World Journal of Pharmaceutical and Life Sciences
WJPLS
www.wjpls.org
SJIF Impact Factor: 6.129

ESTABLISHING REFERENCE INTERVAL (RI): AN INTRODUCTION

R. V. Mahato*[1], M. Lamsal[2] and A. M. Dutta[3]

[1]Tribhuvan University, Institute of Medicine, Ayurveda Campus and Teaching Hospital, Kathmandu, Nepal.
[2]B.P Koirala Institute of Health Science, Department of Biochemistry, Dharan, Nepal.
[3]Assam Down Town University, Guwahati, Assam, India.

Corresponding Author: R. V. Mahato
Tribhuvan University, Institute of Medicine, Ayurveda Campus and Teaching Hospital, Kathmandu, Nepal.

Article Received on 24/08/2020 **Article Revised on 14/09/2020** **Article Accepted on 04/10/2020**

ABSTRACT

Reliable and accurate reference intervals (RIs) for laboratory analyses are an integral part of the process of correct interpretation of clinical laboratory test results. RIs given in laboratory reports have key role for a clinician in interpreting test results in reference to values for healthy populations. In Nepal reference interval used in laboratories have been established in the western population. It is not appropriate to use RIs derived for other population due to difference in genetic load, lifestyle and diet. This review highlights the approach for establishing reference interval in local population using the IFCC guidelines.

KEYWORDS: common reference intervals; reference values; genetic information, reference individual, selection criteria, normal values.

INTRODUCTION

The term ‘reference interval’’ was introduced by Grasbeck and Saris in 1969,[1] in response to growing awareness, expressed with great clarity in a paper from Schneider that the concept of normal range, then conceived, was flawed. Current practice at the time is to compare patient results with an ill defined, or at least inconsistently defined, range of values (called the "normal range") derived from an ill defined population of supposedly "normal" meaning healthy individuals.[2]

Grasbeck and Fellman introduced reference interval through publishing a paper 'Normal Values and Statistics' was very first study in the field of reference intervals (RIs) in twentieth century, this was trailed by an introduction by Grasbeck and Saris on ‘‘Establishment and use of normal values.[3] In consequent years it was understood that the terminology of ‘normal values was not suitable and even mostly incorrect so the term 'reference values came into utilization. IFCC published a group of papers between1987-1991 and suggested that every clinical laboratory must pursue defined strategies and produce its own reference intervals.[4-8]

As is the case for all scientific data, the clinical laboratory test result has no value in isolation. There needs to be some control, standard or reference value for comparison. Comparison is as fundamental to clinical medicine as it is to any other scientific discipline. When doctors note clinical signs and symptoms during clinical examination and interview, they consciously or subconsciously make reference to a database of signs and symptoms associated with disease for comparison with those presenting in their patient. Similarly, interpretation of a laboratory test result is a process of comparison.

The kind of reference utilized for correlation relies upon the idea of the clinical inquiry being posed of the laboratory test. For instance, if the test is being utilized to screen a particular disease process, previous test results from that patient may be the most proper reference for correlation; sequential concentration of blood tumor markers to survey response to malignancy treatment.[9]

Some clinical laboratories tests are utilized not for finding or monitoring but to settle on explicit clinical choices. For instance, estimation of serum cholesterol is frequently utilized for appraisal of cardiovascular ailment risk and to decide whether cholesterol-bringing down guidance/drugs are demonstrated. In such conditions a specific concentration of the analyte, known "**decision limit**", must be characterized.[10] Some clinical laboratories tests are utilized to monitor drug treatment. Here patient outcomes are compared with "therapeutic range",[11] which characterize the range of drug concentration in blood consistent of maximum therapeutic and minimum adverse effect.

A number of tools designed for assessment (interpretation) of patient test outcomes, the most widely utilized are the population-based **"health-associated"** reference interval. This is generally implied by the shortened unqualified term "**reference interval**", which is the main focus of this article. Alternative commonly used terms such as **"reference range", "normal range"** and *"***expected values"** are considered improper terms, although they do serve the useful purpose here of broadly conveying what is meant when we use the correct (expertly based), but maybe less familiar term **"reference interval".**

Concept of The Reference Interval

In most clinical circumstances when a specialist is confronted with a laboratory test result of a patient, he/she will presumably first like a response to the fundamental inquiry: "if this specific patient were in good health, would this test result is the same?" A conclusive response to this inquiry is not possible because it depends at the very least on an objective definition of health and test results from the patient when in a state of *"***good health***"*, both of which are deficient.

Although a conclusive answer is preposterous, the reference interval is intended to provide the most ideal answer, and the *"correctness"* of the response depends on the quality or *"goodness"* of the reference interval. A "decent" reference interval is one that, when applied to the local population, includes majority of the subjects with uniqueness similar to the reference group.[12] Good *"health-associated"* reference intervals, with a clinically satisfactory degree of statistical probability, include each one of those from the reference population, who are healthy with respect to the particular tests being considered and exclude all those with disease for which there is relationship with the estimation being considered.

Clinical students and medical laboratory staff were preferential subjects for the derivation of reverence interval this option being born into the world of convenience rather than any real scientific belief or proof of representation of the population with which to be compared. The assumption contained in the term "normal range" that medical students, laboratory staff or any other chosen "normal" population are healthy and generally unchallenged.

Pretty apart from supposed lack of statistical inflexibility deployed in deriving and utilizing normal ranges, the term "normal range" itself was considered inaccurate and ambiguous, because "normal" has several meanings: statistical, epidemiological and clinical.[13] Statistical use of the expression "normal" infers that values (e.g. serum sodium, cholesterol, albumin, etc.) are disseminated in the population in accordance with the hypothetical bell-shaped, perfectly symmetrical curve, known as "Normal" or "Gaussian" distribution (figure 1).For some analytes it was observed that distribution near to normal distribution, but that is by no means always the case and for many analytes the distribution curve is skewed, either to the left or right (Figure 2).

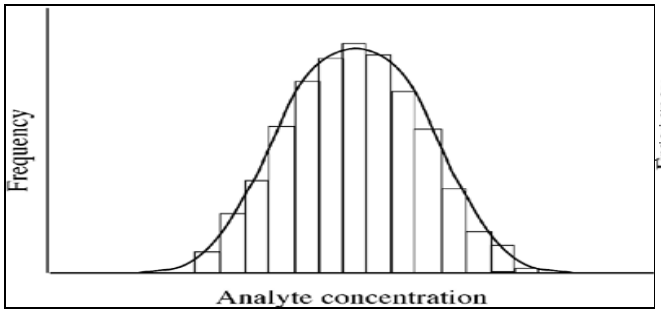

Figure 1: Normal (Gaussian) distribution of analyte

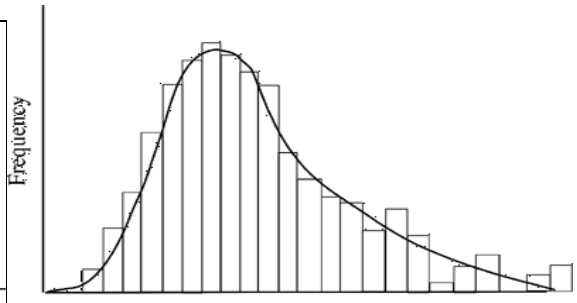

Figure 2: Skewed (non-Gausian) distribution of analyte.

From an epidemiological perspective it may be "normal" for serum cholesterol to be greater than 5.5 mmol/L, but from a clinical viewpoint it certainly is not normal for healthy. In short, "normal range" is an contradictory term, incompatible with the logical rigor required for development of the most precise interpretive tool. Ultimate objective of introducing scientific rigidity, a clear unambiguous definition of terms for a unifying concept of reference intervals was required and in 1986, after much expert forethought and consultation, the International Federation of Clinical Chemistry (IFCC) agreed on a set of definitions,[14] that continue to strengthen the theory and practice of reference intervals today.

IFCC, Terminology related to reference interval

Reference individual: Healthy subjects selected for comparison using defined criteria.

Reference population: Comprises of all probable reference individuals. It usually has an indefinite number and is therefore a imaginary entity.

Reference sample group: An adequate number of reference individuals taken to represent the reference population. Preferably they should be arbitrarily drawn from the reference population.

Reference value: The test result obtained by examination or measurement of a particular quantity on an individual belonging to a reference sample group.

Reference distribution: It is statistical distribution of reference values. Assumption regarding reference distribution obtained from a reference population can be tested using the reference distribution of the sample group and adequate statistical methods.

Reference limit: It is derived from the reference distribution and is used for explanatory purposes. It is common practice to define a reference limit in order that a stated fraction of the reference values is smaller than or equal to or more than or equal to the respective upper or lower limit. A reference limit is descriptive only of reference values and will not be confused with the term "decision limit".

Reference interval: It is the interval between and including two reference limits. The term "reference range" was rejected because strictly (statistically) speaking range is that the difference between the very best and lowest value; it is one value.

Observed value: Patient test result is the value of a particular type of tests obtained by either observation or measurement and produced to make a medical decision. This might be compared with reference values, reference distributions, reference limits or reference intervals. The working relationship between these terms is described in fig 3.

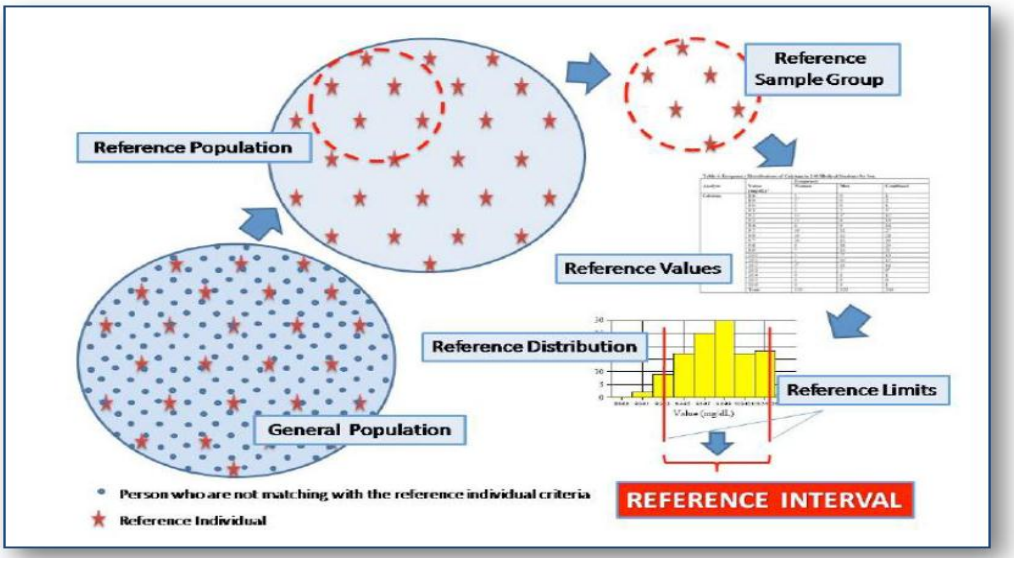

Fig. 1: Working relationship between IFCC defined terminology.

Reference Interval can be derived in following four steps

1) Defining the reference population

The IFCC-recommended use of the term "reference population" which does not define or describe the reference population. For example, presence of health is not implied, allowing the derivation of reference intervals for both the healthy and the sick. Defining the reference population is vital for the preparation of effective reference intervals. This definition must be supportive in clear understanding of how the reference interval is to be used, which in turn must be supporting the measurant in question as regards, for example, its pathophysiological significance and biological variance. Clearly, for "health-associated" reference intervals the reference population must be healthy but there are other considerations, the most significant are age and gender while ethnicity, socioeconomic and life style factors may in some circumstances be significant. The vital point is that the reference population should be an acceptable "control" for patients, having due regard for the way in which the test result is to be used.

2) Selection of reference individuals

Normally, the reference sample group should perfectly reflect the reference population. It can be achieved if reference individuals are selected randomly from the reference population. Since random selection demands that every member of the reference population has an equal chance of being selected out of thousands or millions. For the derivation of "health-related" reference intervals, reference individuals must be healthy, but health may be a relative concept, difficult to define and even hard to pin down individuals.[15] For example, adults may be suffering from latent or subclinical disease (e.g. atherosclerosis) although they may well be in apparent good health. A subjective feeling of good health ("I feel fine") is not guarantee of healthy status. Given that it's difficult to define health in any meaningful way, the standard pragmatic solution is to aim to exclude all those with disease and may be those with an unhealthy lifestyle.

To this end, exclusion criteria for the selection of reference individuals might include: current illness, recent hospitalization, use of prescription or **recreational**

drugs, obesity, smoking habit, raised vitals etc. Whatever the exclusion criteria applied to select reference individuals, these will vary with the pathophysiological significance of the analyte concerned; they need to be appropriate and justified. For example, past history of jaundice could be considered an appropriate exclusion criterion when determining a reference interval for plasma bilirubin but probably should not be considered appropriate (necessary) if the objective is to derive reference interval for plasma sodium. Other inclusion/exclusion criteria (e.g. age, gender ethnicity, etc.) might be applied to make sure that reference individuals have thus far as same characteristics as those of the defined reference population.

Apart from qualitative considerations to select reference individuals, it's important to think about the size of the reference sample group. The greater the sample size, the greater is the statistical confidence, that the derived reference interval is "true" reference interval for that population. An absolute minimum of 40 samples is required to compute a reference interval which includes 95 % from the mid range of data set and excludes 2.5 % at either end of the range.[16]

The IFCC recommends that a reference sample size should not be less than 120 individuals. This is the minimum number needed to calculate the 90 % confidence limits of a 95 % reference interval determined by non-parametric statistics.[17,18] Larger numbers of reference individuals (up to 700) are required if the analyte being considered displays particularly marked skewness.[12,18] It may be considered necessary to partition a reference group with regard to age or perhaps sex in order to provide age- or gender-specific reference intervals.[19] In such cases each partitioned population should comprise a minimum of 120 individuals.

3. Measurement of analytes in reference individuals

Having selected a reference sample group of adequate size, attention turns to measurement of the particular analytes under study, within the selected reference individuals. An important consideration here is that the reduction of unnecessary or avoidable variation.[20] This reduces the "biological noise" of a reference interval, making it more likely that the "biological signal" of disease in patient samples is detected. Variability often considered under two headings: pre-analytical and analytical variation.

Pre-analytical variability is further divided into in vivo variability beaus of biological factors, and in vitro variability due to non-biological factors. Biological factors which affect analyte concentration include: Sample type, chronobiological rhythms (daily, weekly, monthly, seasonal), fasting, time since last food, posture (standing, sitting, lying), recent exercise and use of tourniquet during sample collection. In vitro variability concerned to sample collection and handling. The factors of interest here include, hemolysis, sample container type, preservatives in sample container, time duration between sample collection and centrifugation /analysis and sample conditions.

The study required for the derivation of reference intervals needs consideration of all possible pre-analytical sources of variations and an assessment of their individual significance for the analytes under study. This enables production of a selected protocol that defines reference-individual preparation, timing of sample collection, type of sample, detail of sample collection and sample-handling, etc. In line with the philosophical stance that reference individuals are "controls" for patients, it is essential that the protocol applied to reference individuals must be applied to patients with equal diligence during sample collection and handling.

The methods to derive reference interval should perfectly be matching to that applied to generate patients test results. If not identical, methods must be comparable in terms of precision and accuracy, traceable to a standard.[21] It is in fact important that the analytical variability of observed values is that the same as that of reference values. Reference values should be determined by analyzing samples alongside patient samples. Samples should be analyzed in several batches to measure analytical variability over time (between-batch variability).

4. Statistical examination of measured data

In this final section we glance at the reference values generated by measurement in reference individuals are used to derive reference intervals. It is an arbitrary but long-held and widely applied convention that observed values (patient test results) be compared with the reference values truncated to 95 % of observed results that lie in the mid range of the reference distribution.[22,23] The 2.5 % of values at both ends are excluded in order that the two reference limits that outline the reference interval are the values of the 2.5th and 97.5th percentile of the reference distribution.

Reference limits are often estimated by parametric or non-parametric statistical methods.[22] Parametric methods can only be applied to bell shaped distributions, and if the analyte displays skewed (non-Guassian) distribution, reference values must be transformed (e.g. by log transformation) to a log-Gaussian distribution for parametric methods to be applied.[24]

Once normal distribution is confirmed, the mean (x) and standard deviation (SD) of reference values are calculated and these parameters are applied to derive reference intervals. For a Normal distribution, 95 % of values lie within ± 1.96 standard deviations of the mean, in order that the two 2.5 and 97.5 % reference limits are (x–1.96 SD) and (x+1.96 SD) respectively, shown in (Fig. 3).

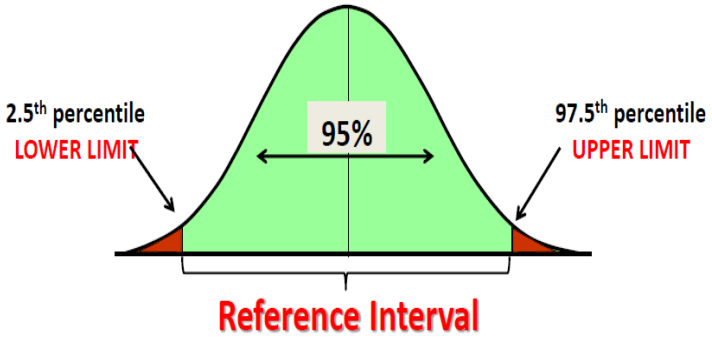

Figure 3: Estimation of reference interval (parametric method).

Non-parametric statistical methods are much simpler and may be applied to data regardless of distribution characteristics. The IFCC-recommended method for estimating reference intervals is a non-parametric method that essentially involves simply excluding the lower and higher 2.5 % of reference values.

It is common practice to calculate the 90% confidence interval (CI) for every of the two estimated reference limits. This means with 90% confidence the interval within which the "true" reference limit would fall if reference values from the entire reference population had been used to estimate it, providing a sign of the reliability of the estimated reference limits.[25]

CONCLUSION

This is a basic review on reference interval which can lay down the foundation to decipher laboratory test outcomes. The IFCC meanings of terms that support the study of reference intervals are featured and some of the problems related with development and utilization of reference values explained. It ideally gives a sound idea for derivation of reference intervals for a population.

REFERENCES

1. Grasbeck R, Saris NE. Establishment and use of normal values.Scand J Clin Lab Invest, 1969; 26 (Suppl 110): 62-63.
2. Schneider AJ. Some thoughts on normal or standard values in clinical medicine.Pediatrics, 1960; 26: 973-84.
3. Gräsbeck R, Fellman J. Normal values and statistics. Scand J Clin Lab Invest, 1968; 21: 193-5. http://dx.doi. org/10.3109/00365516809076984.
4. Grasbeck R, Saris NE. Establishment and use of normal values.Scand J Clin Lab Invest, 1969; 26 (Suppl 110): 62-63
5. Solberg HE. International Federation of Clinical Chemistry. Scientific committee, Clinical Section. Expert Panel on The-ory of Reference Values and International Committee for Standardization in Haematology Standing Committee on Reference Values. Approved recommendation (1986) on the theory of reference values. Part 1.The 84 concept of refe-rence values. Clin Chim Acta, 1987; 165: 111-8. http://dx.doi. org/10.1016/0009-8981 (87)90224-5.
6. Petit Clerc C, Solberg HE.International Federation of Clinical Chemistry (IFCC). Approved recommendation on the theory of reference values. Part 2. Selection of individual for the produ ction of reference value. Clin Chim Acta, 1987; 170: S1-S12. http://dx.doi.org/10. 1016/0009-8981 (87)90150-1.
7. Solberg HE, Petit Clerc C. International Federation of Clini-cal Chemistry (IFCC). Approved recommendation on the theory of reference values. Part 3. Preparation of individuals and collection of the specimens for the production of refe-rence values. Clin Chim Acta, 1988; 177: S3-S11. http://dx.doi. org/10.1016/0009-8981(88)90074-5.
8. Solberg HE. International Federation of Clinical Chemistry (IFCC). Approved recommend- ation on the theory of reference values. Part 5. Statistical treatment of collected reference values. Clin Chim Acta, 1987; 170: S13-S32. http://dx.doi.org/10.1016/0009-8981 (87)90151-3.
9. Perkins G, Slater E, Sanders G *et al.* Serum tumour markers. Am Fam Physician, 2003; 68: 1075-82.
10. Appleton C, Caldwell G, McNeil A *et al.* Recommendation for lipid esting and reporting by Australian pathology laboratories. Clin Biochem Review, 2007; 28: 32-45.
11. Amisden A. Serum concentration and clinical supervision in monitoring of lithium treatment. Ther Drug Monit, 1980; 2: 73-83.
12. Cerriotti F. Pre-requisites for use of common reference intervals. Clin Biochem Rev., 2007; 28: 115-21.
13. Solberg H, Grasbeck R. Reference values. Adv Clin Chem, 1989; 27: 1-79.
14. Solberg H (on behalf of IFCC). Approved recommendation (1986) on the theory of reference values. Part 1 The concept of reference values. Clin Chim Acta, 1987; 167: 111-18.
15. Grasbeck R. Reference values, why and how. Scand J Clin Lab Invest, 1990; 50 Suppl 210: 45-53.
16. Jones R, Payne B. Data for diagnosis and monitoring (Chapter 3) In: Clinical investigations and statistics in laboratory medicine. ACB Venture Publications, 1997.

17. Horn PS, Pesce AJ. Reference intervals: an update. Clin Chim Acta, 2003; 334: 5-23.
18. Linnet K. Two-stage transformations for normalization of reference distributions evaluated. Clin Chem, 1987; 33: 381-86.
19. Harris EK, Boyd J. On dividing reference data into subgroups to produce separate reference ranges, 1990; 36: 265-70
20. Fraser CG. Inherent biological variation and reference values. Clin Chem Lab Med, 2004; 42: 758-64.
21. Koumantakis G. Traceability of measurement results. Clin Biochem Rev, 2008; 29: S61-S66.
22. Solberg H, Grasbeck R. Reference values. Adv Clin Chem, 1989; 27: 1-79.
23. Jones R, Payne B. Data for diagnosis and monitoring (Chapter 3) In: Clinical investigations and statistics in laboratory medicine. ACB Venture Publications, 1997.
24. Peterson P, Gowans EMS, Blaabjerg O *et al.* Analytical goals for the estimation of non-Gaussian reference intervals. Scand J Clin Lab Invest, 1989; 49: 727-37.
25. Solberg HE. Establishment and use of reference values (Chapter 16) In: Burtis CA, Ashwood E, Bruns Tietz Textbook of Clinical Chemistry and Molecular Diagnostics (4th Ed) Saunders, 2005.