

“Are my laboratory results normal?” Considerations to be Made Concerning Reference Intervals and Decision Limits.

Ferruccio Ceriotti^{1*}, Joseph Henny²

¹ Diagnostica e Ricerca San Raffaele S.p.A., IRCCS San Raffaele Hospital, Milan, Italy, Chair IFCC Committee on Reference Intervals and Decision Limits

² Laboratoire de Biologie Clinique, Centre de Médecine Préventive, Vandoeuvre-lès-Nancy, France, member IFCC Committee on Reference Intervals and Decision Limits

*Author for correspondence

Ferruccio Ceriotti

Diagnostica e Ricerca San Raffaele

Via Olgettina 60, 20132 Milano, ITALY

Phone +39 0226432282

Fax +39 0226432640

Email: ceriotti.ferruccio@hsr.it

Summary: This paper looks at the topic of reference intervals from the point of view of the patient or the clinician. The differences between the concepts of reference intervals (biological characteristic of a well defined population) and the various types of decision limits are illustrated and discussed. Decision limits can be defined in different ways: based on a Bayesian approach, on epidemiological studies or on clinical experience, but differ from reference intervals because, while the latter deals with physiology, decision limits are related to some kind of disease or risk of developing it.

“Are my laboratory results normal?” This simple, straightforward question can be extremely difficult to answer. As Arky pointed out in his editorial (1), the concept of “normal laboratory results” has evolved with time, which explains why the question is so difficult to answer and why the replies can be different.

An aspect to consider first is the ambiguity of the word “normal”. It is clear that this term is not scientific because, as Murphy stated many years ago, “try as we may, we cannot come up with anything like an absolute definition of normal from the scientific viewpoint” (2). The papers of Grasbeck, Saris and Dybkaer (3,4) in the late 1960s and the beginning of the 1970s clarified this concept.

We can decode this question as “Am I healthy or sick?” or “Am I at risk for some disease?” Obviously this already appears a much more difficult question and explains why the laboratory and the clinician can give, in most cases, only partial and sometimes contradictory answers. The laboratory and the clinician can reply in different ways because the laboratory usually knows neither the reason for the request of a specific test nor the clinical situation of the patient.

The reply depends upon the context, such as the reason that generated the request for the test:

- screening / health check
- assessment of a risk factor (probability to have a disease in the future, prediction of risk)
- diagnosis of disease
- disease management, “treatment goals”

- follow up.

The results of a laboratory test can only be judged by making a comparison, but different clinical questions require different types of references to be used. We can use 1) a set of results from supposedly healthy individuals (reference interval) for screening or health checks; 2) a limit to decide on a specific level of risk or probability for the presence of a certain disease (decision limits); or 3) the previous results of the patient himself (reference change value) for follow up.

Now we start to understand that the apparently simple question “Are my laboratory results normal?” is not a single question, but includes many. A first question is: “Are my laboratory results within the reference interval?” The reply to this question is present in any report from the laboratory, but the clinician must be cautious in the interpretation.

It is probably useful to recall the concept and the definition of reference intervals (5,6). A reference interval is the interval between, and including, two reference limits, which are values derived from the distribution of results obtained from a sample of the reference population. The reference population is a hypothetical entity including an unknown number of reference individuals. A reference individual is a person selected for testing on the basis of well-defined criteria; in the case of health related reference intervals (also called physiological or biological reference intervals), the selection criteria are designed to exclude the most usual pathological conditions known to affect the concentration values of the analyte under investigation. It is essential to notice that the reference limits are defined through various statistical methods so that a stated fraction (usually 2.5%) of the reference values is less than or equal, or more than or equal, to the respective upper and lower limits. Another important observation is that selecting a different sample of reference individuals from the reference population would give slightly different limits. This implies that the reference limits can swing within a range of values with a stated probability and the width of this range depends upon the number of the individuals selected from the population (the higher the number, the smaller the range), the shape of the distribution (wider range in case of skewed distributions) and the accuracy of the selection of the reference individuals (number of unhealthy subjects erroneously included in the reference group). All these technical problems, relatively difficult to understand for the patient or the clinician, have just one practical consequence: there is a “grey zone” and results close to the limits (both within or outside the interval) are difficult to interpret and classify.

But there are also other factors that can influence the reply: the analytical variability (imprecision and bias of the analytical method in use, leading to uncertainty in the result), the intraindividual biological variability of the subject under examination, and, finally, the correctness of the reference interval applied by the laboratory.

As mentioned above, the position of the result of the patient relative to the reference limits is critical for our discussion. If the result is in the middle of the range the answer is easy (provided that the range in use is correct; see the following creatinine example), but if it is close to the upper or lower limit (either within or outside) a yes or no reply could be questionable. In this case the better answer could be that the probability that your result belongs to the population of values obtainable from a group of individuals not affected by a specific disease is fairly high (if inside) or fairly low (if outside). This kind of answer means nothing to the patient and has to be interpreted by taking into consideration, for example, the individuality index of that particular analyte (7) and thus the previous results of the patient. For these reasons the use on the report of flags (H, L, *, >, <, etc.) that identify a result outside the limits is questionable. They are useful to present the highly pathological results and make the check of the report easier for the laboratory side, but tend to drive

the interest only to some components, possibly overlooking other borderline data, which may be more significant in that clinical context. Moreover, they usually terrify the patient who is not aware of the fact that the reference intervals include only the 95% of the unaffected individuals and that he has a 5% statistical probability of being outside the limits even if he is unaffected.

The appropriateness of the interval used by the laboratory is another variable that is often relevant for paediatric or geriatric subjects. For example, 0.8 mg/dL (71 μ mol/L) of serum creatinine can be a result indicating a regular glomerular filtration rate if obtained for an adult male, but an impaired renal function in a 5-year-old boy (8). This kind of problem is more relevant for analytes that have reference intervals changing with age or sex (like creatinine, alkaline phosphatase and other enzymes, hormones etc.).

The choice of a correct reference interval is also problematic in the case of vitamins or markers of iron status such as ferritin. In these cases the real problem is the criteria used to define the characteristics of the reference group. Just the subjective wellbeing status and some generic markers of good health such as haemoglobin concentration are not sufficient to exclude subclinical deficiencies. This can explain why many discrepancies exist in the literature and why the data of Dorizzi et al. (9) are different from the numbers currently in use (10). In conclusion, reference intervals have to be used as a guide to help in making medical decisions, not as an absolute indicator of health or disease.

A second question for consideration should be: “Are my laboratory results lower (or higher) than the decision limit?” The two questions appear to be similar; the difference is essentially that health related reference intervals depend on physiology, whereas decision limits depend on the type of pathological condition you are dealing with. Reference intervals are a biological characteristic of the population and, if defined with the necessary attention to the choice of subjects, with partitioning for gender, age, etc., combined with the use of standardized analytical methods producing results traceable to the reference measurement system, they could (and should) remain stable. Decision limits, on the contrary, depend upon the type of decision to be made; they can be changed from time to time as new scientific information becomes available (e.g. the decision limit for triglycerides shifted from 250 mg/dL to 200 and eventually to 150 mg/dL in the three consecutive ATP documents (11-13)). For the same individual there can be different decision limits for the same measurand (e.g. for total cholesterol). Murphy and Abbey (14) defined decision limits as “the best dividing lines between the ‘normal’ and the ‘diseased’ or between ‘those who need not be investigated further’ and ‘those who do’”. The differences between reference intervals and decision limits are summarized in table 1, modified from Ceriotti et al. (15).

There are different types of decision limits and they can be classified according to different criteria such as the following:

1. Bayesian approach. As proposed by Sunderman in 1975 (16), to define what he called “discrimination values”, six basic requirements are needed:

- a) definition of the disease being sought
- b) demarcation of the stage in the pathogenesis of the disease that is the objective for diagnosis by the test to be evaluated.
- c) knowledge of the clinical (diagnostic) sensitivity of the diagnostic test
- d) knowledge of the clinical (diagnostic) specificity of the diagnostic test
- e) knowledge of the prevalence of the disease
- f) knowledge of the consequences (clinical cost) of positive or negative misdiagnoses.

Following these criteria, the definition of decision limit is stated in the paper by Sunderman (16) to be “a value for the result of a diagnostic test that serves as a criterion for distinguishing between two subgroups of a population being tested, based upon stated assumptions regarding (a) the clinical sensitivity of the diagnostic test; (b) the clinical specificity of the diagnostic test; (c) the relative distribution of individuals between the two subgroups; and (d) the clinical costs of misclassification.” This is probably the most useful and evidence based approach that can modify the management of the patient, but, up to now, 33 years later, not many decision limits have been defined following these criteria. The concept has been applied to troponin I (17), ALT (18), and few others.

2. Epidemiological approach. This approach is based on population studies and is typically applied in the lipid field (total cholesterol, LDL cholesterol etc.) or to measure glucose or glycated haemoglobin. The decision limits were defined during consensus conferences or based on guidelines. The precise choice of the number is arbitrary (often based on easiness to remember: e.g. the decision limit for total cholesterol is 200 mg/dL (5.17 mmol/L) for the US based guideline (13) and 5.00 mmol/L (190 mg/dL) for the European based guideline (19)), but the rationale is based on outcome studies that demonstrated a different level of survival or a different incidence of complications for patients with concentrations below or above the limit. A nice example of this approach is given by Alehagen et al. for BNP and NT-proBNP (20).

3. Physiopathological approach. This approach involves the use of so called “critical values” or “panic values”. These are results representing a pathophysiological state with such variance from normal as to be life-threatening unless something is done promptly and for which some corrective action could be taken (21). The original concept has been widened to include “alert values” for non-life-threatening situations (22). These limits are somewhat arbitrary, being based on clinical experience, usually without the support of any statistical mean (23-25).

The third question to consider is: “Are my present laboratory results different from those I had in the past?”. This question can be particularly relevant when dealing with analytes with low individuality indices. In this case the intraindividual variability is much lower than the interindividual variability and thus variations in the concentration of the analyte still within the reference interval can be significantly outside the subject’s usual values. In this case it is useful to calculate if the reference change value has been surpassed or, even better, to calculate the statistical significance of a trend. The graphical representation of the data can be extremely effective in identifying trends even without any statistical evaluation. This approach can be applied to the interpretation of results of tumor markers, but can also be used for any analyte with a low intraindividual variability (e.g. creatinine).

Now we can imagine some paradigmatic situations where a clinician is faced with the question raised at the beginning of the paper.

Case 1: Haemoglobin (Table 2)

As a first step the clinician compares the observed value (OV) to the reference interval. If the OV is within the interval he will look for a difference from the previous results of the patient to identify a possible trend. If lower than the low reference limit (RL), he will compare this OV to a medical decision limit (MDL): i.e. 13.0 g/dL for anaemia (26) and/or 8.0 g/dL (or 7.0 g/dL) for blood transfusion (27). It means that the RLs have only a screening value at first; they don’t play a role in the medical decision making unless the OV is inside the reference interval.

Case 2: Fasting Glucose (Table 3)

The reference intervals for fasting glucose in adults (70 - 115 mg/dL) (28) have a limited clinical value: they overlap with the DLs for the diagnosis of diabetes. The lowest DL for determining if an Observed Value is located inside the “grey zone” (100 mg/dL) (29) is inside the RI, and the highest DL (for diagnosis of diabetes) is outside the RI. What shall the laboratory worker print on his laboratory report: 100, 115 or 126 mg/dL (5.5, 6.4 or 7.0 mmol/L)? In this case we suggest that he adopt the DL (100 mg/dL), adding a note that specifies that it is a decision limit to define the presence of an “impaired fasting glucose” condition.

Case 3: Calcium and Potassium (Table 3)

For these two analytes there is a similar situation in that clinicians usually do not react to borderline values, but rather use DL limits or critical values.

In conclusion, shall we replace Reference Intervals with Decision Limits? Certainly yes for all the analytes that have well defined and accepted DL (glucose, cholesterol etc.) and probably yes for many others in the future when evidence based DL will be available. For a number of years reference intervals will still play a major role for most analytes as a guide in the interpretation of laboratory tests. They will also continue to have a relevant function as control group to calculate the clinical sensitivity and specificity of a test and as an essential prerequisite in the definition of a DL.

Coming back to the initial question, “Are my laboratory results normal?”, it now appears clearer why the reply cannot always be simple. The laboratory can significantly help the clinician by providing information on reference intervals or decision limits and by indicating significant deviation from previous results, but the clinician has to use his/her clinical knowledge, experience, and wisdom to use these supports in the correct way. The collaboration between laboratory and clinician is essential to put together all the pieces of knowledge needed to reply to this apparently “simple” question.

References:

1. Arky RA. "Doctor, Is My Sugar Normal?" *N Engl J Med* 2005;353:1511-3.
2. Murphy EA. A scientific viewpoint on normalcy. *Perspect Biol Med* 1966;9:333-48.
3. Gräsbeck R, Saris NE. Establishment and use of normal values. *Scand J Clin Lab Invest* 1969;26 (Suppl 110):62-3.
4. Dybkaer, R., Concepts and nomenclature in theory of reference values. *Scand J Clin Lab Invest.* 1972;29 (Suppl. 126),19-21.
5. Solberg HE. International Federation of Clinical Chemistry (IFCC). Scientific Committee, Clinical Section. Expert Panel on Theory of Reference Values and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *J Clin Chem Clin Biochem* 1987;25:337-42; *Clin Chim Acta* 1987;165:111-8; *Labmedica* 1987;4:27-31; *Ann Biol Clin* 1987;45:237-41.
6. CLSI document C28-A2. How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline — Second Edition. Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2000.
7. Harris EK. Effects of intra- and interindividual variation on the appropriate use of normal ranges. *Clin Chem* 1974;20:1535:42.
8. Ceriotti F, Boyd JC, Klein G, et al. Reference intervals for creatinine concentrations in serum: assessment of the available data for global application. *Clin Chem* 2008;54:559-66.
9. Dorizzi RM, Fortunato A, Marchi G, Scattolo N. Reference interval of ferritin in premenopausal women calculated in four laboratories using three different analyzers. *Clin Biochem.* 2000;33:75-7.
10. Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of clinical chemistry and molecular diagnostics*. Fourth edition, St. Louis MO. Elsevier Saunders, 2006, page 1191.
11. Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. The Expert Panel. *Arch Intern Med.* 1988;148:36-69.
12. Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II) *JAMA.* 1993;269:3015-23.
13. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) *JAMA.* 2001;285:2486-97.
14. Murphy EA, Abbey H. The normal range—a common misuse. *J Chronic Dis* 1967; 20:79-88.
15. Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. *Ann Clin Biochem* in press
16. Sunderman F W Jr. Current concepts of “Normal values”, “Reference values”, and “Discrimination values” in clinical chemistry. *Clin Chem* 1975;21:1873-7.
17. Pagani F, Sonetti G, Stefini F, Cuccia C, Panteghini M. Determination of decision limits for ACS:System Cardiac Troponin I. *Clin Chem Lab Med* 2000;38:1155-7.
18. Prati D, Taioli E, Zanella A, Della Torre E, Butelli S, Del Vecchio E, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. *Ann Intern Med.* 2002;137:1-10.
19. Executive Summary. European guidelines on cardiovascular disease prevention in clinical practice. Third joint task force of European and other Societies on cardiovascular disease prevention in clinical practice. *Eur Heart J* 2003;24:1601-10
20. Alehagen U, Goetze JP, Dahlström U. Reference intervals and decision limits for B-type natriuretic peptide (BNP) and its precursor (Nt-proBNP) in the elderly. *Clin Chim Acta* 2007;382:8-14.
21. Lundberg GD. When to panic over abnormal values. *Med Lab Observ* 1972;4:47-54.

22. Howanitz PJ, Steindel SJ, Heard NV. Laboratory critical values policies and procedures: a college of American Pathologists Q-Probes Study in 623 institutions. *Arch Pathol Lab Med* 2002;126:663-669.
23. Emancipator K. Critical values: ASCP practice parameter. *Am J Clin Pathol* 1997;108:147-153.
24. Dighe AS, Rao A, Coakley AB, Lewandrowski KB. Analysis of laboratory critical value reporting at a large academic medical center. *Am J Clin Pathol* 2006;125:758-764.
25. Tillman J, Barth JH on behalf of the ACB National Audit Group. A survey of laboratory 'critical (alert) limits' in the UK. *Ann Clin Biochem* 2003;40:181-184.
26. World Health Organisation. Nutritional anemia. Report of a WHO scientific group. Geneva Switzerland: WHO, 1968
27. Corwin HL, Gettinger A, Pearl RG, Fink MP, Levy MM, Abraham E, et al. The CRIT Study: Anemia and blood transfusion in the critically ill--current clinical practice in the United States. *Crit Care Med.* 2004;32:39-52.
28. Thomas L editor. *Clinical Laboratory diagnostics*. TH – Books, Frankfurt/Main, Germany, First edition English 1998, page 132.
29. The expert committee on the diagnosis and classification of diabetes mellitus. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care.* 2003;26:3160-7.
30. Thomas L editor. *Clinical Laboratory diagnostics*. TH – Books, Frankfurt/Main, Germany, First edition English 1998, page 475.
31. Thomas L editor. *Clinical Laboratory diagnostics*. TH – Books, Frankfurt/Main, Germany, First edition English 1998, page 232.
32. Tietz NW. *Clinical guide to laboratory tests*. Saunders. Third edition. 1995, page 103
33. Thomas L editor. *Clinical Laboratory diagnostics*. TH – Books, Frankfurt/Main, Germany, First edition English 1998, page 307.

Table 1. Differences between reference intervals and decision limits

	Reference intervals	Decision limits
Brief definition	The interval between, and including, two reference limits, which are values derived from the distribution of the results obtained from a sample of the reference population.	The best dividing lines between the diseased and the not diseased or between “those who need not be investigated further” and “those who do”.
Conditions influencing them	<ul style="list-style-type: none"> ▪ Type of population ▪ Age group ▪ Gender 	<ul style="list-style-type: none"> ▪ Clinical question ▪ Patient category
Information gathered	Whether or not the patient is part of the reference population	Whether or not the patient is eligible for a certain procedure ("treatment")
Statistics	95% central range of the distribution curve	<ul style="list-style-type: none"> ▪ None (consensus values) ▪ ROC curves ▪ Predictive values
Data number	Two (lower and upper limits) each one with an associated 90% confidence interval	One, without any confidence interval There can be several, but according to the likelihood of various clinical situations or different clinical questions

Table 2: Reference Limits and Decision Limits for haemoglobin. Implications for the diagnosis of Anaemia.

	Males	Females	Ref.
<i>Reference interval for adults (g/dL)</i>	14.0 - 17.5	12.3 - 15.3	(30)
<i>Decision limit for anaemia (g/dL)</i>	13.0	12.0	(26)
<i>Decision limit for blood transfusion (g/dL)</i>	7.0	7.0	(27)

Table 3. Relationship between adult Reference Intervals and Decision Limits for three laboratory tests

GLUCOSE			
(fasting plasma)			
<i>Reference Intervals Adults</i>		70 – 115 mg/dL (3.9 – 6.4 mmol/L)	(28)
<i>Decision Limits</i>	Hypoglycemia	55 – 70 mg/dL (3.0 – 3.9 mmol/L)	Possible impairment of cognitive function
	Life-threatening hypoglycemia	≤40 - 55 mg/dL (≤ 2.2 - 3.0 mmol/L)	Neurological symptoms (22,25)
	Hyperglycemia - Grey zone	100 – 125 mg/dL (5.5-6.95 mmol/L)	Impaired fasting glucose (29)
	- Diabetes	≥ 126 mg/dL (≥ 7.0 mmol/L)	(29)
CALCIUM (total)			
<i>Reference Intervals Adults</i>		8.8 – 10.6 mg/dL (2.20 – 2.65 mmol/L)	(31)
<i>Decision Limits</i>	Mild Hypocalcemia	< 8.0 mg/dL (2.0 mmol/L)	Associated with risk of convulsion
	Hypocalcemic crisis	< 6.0 – 7.0 mg/dL (< 1.50 – 1.75 mmol/L)	(22,32)
	Mild Hypercalcemia	<12 mg/dL (<3.0 mmol/L)	Associated with primary hyperparathyroidism etc...
	Hypercalcemic crisis	≥ 12.0 - 14.0 mg/dL (3.0 - 3.5 mmol/L)	Associated with other symptoms. Most common causes are hyperparathyroidism and tumor associated hypercalcemia (22)
POTASSIUM			
(heparin plasma)			
<i>Reference Intervals Adults</i>		3.6 – 4.8 mmol/L	(33)
<i>Decision Limits</i>	Hypokalemia	3.0 – 3.5 mmol/L	If normal cardiac function, generally does not cause any cardiac problems
	Life-threatening Hypokalemia	< 2.5 – 3.0 mmol/L	Causes clinical symptoms; May be associated with cardiac arrhythmia (22)
	Hyperkalemia	≥ 5 mmol/L	Sign that the regulatory mechanism for potassium balance is failing. Cardiovascular and neuromuscular symptoms
	Life-threatening Hyperkalemia	≥ 6.0 - 6.5 mmol/L	(22)