Evaluation of Diagnostest Reagent Sets

To the Editor:

In the publication, “Assessment of Two Blood-Analyzer Systems Intended for Diagnostic Testing in the Physician’s Office” [CLIN. CHEM. 16, 990 (1970)], J. E. Logan and M. L. E. Sunderland present data of their evaluation of Diagnostest Reagent Sets manufactured and sold by Dow Diagnostics, The Dow Chemical Co., Indianapolis, Ind. 46206. They conclude that some of the Diagnostest methods give statistically biased results as compared to their “reference” AutoAnalyzer methods, and that inter-run precision was not acceptable for some of the methods at certain concentration levels.

The purpose of this letter is to show (a) that the bias noted by these authors reflects their improper choice of AutoAnalyzer methods as reference methods, and (b) that their comments regarding unacceptable between-run reproducibility for some Diagnostest reagent sets are unduly critical. Our remarks are concerned only with the statements made and conclusions reached by Logan and Sunderland concerning Diagnostest products.

The following observations regarding the day-to-day reproducibility data are pertinent with respect to some of Logan and Sunderland’s conclusions. The acceptable reproducibilities at low and high levels, but not at intermediate levels, for uric acid (Table 1) suggest difficulties other than methodologic inadequacy of the Diagnostest method. Also, the high-level (396 mg/100 ml) study for cholesterol reportedly yielded an unacceptable precision of ±11%, because Tons’ acceptable limit of error is ±10%. An important fact to point out is the lack of details pertaining to their experimental design that yielded the data in Table 1. We refer to the absence of the following important information: number of runs (days), how specimens were stored between runs and for how long, and what commercial control serum was used. Logan and Sunderland adopted Barnett’s (I) method for evaluating the performance of kits. If in fact they followed Barnett’s suggestion of 10 runs on 10 separate days at each of the three levels, we want to emphasize the inadequacy of 10 values for obtaining an accurate measure of precision. Any number less than approximately 30 will not provide a reliable estimate of precision (2).

The question of bias will be considered next since it involves four of the six tests evaluated (Table 4). Logan and Sunderland used AutoAnalyzer methods as “reference” methods for urea nitrogen, uric acid, bilirubin, and cholesterol. Manual reference methods—o-toluidine for glucose and cyanmethemoglobin for hemoglobin—were used for the other two tests. It is pertinent to note that Logan and Sunderland found acceptable the data concerning day-to-day reproducibility, recoveries, and Student’s t test for bias for the glucose and hemoglobin methods. But the methods for urea nitrogen, uric acid, bilirubin, and especially cholesterol were stated to exhibit statistically significant but not necessarily clinically significant bias in either the normal or abnormal range, or both. It is appropriate to consider each of these tests and discuss the reasons for the biases as reported in Table 4.

The urea nitrogen method exhibited no bias in the normal range (4 to 25 mg/100 ml), but a statistically significant bias (low results for Diagnostest) is claimed for the abnormal (26 to 98 mg/100 ml) range. The magnitude of this bias is such that the results appear to be about 9% low. During development of the Diagnostest method, we compared it with a specific urease method as stated in the Diagnostest literature (3). We observed no bias between methods for normal or elevated values. On the other hand, a comparison study with the automated dicetylmonoxime method did reveal higher results by the automated method. Also, in a recent publication by Elser and Savory (4), in which they report on the evaluation of the Diagnostest urea nitrogen kit, they found better agreement between their manual dicetylmonoxime procedure and the Diagnostest method than between the automated dicetylmonoxime procedure and the Diagnostest method. From the above findings, we conclude that the automated dicetylmonoxime method is biased high rather than that the Diagnostest method is biased low.

Logan and Sunderland found no bias for uric acid values in the normal range (3.5 to 5.8), but a statistically significant bias is shown for the values in the 6 to 12 mg/100 ml range. The Diagnostest method is a carbonate-phosphotungstate reduction method, which was compared during its development with a specific uricase method (5). No bias was observed between this accepted reference method and the kit method. Logan and Sunderland, on the other hand, used as their reference the automated method of Crowley and Alton (6), who themselves report that the method yields results that average 0.4 mg/100 ml higher than their reference uricase method. This could account for the bias (higher results for Logan and Sunderland) found.

Their data in Table 4 also show a statistically significant bias between the automated bilirubin method used in their laboratory and the Diagnostest method, although they state that the differences are not necessarily significant clinically. This reference method was the automated version of the Jendrasik–Grob method (7). Several investigators, including Mabry et al. (8), Michælsson et al. (9), and Gambino et al. (10) have stated that this method gives lower total bilirubin results in the higher bilirubin ranges as compared to the manual reference method of Malloy and Evelyn. The Diagnostest method is a modified Malloy–Evelyn method, as stated in the Diagnostest literature, and showed no bias when compared with this reference method (11) during development of the kit. The bias quoted by Logan and Sunderland in the abnormal range (Diagnostest lower by an average 0.4 mg/100 ml) is probably due to the difference between the Malloy–Evelyn and Jendrasik–Grob methods, and is no reflection on the accuracy of the Diagnostest method.

Finally, and perhaps most importantly, is the claim of significant bias (lower results) of the Diagnostest cholesterol method. In a recent publication by Wybenga et al. (12), they clearly show that results for the Diagnostest cholesterol method are in excellent agreement with those obtained by the generally accepted reference method of Abell et al. (13). The AutoAnalyzer cholesterol method is not an accepted reference method, and is known to give results that are biased high when compared with the Abell reference method (14). A modification of the automated method, whereby zeolite is added to the isopropanol extract to remove bilirubin, has been reported to improve the agreement between it and the Abell method (14). Since Logan and Sunderland did not use this modified automated method, we conclude that the Diagnostest cholesterol results are accurate and that the “reference” method values are biased high.

We would like to comment also on
some speculative statements made by Logan and Sunderland regarding the premature introduction of kits by manufacturers. We have participated in the development and evaluation of the Diagnostest Reagent Sets from their beginning and can state, without reservation, that none of these sets were introduced by Dow Diagnostics without their being first thoroughly evaluated, and then having a rigid quality control program designed for production. These included such items as determining the precision and accuracy of all pipetting or transfer devices supplied, comparisons of test methodology with true reference methods utilizing normal and abnormal patient sera, reinvestigation of normal ranges, and the performance of stability studies on all reagents at temperatures ranging from 8° to 55°C. Furthermore, a continuing research program in the areas of improving reagent stability, method specificity, and precision exists for all reagent sets being marketed.

These products are offered to serve a need in the field of clinical laboratory medicine for convenient, rapid, and reliable reagent sets. They are useful not only in the individual physician’s office but in the largest hospitals as emergency procedures and back-up methods for automated equipment. Dow Diagnostics and Bio-Science Laboratories recognize and accept the deep responsibility that the marketing of products used in human health care carries with it.

References

VINCENT J. PILEGGI  
JOHN DI GIORGIO  
RICHARD J. HENRY  

Bio-Science Laboratories  
7800 Tyrone Ave.  
Van Nuys, Calif. 91405

To the Editor:

We wish to respond to the letter of Drs. Vincent J. Pileggi, John DiGiorgio, and Richard J. Henry, in which they criticize some of our methods of assessment of the Diagnostest System and challenge some of the comments set forth in our paper [Clin. Chem. 16, 990 (1970)].

At the outset we wish to say that we are pleased to note the interest shown by the Dow Chemical Co. in our findings and their genuine concern to market high-quality reagent sets through research and development and a rigid quality-control program. We have appreciated the discussions we have had both with representatives of the Canadian firm and the parent company in Indianapolis throughout our study. They have demonstrated an eagerness to cooperate in our evaluation and to incorporate a number of the suggestions we presented. This approach to the marketing of a product is commendable and in our experience one not always practiced by manufacturers of reagent kits and sets.

The policy that we have pursued in our laboratory in the evaluation of reagent kits, sets, and instruments has been to present the data we have ob-