *et al*. The origin of reference intervals. *Arch Pathol Lab Med* 2007; 131: 348-57.
6. Howanitz P. Quality assurance measurements in departments of pathology and laboratory medicine. *Arch Pathol Lab Med* 1990; 114: 1131-35.
7. Clinical and Laboratory Standards Institute (CLSI) Method comparison and bias estimation using patient samples. Approved guideline (2nd ed) CLSI document EP09-A2. Pennsylvania USA 2002.
8. Dixon WJ. Processing data for outliers. *Biometrics* 1953; 9: 74-89.
9. Reed A, Henry R, Manson W. Influence of statistical method used on the resulting estimate of normal range. *Clin Chem* 1971; 17: 275-84.
10. Tukey J. *Exploratory data analysis*. Reading, MA: Addison-Wesley; 1977.