Reference Service. In all six cases, the patients were referred to an oncologist or gynecologic oncologist with a suspected diagnosis of trophoblast disease or postgestational choriocarcinoma.

Four of the six cases received multiple courses of methotrexate chemotherapy, and one of the six received in addition EMACO chemotherapy, with no major quantitative reduction in false-positive hCG results. After chemotherapy, two patients underwent a hysterectomy and one other patient underwent an oophorectomy, all without a major reduction in measured hCG concentrations. One patient developed type 1 diabetes as a complication of the chemotherapy and became comatose. All therapies came to a halt with the finding by the hCG Reference Service that the persistent hCG results, the sole basis for treatment, were in fact false-positive or “phantom” hCG.

We have now heard that in two of our earliest phantom hCG cases (tested in Spring 1998), the false-positive hCG results eventually, after 10–14 months, subsided. The cases came from all parts of the United States (one from the West Coast, three from the Midwest, and two from the Northeast). These cases of phantom hCG found their way through word of mouth to the hCG Reference Service, a new, unadvertised facility, in a 9-month period. We wonder how many other similar cases may exist.

Four of the six false-positive cases had been detected and followed with the Abbott Diagnostics AxSym hCGb test (Table 1), and one of six had been followed with the sister assay, the Abbott Diagnostics IMX hCGb test, which uses the same antibody/chemical set. Thus, five of the six cases were detected with this one type of assay. We do not know if this type of assay, among the >40 quantitative hCG test sold in the US (4), is particularly prone to false-positive results.

Laboratory directors and managers need to be aware of this potential problem, especially if they are performing the AxSym or IMX hCGb type test. They need to be available to help physicians quickly exclude or identify phantom hCG, which can be done by simply running quantitative urine hCG tests. In phantom hCG cases, no hCG immunoreactivity may be detected (<5 IU/L) in urine samples. Other simple methods to exclude phantom hCG are to test serum samples with competitive hCGβ RIAs (which do not detect phantom hCG), or to demonstrate nonlinearity in dilutions of the serum samples in an hCG immunoassay (phantom hCG may give grossly nonlinear results). Alternatively, help can be sought from the hCG Reference Service.

References

Laurence A. Cole*
Kirsi M. Rinne
Shohreh Shahabi
Aziza Omrani
hCG Reference Service
Department of Obstetrics and Gynecology
Yale University
New Haven, CT 06520

*Author for correspondence. Fax 203-785-6367; e-mail laurence.cole@yale.edu.

Limitations of the Paired t-Test for Evaluation of Method Comparison Data

To the Editor:
In recent years, the difference or bias plot for evaluation of method comparison data has become increasingly popular. Originally suggested by Bland and Altman for comparison of measurements in clinical medicine, the procedure also has been adopted in clinical chemistry (1–3). The difference plot is very instructive for the display of differences as functions of the measurement average. In addition to the graphical display, however, it is usual to present some form of summary statistics for a method comparison study. In association with the difference plot, the paired t-test is usually applied (1).

The paired t-test is ideal for evaluation of a constant difference between two sets of values (4, 5). When it is used to analyze other types of differences, however, problems may arise.

For example, consider the case shown below, in which y measurements tend to exceed x measurements in the low range, and vice versa in the high range (Fig. 1). The actual data set of n = 50 (x, y) measurement sets were generated as a random sample based on the relationship y = 20 + 0.8x between the true values (target values), with added measurement errors corresponding to analytical SDs of 5 for both x and y (CV of ~5% at the mean of 100). The y target values were assumed uniformly distributed on the interval (25, 175). In this situation, the overall averages of both sets of measurements are nearly identical, and the paired t-test yields a nonsignificant result because the average paired difference is close to zero: mean of x values, 101.8; SD, 43.8; SE, 6.2; mean of y values, 100.1, SD, 35.4; SE, 5.0; mean of paired (y − x) differences, −1.7; SD, 10.9; SE, 1.5; paired t-test, t = −1.7/1.5 = −1.1 (not significant).

Thus, this test is unsuitable for characterization of the measurement relationship in the present situation, which may arise frequently in the context of method comparison studies. Rather, subjecting the data to a type of regression analysis (e.g., the Deming approach) clearly discloses the relationship (6): slope (b), 0.81; SE, 0.026; test against 1.00, t = (0.81 − 1.00)/0.026 = −7.4 (P < 0.001); intercept (a0), 18; SE, 3.1; test against zero, t = (18 − 0)/3.1 = 5.7 (P < 0.001).

The results of the regression analysis confirm the existence of both a sys-
tematic constant difference (intercept different from zero) and a systematic proportional difference (slope different from 1). Therefore, the paired t-test should not be applied uncritically to method comparison data. Only when the graphical display suggests that a systematic constant difference, but not a systematic proportional difference, is involved should this test be applied. With this background, it appears surprising that a clinical chemistry journal has directly prohibited the use of regression analysis in method comparison studies, a point of view also expressed in another journal (7, 8). Opposition against this practice has previously been put forward (9).

References

Kristian Linnet
Laboratory of Clinical Biochemistry
Psychiatric University Hospital
DK-8240 Risskov, Denmark
E-mail linnet@post7.tele.dk

Interlaboratory Variability for Total Homocysteine Analysis in Plasma

To the Editor:
Total plasma homocysteine consists of free homocysteine and homocysteine that is complexed with itself or with other amino acids or proteins. Free homocysteine has been measured previously as part of a biochemical screen for inherited metabolic disorders. More recently, increased total plasma homocysteine has been suggested as an independent risk factor for atherosclerotic coronary artery disease [reviewed in Ref. (1)]. In addition, increased total homocysteine is associated with a poor prognosis in patients with angiographically demonstrated coronary artery disease (2). These studies have prompted clinicians to include total homocysteine analysis as part of the risk assessment profile of patients with premature coronary artery disease. However, it has yet to be shown that reducing total plasma homocysteine concentration leads to a decrease in cardiovascular risk, although vitamin supplementation may effectively lower or normalize circulating homocysteine concentrations (3).

From a laboratory standpoint, problems exist in homocysteine analysis. External quality-assurance programs for total homocysteine (proficiency testing) are not available at present, and interlaboratory correlations of total homocysteine measurements have not been evaluated formally. We have contacted many of the reference laboratories in the US that offer this assay and have found that the reference ranges vary considerably between laboratories. Although most laboratories offer reference intervals based on in-house studies, a few base reference intervals on a review of the literature. Other laboratories offer a “target range” based on prospective studies, which correlate total homocysteine concentrations with risk for cardiovascular disease or with mortality (2).

To evaluate laboratory variability of total homocysteine analysis, we sent five samples of frozen plasma in EDTA tubes, all drawn and pooled from the same fasting subject, to five different reference laboratories. The analytical variation of the results produced a wide range of possible risk estimates for that subject (Table 1).

We evaluated the extent of variability attributable to inherent test imprecision as opposed to bias between laboratories. The variance of the reported homocysteine test results is composed of the sum of variance within the laboratories and variance between laboratories. To estimate variance within the laboratories, we obtained the coefficients of