cule to market and increasing the number of new chemical entities registered each year (3). This also is a major concern for the pharmaceutical industry because research and development costs have increased from $54 million to approximately $600 million per registered molecule during the past 20 years (3).

However, as an indirect consequence of ATG, clinical diagnostics will be forever altered, and there are at least three reasons behind this change: First, one by-product of identifying new targets for drug development is identification of unambiguous genetic and protein markers of disease. After the genome, these new markers will have unparalleled sensitivity and specificity for diagnosing inherited as well as acquired diseases. Furthermore, these markers will be necessary because disease management strategies will be gene-based. In the gene-based era of medicine, diseases will be redifined according their genotype rather than phenotype. Another by-product of drug discovery will be the identification of genetic markers that predispose one to disease. Such predisposition markers will presage the growth of predictive medicine and thereby permit effective screening, early intervention, and disease prevention. Finally, because genetic polymorphisms underlie individual responsiveness to drugs, both the safety and efficacy of drug use will be significantly increased by the use of pharmacogenetics to effectively tailor drug prescriptions to individual genotypes.

What ATG means to clinical diagnostics therefore is becoming clear—many new things to come and the ushering-in of an entirely new season.

References
1. After the genome IV: envisioning biology in the year 2010. 10–14 October 1998; Jackson, WY.

P. Patrick Hess
Quest Diagnostics Nichols Institute
33608 Ortega Highway
San Juan Capistrano, CA 92690
Fax 949-728-4930

False-Positive hCG Assay Results Leading to Unnecessary Surgery and Chemotherapy and Needless Occurrences of Diabetes and Coma

To the Editor:

Much concern has been raised by the unraveling at the hCG Reference Service of six cases of persistent phantom human chorionic gonadotropin (hCG). These are false-positive hCG results, which are likely attributable to human anti-mouse IgG or to heterophilic antibodies (1–3). The hCG Reference Service, started in January 1998 to aid with the interpretation of irregular or discordant hCG immunoassay results, requests parallel serum and urine samples. Each is tested in four separate two-step microtiter plate ELISAs (assay 1 detects intact hCG, assay 2 detects non-nicked or bioactive hCG only, assay 3 detects the hCG free β-subunit only, and assay 4 detects the hCG β-core fragment only) at three different concentrations (undiluted, a 1:2 dilution, and a 1:5 dilution). From the data, inferences are made about the nature (nonnicked or nicked hCG, free β-subunit, and β-core fragment) and likely source (trophoblast disease, pitiary hCG, cancer, or phantom hCG) of the hCG immunoreactivity.

In all six cases of phantom hCG tested by the service, increased hCG concentrations were detected in serum samples (Table 1), but no detectable hCG, free β-subunit, or β-core fragment was found in the parallel urine samples (<3 IU/L). In all cases, the presence of phantom hCG was confirmed by at least two of the following three criteria: the finding of serum concentrations that were nonlinear on dilution; the finding of hCG concentrations in a two-step assay that were 20% or less of values in a one-step (single incubation with both coating and tracer antibodies) sandwich assay or that differed by 80% in two different hCG assays; and the finding of measurable β-core fragment immunoreactivity (not usually detectable in serum) in serum samples.

The six cases had similar histories. Each started with an incidental pregnancy test (Table 1). The pregnancy test was positive (69–285 IU/L), and the patient was sent to an obstetrician. In each case, ultrasound failed to reveal a fetal sac, laparoscopy did not reveal an ectopic pregnancy, and dilation and curettage revealed no recent history of pregnancy or trophoblast disease. In each case, the false-positive hCG persisted for an additional 3–11 months (5–451 IU/L) before samples were sent to the hCG

<table>
<thead>
<tr>
<th>Table 1. Summary of clinical findings and laboratory data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Age, years</td>
</tr>
<tr>
<td>hCG test</td>
</tr>
<tr>
<td>hCGβ</td>
</tr>
<tr>
<td>Initial hCG, IU/L</td>
</tr>
<tr>
<td>Laparoscopy</td>
</tr>
<tr>
<td>Dilation and curettage</td>
</tr>
<tr>
<td>Oophorectomy</td>
</tr>
<tr>
<td>Hysterectomy</td>
</tr>
<tr>
<td>Methotrexate chemotherapy</td>
</tr>
<tr>
<td>EMACO chemotherapy</td>
</tr>
<tr>
<td>Type 1 diabetes and coma</td>
</tr>
</tbody>
</table>

* Patients i, ii, and iii are described in more detail in Cole (3).

^b Range of hCG concentrations in the 3–11 months (depending on case) after initial detection.
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 hCGβ test (Table 1), and one of six had been followed with the sister assay, the Abbott Diagnostics IMx hCGβ 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 hCGβ 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
5. Laurence A. Cole* Kirs M. Rinne Shoher 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 x 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-