MedicalBiostatistics.com

Diagnostic Thresholds of Medical Measurements

Adapted from Medical Biostatistics, Fourth Edition
by A. Indrayan (http://indrayan.weebly.com)
Chapman & Hall/CRC Press, 2018
(https://www.routledge.com/Medical-Biostatistics/Indrayan-Malhotra/p/book/9781498799539)
Details of the book and reviews available at http://MedicalBiostatistics.synthasite.com

Even though normal is the level generally seen in healthy subjects, there are always persons with very high or very low values who are still absolutely healthy. Thus there could be considerable overlap between normal and abnormal values. Determining threshold that worksin all situations has remained an elusive objective. Following approaches are available.

Disease Threshold

The best course to delineate normal levels is to observe people with different levels for a sustained period, and identify a threshold beyond which people start feeling the burden in some sense—not be able to do work to the capacity, or entailing a risk for an adverse condition down the line in life. This is an extremely complex procedure and requires consultation from experts, who in turn should have full evidence for the threshold they propose. The cut-off 140/90 mmHg for BP is such a threshold. Experts have observed that a higher BP considerably increases the risk of coronary artery disease. Not many examples of this type of cut–point are available. But there would be people with level 145/92 who would be healthy and there would be people with level 136/88 and nonhealthy (with complaints such as headache and irritability). Thus even this threshold does not rule out errors. Call this type of threshold as disease threshold of normal level.

Clinical Threshold

The second alternative is to compare levels of those who are in perfect health with those who are not. Since each of these groups will have a distribution of its own, the situation typically will be as in Fig. 1.

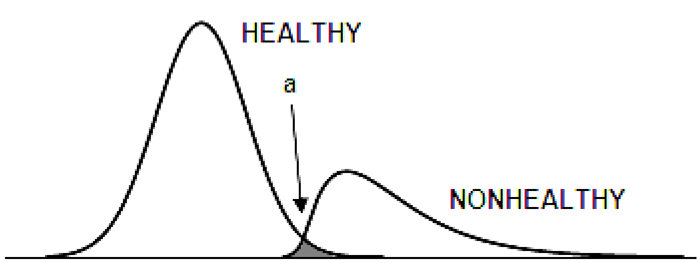

Figure 1 The pattern and overlap of measurement in healthy and nonhealthy subjects

This figure has the following features.

1. The number of healthy persons far exceeds the nonhealthy persons.
2. The variation in the levels is smaller for healthy than for nonhealthy persons. Note how scatteredness in the levels of nonhealthy is relatively large.
3. The shape of the distribution in healthy subjects is Gaussian whereas in nonhealthy subjects is positively skewed.
4. In this figure, nonhealthy subjects have higher levels. This is true for many measurements such as T_3, T_4, BP, blood glucose, etc., but not for all. Higher levels of Hb, peak expiratory flow rate, HDL cholesterol, etc., indicate good health. For these measurements, the curve for nonhealthy will be on the left side.
5. There is some overlap between levels seen in healthy subjects and the levels seen in nonhealthy subjects. This is shown as shaded area in Fig. 1. If there is no overlap, the healthy levels and nonhealthy levels can be immediately defined. In practice, this overlap is substantial and causes problems in defining healthy levels.

Statisticians have shown that the point where the two curves intersect provides the cut-off with least number of misclassifications. This level is indicated as 'a' in the figure. This is the clinical threshold that could be used to define normal levels. Indeed this is a very convincing approach but can be adopted only when the distribution in the healthy and nonhealthy groups is known and the overlap is minimal. The biggest problem in this approach is the choice of criterion to categorize a person as healthy or nonhealthy for drawing these curves. Threshold obviously will not be known without the curves and the curves cannot be drawn without categorizing subjects as healthy and nonhealthy. Obviously external criteria are needed, and those may or may not work. Nonetheless, such clinical threshold has in-built provision for tolerating error of misclassification as indicated by the shaded area. Errors are not ruled out by this method also. Bigger the overlap, larger is the shaded area and higher is the chance of error.

Statistical Threshold

When the distribution is Gaussian, its popular property is invoked to say that (mean – 2SD, mean + 2SD) are the normal limits. They exclude nearly 2.5 percent of healthy subjects with extreme measurements on either side. This is arbitrary but now accepted around the world. The mean and SD are computed from measurements obtained on a large number of healthy subjects. These are statistical thresholds and popularly known as ±2SD limits. Most of the normal ranges used in medical practice are obtained in this manner. If mean fasting blood glucose level in healthy subjects is 90 mg/dL and SD = 7.5 mg/dL, normal range is 90 – 2×7.5 to 90 + 2×7.5, or 75 to 105 mg/dL. When the distributional shape is far from Gaussian, the range from (2.5)th to (97.5)th percentile points is considered normal instead of ±2SD limits. Note that the ±2SD limits for the Gaussian distribution are also from (2.5)th to (97.5)th percentile. You can therefore forget about ±2SD limits and use the percentile-based range for all measurements irrespective of the shape of the distribution. But ±2SD limits are ingrained in the minds of many clinicians and statisticians alike. One reason for this is that ±2SD limits fit well into the confidence interval and testing hypothesis strategy that are so commonly used. Note the following for such statistical thresholds.

1. No matter how healthy subjects are, there are always 2.5 percent healthy subjects at the lower end and another 2.5 percent at the upper end who will have levels outside such normal range. This is an error but is tolerated because an error of this magnitude may always occur irrespective of the method used to establish normal limits. This error is at least quantitatively known for statistical thresholds, but would not be easily known in other approaches.

2. The ±2SD limits are purely statistical. A level beyond these limits is abnormal only in the sense that such an extreme level is rare in healthy subjects. Whether this translates to medical problems is not known. However, the limits seem to be working well as an aid in most practical situations.

3. A measurement such as 106 mg/dL for fasting blood glucose level is not abnormal when the normal range is from 75 t o 105; just that the chance of this person being normal is small—less than 2.5 percent. Gaussian theory stipulates that this chance reduces steeply as the measurement becomes farther and farther away from mean. No miracle happens at the cut–off, such as 105, that would suddenly make a measurement abnormal. Nonetheless such cut–off is needed somewhere as a guideline to start suspicion. (Mean ± 2SD) provides such cut-off. But it is applicable to one type of measurement at a time. If there are five different types of measurements such as different components of lipid profile, the chance of a healthy person labeled as healthy by such statistical criteria for all measurements together is not large.

4. As a corollary to the point 3, a measurement in a person beyond ±2SD limits does not necessarily indicate a pathological condition. A health problem may or may not exist. A complete picture is available when the person is considered in totality along with signs and symptoms, if any.

5. Some disease entities are based almost exclusively on a single parameter. Diagnosis of iron deficiency anemia is based on hemoglobin (Hb) level, hypertension on BP levels, diabetes mellitus on serum glucose level, and glaucoma on intraocular pressure. Other indications such as signs-symptoms play a minor role for classifying such diseases. Evidence exists that persons with statistically abnormal levels do have an increased risk of the concerned morbidity and mortality. An intervention, such as therapy, to bring the level back to the normal range helps to reduce this risk.

6. The normal levels, whether statistical ±2SD limits or based on the healthy-sick dichotomy, should be determined by measuring a large number of subjects. Only then they command confidence. Normative data based on small samples can at best be indicative for confirmation in subsequent testing. Small sample based 'normals' can seldom be used for diagnostic or prognostic purposes.

7. The risk of misdiagnosis and missed diagnosis seems to be universally present irrespective of the procedure used to delineate reference values. As already explained, if the reference values are not statistical but based on values actually present in healthy and diseased subjects, then also some overlap is inevitable. If the diagnosis is based on clinical signs and symptoms instead of the value of a single parameter, there will also be cases with a nontypical picture. Even a composite picture jointly obtained by several

measurements, investigations, and signs and symptoms can turn out false in some cases. As the information on a patient increases, the risk of error decreases, but it would rarely vanish at the diagnosis stage. This is where the acumen of a clinician comes in handy. The human brain is always superior to any technological inputs, particularly in the case of medicine. The answer lies in putting together the pieces of a jigsaw puzzle in as efficient a manner as possible.