How to do it

Determination of a reference interval in a population

VIDYA MAVE, VANDANA KULKARNI, RENU BHARADWAJ, MEDHA KHANDEKAR, AMITA GUPTA, NIKHIL GUPTE

ABSTRACT

The reference interval is the most widely used medical decision-making tool that separates healthy from diseased individuals. We briefly discuss the methods used to determine reference interval and its limitations.

Natl Med J India 2012;25:33–34

INTRODUCTION

The theory of reference interval (RI) was first put forth by Schneider in his 1960s paper entitled 'Some thoughts on normal, or standard, values in clinical medicine'. He states: '… practical medicine is basically founded on comparison. If medicine is to be scientific, we must not only understand the structural, functional and chemical relations operating in individuals, but we must also understand the basis of our comparisons.'[1]

The RI is the most widely used medical decision-making tool.[2] It is central to the determination as to whether or not an individual is healthy. It is generally used as a health-related term to determine the statistical probability of having a specific disease when values fall within the RI. Those with a value outside the RI have a higher statistical probability of having the disease or at least the observed value is not normal for a healthy person. It is an interval that, when applied to the population serviced by the laboratory, correctly includes most of the subjects with characteristics similar to the reference group to be identified as 'healthy' and excludes the others.[3]

ESTABLISHING THE REFERENCE INTERVAL

The National Committee for Clinical Laboratory Standards (NCCLS) now called Clinical and Laboratory Standards Institute (CLSI) has defined a scheme for the determination of RI, which is as follows: Reference individuals comprise a reference population from which is selected a reference sample group on which are determined reference values on which is observed a reference distribution from which are calculated reference limits that may define RIs.[4]

Several methods are used to establish the RIs. However, most methods need healthy individuals as reference individuals. The standard approach uses 120 healthy individuals from a reference sample group. Stratification by age and gender are usually the minimum prerequisite. The values obtained from the sample group are then analysed by non-parametric methods to determine the 2.5 and 97.5 percentiles that form the 95% RI. However, this method of calculating RIs has some disadvantages. It does not allow exclusion of extreme values from the healthy group. Another method to overcome this is called truncation method, which allows removal of extreme values and uses the 80% confidence interval as the reference range. The most scientific method, also called the robust statistical method, estimates the centre of distribution and gives lower points to extreme values. In any analysis of RI data, a visual plot of all the data, such as histogram, a stem-and-leaf plot and/or a boxplot should be done first with the consideration that skewed values may be caused by diseased individuals in the sample.[2]

The best approach to determine the reference range is by a multicentre RI study[5–11] by using samples from different centres consisting of varying distribution of age, gender and race. However, this approach requires separate studies and therefore involves a large amount of time and money. Thus, it is often not feasible; so most laboratories do not derive their own RIs and rely on manufacturer's values. In order to use vendor-published RIs, one must verify that these values are applicable to the specific local population the laboratory serves. This is often termed transference of reference values.[2]

The transference of reference values, a well-known method to determine if one can use the manufacturer supplied RI, measures the analytes on 20 healthy people and compares the values with the manufacturer's 95% interval. If three or more of the values lie outside the interval, then one cannot use the manufacturer's RI.[12] If all tested values fall within reference ranges with a greater number of samples (40 or more), then this should raise concern that manufacturers range is too wide.[2]

The paediatric reference ranges are generally a transference of reference ranges without verification. As per CLSI, there are different ways for a laboratory to validate the transference of established RIs. Paediatric RIs often adapt this approach because of the difficulty in obtaining sufficient specimens to establish or verify RIs. If a laboratory wishes to transfer an RI established by another laboratory, manufacturer's or published ranges, the acceptability should be assessed on the basis of the following factors: similarity of geographics and demographics, similarity of test methodology, and considering clinical judgement of local medical professionals. The approval by the laboratory medical director is required and must be documented. The documentation needs to include at least: (i) the source and reasons for range

Byramjee-Jeejeebhoy Medical College, Pune, Maharashtra, India
VIDYA MAVE, VANDANA KULKARNI, RENU BHARADWAJ, MEDHA KHANDEKAR, AMITA GUPTA, NIKHIL GUPTE
Clinical Trials Unit

Johns Hopkins School of Medicine, Baltimore, USA
VIDYA MAVE, AMITA GUPTA, NIKHIL GUPTE

Byramjee-Jeejeebhoy Medical College, Pune, Maharashtra, India
RENU BHARADWAJ, MEDHA KHANDEKAR

Correspondence to VIDYA MAVE, Pathology Museum, 1st Floor, BJ Medical College, Jai Prakash Narayan Road, Pune, Maharashtra 411001; *vidyamave@gmail.com*

adoption for each analyte, and (ii) a written plan of review over time of the continued appropriateness of the adopted ranges.[13]

The CLSI working group encourages laboratories to focus on verifying the established RIs and to use computerized software procedures to permit increased precision and less stringent sample requirements, which will help laboratories capable of having their own population-based reference ranges to establish reference ranges.[14]

It is also critical to employ standard infection control precautions while working with the samples used to determine reference ranges as it is not possible to know what specimens may be infectious or associated with harm to personnel handling the specimens. An example of a comprehensive guideline is available from the US Centers for Disease Control and Prevention.[14]

STATISTICAL CONSIDERATIONS FOR REFERENCE RANGE DETERMINATION

An important component of design of reference range intervals is the sample size. Sample size is the function of how precise the reference ranges need to be estimated. A 95% CI with a higher sample size (e.g. 2000) will produce an interval that is narrow compared to an interval with a smaller sample size (e.g. 200). Calculating 95% CI is a reliable procedure if the data are normally distributed; if not, some power transformations can be used to normalize the data. Before calculating the confidence intervals, data can be tested for normality using either a chi-square or a Shapiro–Wilk test.[15] If data are not normally distributed, a non-parametric method can be used to calculate a 2.5th and 97.5th percentile range. Variation in the test results can be assessed using multiple regression analysis to explore and account for confounding factors.

LIMITATIONS OF DETERMINATION OF RIs

The processes described above are not easy, fast or straightforward. In transference methods, the laboratory relies on manufacturer's values with the assumption that the manufacturer complied with international standards and that the populations tested are symmetrical to the local population. The challenge is that significant time, effort and money are required to establish RIs and most clinical laboratories are not readily able to modify RIs as this demanding task requires involvement of appropriate ethical approvals and recruitment of clinicians and patients to obtain specimens.[3]

Defining the population for establishing the RI is another limitation. There are several well-defined populations such as the National Health and Nutrition Examination Survey (NHANES) and Fernald population.[16,17] However, these data are from a sample of the US population and may not be generalizable to populations in other countries. Another issue is whether or not one should derive separate RIs for different demographic groups, though it is generally advised.[2]

There are several bad practices for the estimation of RIs. A common practice is to compile an RI without visually reviewing the data. Another current false practice is to use mean+standard deviations without establishing the normal distribution of values. In paediatrics, it is not possible to obtain a healthy group so the test group includes both healthy and non-healthy individuals. Generally, in such a situation, truncation methods are used to compute RIs by using arbitrary cut-off for intervals. Lastly, some failures of current clinical practice include: (i) not keeping the RI data with the patient record, (ii) changing analytical measurements without a long-term comparable record, and (iii) using analytical methods which are not comparable to a reference standard.[2]

In conclusion, the determination of reference values depends on the size of the dataset and the method of evaluation. Although multicentre studies are the best approach to estimate reference values, in most circumstances, transference of RI is commonly performed with a small sample validation of the manufacturer's provided values. Large well-defined populations such as NHANES and the Fernald population data exist from which RIs of many analytes can be transferred by conducting a small verification (confirmation) study.

Conflict of interest: None declared

REFERENCES

1 Schneider AJ. Some thoughts on normal, or standard, values in clinical medicine. *Pediatrics* 1960;**26:**973–84.
2 Horn PS, Pesce AJ. Reference intervals: An update. *Clin Chim Acta* 2003;**334:**5–23.
3 Ceriotti F. Prerequisites for use of common reference intervals. *Clin Biochem Rev* 2007;**28:**115–21.
4 National Committee for Clinical Laboratory Standards. *How to define and determine reference intervals in the clinical laboratory: Approved guideline, NCCLS document C28-A and C28-A2, 1995.* NCCLS:Pennsylvania, USA; 2001.
5 Ferré-Masferrer M, Fuentes-Arderiu X, Gomà-Llongueras M, Alumà-Trullàs A, Aramendi-Ramos M, Castaño-Vidriales JL, *et al.* Regional reference values for some quantities measured with the ADVIA Centaur analyser: A model of co-operation between the *in vitro* diagnostic industry and clinical laboratories. *Clin Chem Lab Med* 2001;**39:**166–9.
6 Fuentes-Arderiu X, Ferrer-Masferrer M, González-Alba JM, Villarino-González MI, Arrimadas-Esteban E, Cabrero-Olivé D, *et al.* Multicentric reference values for some quantities measured with the Elecsys 2010 analyser. *Clin Chim Acta* 2001;**304:**143–6.
7 Fuentes-Arderiu X, Mas-Serra R, Alumà-Trullàs A, Martí-Marcet MI, Dot-Bach D; Catalan Association for Clinical Laboratory; Sciences. Guideline for the production of multicentre physiological reference values using the same measurement system: A proposal of the Catalan Association for Clinical Laboratory Sciences. *Clin Chem Lab Med* 2004;**42:**778–82.
8 Bäck SE, Nilsson JE, Fex G, Jeppson JO, Rosén U, Tryding N, *et al.* Towards common reference intervals in clinical chemistry: An attempt at harmonization between three hospital laboratories in Skåne, Sweden. *Clin Chem Lab Med* 1999;**37:**573–92.
9 Rustad P, Felding P, Lahti A; Nordic Reference Interval Project 2000. Proposal for guidelines to establish common biological reference intervals in large geographical areas for biochemical quantities measured frequently in serum and plasma. *Clin Chem Lab Med* 2004;**42:**783–91.
10 Hyltoft Petersen P, Rustad P. Prerequisites for establishing common reference intervals. *Scand J Clin Lab Invest* 2004;**64:**285–92.
11 Rustad P, Felding P, Franzson L, Kairisto V, Lahti A, Mårtensson A, *et al.* The Nordic Reference Interval Project 2000: Recommended reference intervals for 25 common biochemical properties. *Scand J Clin Lab Invest* 2004;**64:**271–84.
12 Cembrowski GS, Juco JW. Establishing and verifying reference intervals. *Clin Chem* 1999;**39:**91–108.
13 College of American Pathologists (CAP), Commission on Laboratory Accreditation, Chemistry and Toxicology Checklist, REVISED 09/27/2007.
14 Clinical and Laboratory Standards Institute (CLSI). *Defining, establishing and verifying reference intervals in the clinical laboratory: Approved guideline,* 3rd ed, C28-A3, Vol. 28 No. 30 CLSI: Pennsylvania, USA; 2008.
15 Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika* 1965;**52:**591–611.
16 U.S. Department of Health and Human Services (DHHS). *National Center for Health Statistics. Third National Health and Nutrition Examination Survey, 1988–1994, NHANES III Laboratory Data File (CD-ROM). Public Use Data File Documentation Number 76200.* Hyattsville, MD: Centers for Disease Control and Prevention, 1996. Available from National Technical Information Service (NTIS), Springfield, VA. Acrobat. PDF format; includes access software: Adobe Systems, Acrobat Reader 2.1. Available at *wonder.cdc.gov/wonder/sci_data/surveys/hanes/hanes2/type…/lab.asp*
17 Horn PS, Pesce AJ, Copeland BE. Reference interval computation using robust vs parametric and nonparametric analyses. *Clin Chem* 1999;**45:**2284–5.