How Many Measurements to Make a Decision?

Pascal Houillier,*†‡ Joël Coste,*§ and Marc Froissant*†||

*Department of Physiology, Paris Descartes University, Paris, France; †Department of Physiology, Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris, Paris, France; ‡INSERM, Centre de Recherche des Cordeliers, UMRS 872, Paris, France; §Biostatistics and Epidemiology Unit, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, Paris, France; and ||Department of Biostatistics and Epidemiology, INSERM, Unit 1018, Villejuif, France


Every day, nephrologists rely on biological values to diagnose kidney disease or complications, to make the decision to treat or not to treat, and to assess the efficacy of treatments. At first glance, a biological criterion seems easy to use because the result of the measurement is almost always quantitative and provided with a reference interval that helps in the recognition of whether a given result falls within or outside the “normal range” and how far it falls outside the range. More difficult is to determine, during a patient’s follow-up, whether a biological value significantly differs from the previously measured value in the same patient.

The reality is much more complex because of the inherent variability of biological values. Laboratory analytes are subject to three main sources of variation: Preanalytical, analytical, and biological. Preanalytical variation can be minimized by standardizing sampling procedures and handling of specimens. A recent study (1) revealed that the nature of the blood specimen type (plasma or serum) can induce a large variability in parathyroid hormone (PTH) concentrations; consequently, it is advisable to use one specified collection method to reduce the variability in PTH measurement.

Analytical variation can be a major source of variation in biological analytes. Souberbielle et al. (2) compared the PTH concentrations measured with 15 commercial immunoassays in 47 serum pools from dialysis patients, using the Allegro intact PTH assay as the reference. The median bias between the Allegro intact PTH assay and the tested assays ranged from −44.9 to 123.0%. This important intermethod variability in PTH results owed to both antibody specificity and standardization reasons. The potential deleterious consequences of such a variability are obvious. As with the preanalytical variation, the analytical variation can be decreased by improving the specificity and the standardization of the assay, and the use of standard reference materials has to be promoted whenever available.

By contrast, the third source of variation cannot be reduced. Indeed, the intraindividual biological variation results from random fluctuation around an individual homeostatic set point. This intraindividual variation has been estimated for a number of biological analytes both in healthy individuals and in patients with disease (3) and is available at http://www.westgard.com/biological-variation-in-patients-with-disease.htm. It ranges from 0.7% for serum sodium concentration in healthy individuals to 61% for urinary albumin in first morning urine in patients with diabetes. The intraindividual biological variation is useful in clinical practice, because it allows determination of how many measurements are necessary to know precisely the “true” value of any analyte. Obviously, the lower the intraindividual biological variation, the less the necessary number of measurements. Maybe more important, the intraindividual biological variation is used to calculate the reference change value or critical difference (4): It is required to determine how much any analyte has to vary for the change to be considered significant.

In this issue of the Clinical Journal of the American Society of Nephrology, Gardham et al. (5) studied the variability of calcium, phosphate, PTH, and alkaline phosphatase in stable patients who were receiving hemodialysis and healthy volunteers. They report that the intraindividual biological variation is significantly larger for phosphate than for calcium and is the largest for PTH. Interestingly, every intraindividual biological variation is larger in hemodialysis patients than in healthy volunteers. The number of independent measurements required to determine the “true” value of any analyte increases with the intraindividual biological variation (6,7); therefore, in stable hemodialysis patients, estimation of the true value of blood calcium concentration with a 10% accuracy requires one single measurement. By contrast, because of a larger intraindividual biological variation, four measurements are necessary for alkaline phosphatase, eight for phosphate, and 26 for intact PTH. In addition, because the reference change value directly depends on the intraindividual biological variation, plasma phosphate concentration has to vary by at least 41% and intact PTH by 72% before a change can be considered significant, whereas a change of only 8% in calcium concentration can be viewed as significant.

For practical considerations, it is obvious that a decision to treat secondary hyperparathyroidism or hyperphosphatemia cannot be made on the basis of a single abnormal value of those analytes. By contrast, the results of this study underline that
both the number of consecutive abnormal values and the extent to which they differ from the normal range should be considered in the process of making a decision to treat.

Such a concept should be kept in mind and extended in nephrology practice and research to other application fields than mineral metabolism. Assessing renal function is one of them. Apart from discussion on predictive performance of estimated creatinine–based GFR formulas, which by nature have to handle creatinine measurement variability, these concerns are to be pointed out on what we call “gold standards,” or reference measurements. For instance, measured GFR variability, rarely available among publications, has been characterized using coefficient of variation as 11.7, 9.4, and 8.4% in the Diabetes Control and Complications Trial (DCCT), the Modification of Diet in Renal Disease (MDRD) study, and the NephroTest CKD cohort study, respectively (8,9). Such GFR measurement variability has to be taken into account when evaluating estimated GFR performance by the use of biostatistics methods managing both errors on estimates and reference values. Similarly, hemoglobin variability in response to an erythropoiesis-stimulating agent regimen should take into account hemoglobin determination variability in the analyses.

These problems that actually concern many medical measurements, either clinical (e.g., BP, tuberculosis skin test) or biological (e.g., blood glucose, hormones, viral markers), have been known for a long time. It is regrettable that they are so promptly forgotten and so rarely rediscovered. The article by Gardham et al. (5) has the merit to help nephrologists think about the medical meaning of an “abnormal” value before making a decision.

Disclosures
None.

References