insulin species they use, further serious doubt on the validity of data reported on the alleged hypoglycemia unawareness associated with human insulin in his center is warranted.

INGRID MÜHLHAUSER, MD
MICHAEL BERGER, MD

FROM THE HEINRICH-HEINE-UNIVERSITY OF DÜSSELDORF, DÜSSELDORF GERMANY.

ADDRESS CORRESPONDENCE TO INGRID MÜHLHAUSER, MD, UNIVERSITÄT DÜSSELDORF, MEDIZINISCHE KLINIK, ABTEILUNG STOFFWECHSEL UND ERNÄHRUNG, MOORENSTRABE 5, D-4000 DÜSSELDORF, GERMANY.

TYPE I DIABETES, INSULIN-DEPENDENT DIABETES MELLITUS.

References

1. Smith GD, Egger M, Teuscher A: Confusion over human Insulin (Letter). *Lancet* 2:1214, 1991
2. Egger M, Smith GD, Diem P, Teuscher A: Human versus porcine insulin gone wrong (Letter). *Diabetes Care* 15:585–86, 1992
3. Mühlhauser I, Berger M: Endless debate about an unsubstantiated cause of severe hypoglycemia: reply to human versus porcine insulin gone wrong (Letter). *Diabetes Care* 15:1110, 1992
4. Mühlhauser I, Heinemann L, Fritsche E, Von Lennep K, Berger M: Hypoglycemic symptoms and frequency of severe hypoglycemia in patients treated with human and animal insulin preparations. *Diabetes Care* 14:745–49, 1991

Questionable Accuracy of a Filter Paper Method for Measuring GHb

Gay et al. (1) report that a fingerstick capillary blood GHb method was highly correlated ($r = 0.89$, $P = 0.0001$) with venous plasma HbA_1 by an ion-exchange chromatography method (HbA_1). They conclude that the GHb method is accurate and should be considered a credible alternative research and clinical tool to HbA_1. However, a high correlation coefficient, r, a measure of association between variables, does not necessarily imply good agreement in terms of their clinical significance (2). Similarly, as the methods both measured glycated hemoglobin, it would be most surprising if the P value did not indicate a highly significant relationship.

Linear regression is not the best approach for comparing two methods because it assumes that all the error is associated with the dependent (y) variable. In practice both methods have imprecision. The quoted coefficients of variation at equivalent concentrations, and therefore the variances, indicated slightly more variability in HbA_1 than in GHb. A more representative comparison would be achieved with a regression technique, which takes this into account using a ratio of the method variances (3). Alternatively, the difference between methods can be plotted against the mean for each pair of results (4). This difference plot is simple to perform and lends itself to straightforward clinical interpretation.

The classification as multiples of the method SD is a valid indicator of clinical significance, but the data presented showed that 24 of 58 samples (41%) were classified differently. In the range between the upper limit of the reference range and +10 SD, which should be sensitive to changes in glycemic control in average to well-controlled patients, the two methods would give the clinician different messages on 57% of occasions. Treatment decisions in patients whose samples were classified differently may be similar, as suggested, but the acceptability of such a high discordance rate is questionable. If glycated hemoglobin is to be employed as an accurate and objective measure of glycemia, then different methods for its measurement should agree. The authors suggest that much of the error may be attributable to the known problems of the ion exchange assay. However, they eluted their filter paper samples at 7 days after collection rather than 14 as originally reported (5). This could have contributed to the discrepancy, because at 7 days continuing and unpredictable in vitro glycosylation was demonstrated, which introduced variability and persisted until 14 days. It is not stated whether the change in the GHb protocol has been validated. Changes to the affinity chromatography method, if introduced in other laboratories, could lead to further variability (6).

Important differences in clinical classification have recently been reported between paired filter paper GHb by affinity chromatography and plasma HbA_1 by microcolumn ion-exchange chromatography (7). In view of the different analytical principles and the difficulties in calibrating glycated hemoglobin assays (8,9), method differences should be assessed carefully, with appropriate use of statistical tests. On the evidence of Gay et al. (1), the accuracy of the filter paper method remains open to question.

RICHARD P. TAYLOR, PHD, MRCPATH

FROM THE DEPT. OF CLINICAL BIOCHEMISTRY, JOHN RADCLIFFE HOSPITAL, HEADINGTON, OXFORD, UK.

SEND CORRESPONDENCE TO RICHARD P. TAYLOR, PHD, DEPT. OF CLINICAL BIOCHEMISTRY, JOHN RADCLIFFE HOSPITAL, HEADINGTON, OXFORD, OX3 9DU, U.K.

References

1. Gay EC, Cruickshanks KJ, Chase HP, Klingensmith G, Hamman RF: Accuracy of a filter paper method for measuring glycosylated hemoglobin. *Diabetes Care* 15:108–10, 1992
2. Dedrick RF, Davis WK: What do statistics really tell us about the quality of the data from self-monitoring of blood glucose? *Di-*

abetic Medicine 6:267–73, 1989
3. Strike PW: *Medical laboratory statistics*. Bristol, UK, J. Wright & Sons Ltd., 1981
4. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–10, 1986
5. Little RR, McKenzie EM, Wiedmeyer H-M, England JD, Goldstein DE: Collection of blood on filter paper for measurement of glycated hemoglobin by affinity chromatography. *Clin Chem* 32:869–71, 1986
6. John WG: Glycated haemoglobin analyses—assessment of within- and between-laboratory performance in a large UK region. *Ann Clin Biochem* 24:453–60, 1987
7. Slemenda CW, Marrero DG, Fineberg SE, Moore PS, Gibson R: Mail-in paper strip vs. microcolumn technique for measurement of glycosylated hemoglobin. *Diabetes Care* 13:886–88, 1990
8. Standing SJ, Taylor RP: Glycated haemoglobin: an assessment of high capacity liquid chromatographic and immunoassay methods. *Ann Clin Biochem* 29:494–505, 1992
9. John WG, Bullock DG, MacKenzie F: Methods for the analysis of glycated haemoglobins: what is being measured? *Diabetic Medicine* 9:15–19, 1992

Response From Authors

Dr. Taylor (referencing Dedrick and Davis [1]) states that the correlation coefficient r is a measure of association between variables but does not necessarily imply agreement. He suggests that our highly significant P value (which we obtained when we compared FP and VAC results for GHb [2]) does not necessarily indicate good clinical agreement. Furthermore, he states that linear regression is not the best approach for comparing two methods where both are imprecise, and suggests that a plot of the difference against the mean for each pair of results would allow a straightforward clinical interpretation. Such a plot is appropriate to examine agreement between the two measures but not the association that was our intent.

Our purpose was to examine the relative position of a subject on the two measures as quantified by the correlation coefficient. That is, we wished to investigate whether a subject with a high reading on FP would have a high reading on VAC, and whether one might expect intervention leading to a decrease in VAC would be reflected, similarly, in a decreased FP reading. Figure 1 in our paper presented the pictorial view of this association, which can be used to predict one value from the other via the regression equation (2). The laboratories' established means and ranges of normal controls are not the same for the two tests (FP: range 5.82–7.58%, mean 6.70%; VAC: range 6.3–8.2%, mean 7.25%). This tells us that values are not interchangeable and do not meet Taylor's definition of "agreement" a priori. Thus, the regression equation was included to allow interested readers to evaluate the correspondence themselves, and our suggestion that one measure should be used consistently was intended to highlight the lack of equivalence in results.

Our classification of scores as multiples of the method SD results in 59% comparably classified, with a kappa statistic of 0.44 (3) indicating fair to good classification agreement. Dr. Taylor suggests excluding the most extreme group of 16 subjects with results >10 SD above the mean. These 16 subjects are 27.6% or our sample. Given that the mean for FP results was 9.3 SD above the mean of controls and the average VAC reading was 8.1 SD above the mean of controls, eliminating these extremes would not appear to be justified in evaluating the test for use with diabetic subjects. Furthermore, in an earlier publication (4), we reported that among 588 diabetic subjects in the Colorado IDDM Registry average HbA_1 levels as measured by the FP method were 10 SD above the mean for normals.

We recognize that methods for comparing clinical measurement technique are the subject of considerable discussion. We appreciate Dr. Taylor's attention to this methodology and his recent work assessing GHb methods (5). We agree that differences between methods should be evaluated carefully. However, it was not our intent to say that results from the two methods are interchangeable. Rather we wished to demonstrate that the filter paper method is a useful tool for epidemiological research because it accurately classifies the level of glycemia. We apologize for confusion we may have perpetrated in this regard. The Spearman rank correlation is designed to evaluate relative rank in two samples, not agreement between the two measures; and the SD cross classification results add weight to our assertion that subjects will fall in similar relative positions on the two measures.

ELIZABETH C. GAY, MA
KAREN J. CRUICKSHANKS, PHD
H. PETER CHASE, MD
GEORGEANNA KLINGENSMITH, MD
RICHARD F. HAMMAN, MD, DRPH

FROM THE DEPARTMENTS OF PREVENTIVE MEDICINE AND BIOMETRICS AND PEDIATRICS, UNIVERSITY OF COLORADO SCHOOL OF MEDICINE; AND THE CHILDREN'S HOSPITAL, DENVER, COLORADO.

ADDRESS CORRESPONDENCE TO ELIZABETH C. GAY, MA, DEPARTMENT OF PREVENTIVE MEDICINE AND BIOMETRICS, UNIVERSITY OF COLORADO SCHOOL OF MEDICINE, 4200 EAST NINTH AVENUE, BOX C-245, DENVER, CO 80262.

FP, FILTER PAPER AFFINITY CHROMATOGRAPHY; VAC, VACUTAINER ION-EXCHANGE CHROMATOGRAPHY.

References

1. Dedrick RF, Davis WK: What do statistics really tell us about the quality of the data from self-monitoring of blood glucose? *Diabetic Medicine* 6:167–73, 1989
2. Gay EC, Cruickshanks KJ, Chase HP,