Multicomponent Analysis for Alkaline Phosphatase Isoenzyme Determination by Multiple Linear Regression

Colin R. Tillyer,1,4 Sylvia Rakhorst,2 and C. Michael Colley3

Alkaline phosphatase (EC 3.1.3.1) isoenzymes in serum may be determined by multicomponent analysis of the enzyme activities in the presence of multiple inhibitors. To determine inhibition coefficients of the isoenzymes, we used multiple linear regression analysis to compare alkaline phosphatase activities in the presence of known inhibitors with electrophoretically determined isoenzyme activities in plasma and serum samples. All possible combinations of exactly determined and overdetermined linear systems of inhibitors were ranked according to their prediction error to select an optimum set. The best multicomponent system for prediction included the use of levamisole, phenylalanine, and heat inhibition at 56°C and 65°C to determine bone, hepatic, intestinal, and placental isoenzymes. Consideration of the hepatic isoenzyme as liver and macromolecular fractions resulted in significantly worse predictions. Error analysis involving repeat determinations and a simplex optimization of the inhibition coefficients indicated that the inaccuracy of the comparison electrophoretic method may have been a major factor affecting poor isoenzyme prediction in some samples.

Indexing Terms: statistics/internmethod comparison/enzyme inhibition

The differentiation of alkaline phosphatase (EC 3.1.3.1) isoenzyme activities in serum to determine the tissue sources of increased activity is a common request in clinical laboratories. Selective techniques involving chemical or thermal inhibition have been used for this for some time. Statland et al. (1) were the first to use a multicomponent system of inhibitors for routine analysis of all the known serum isoenzymes, and this approach has since been used by many authors, each using a different combination of inhibitors and differing algorithms based on those inhibitors (2–9). Various degrees of imprecision and accuracy have been reported, and a simple analysis (10) has shown that consideration of the potential error propagation in each particular inhibitor system is an important part of their assessment. The principles of multicomponent analysis underlying all of these assays are well known, particularly in spectroscopy, where much of the theoretical basis for systems with linear responses has been developed (11). The approach used in most multicomponent assays of alkaline phosphatase is to isolate the isoenzymes, determine the inhibition coefficients for each, and then use those coefficients to determine the quantities of each isoenzyme in a sample mixture by solving an exactly determined system of simultaneous equations. In spectroscopy, this approach is known as the "classical least-squares" or "K-matrix" approach, where exactly determined or overdetermined systems may be used (see Appendix). With this approach, instead of isolating the isoenzymes to determine the inhibition coefficients, the coefficients are estimated by multiple linear regression analysis (MLR) of results for calibrators spanning the widest possible set of mixtures of the isoenzymes found in clinical practice.6 The matrix of inhibition coefficients is then used to predict the amount of each isoenzyme in an unknown mixture. Using inhibitors selected from the literature, we used this approach to select an optimum multicomponent system that gives the most accurate prediction of bone, liver, macromolecular, intestinal, and placental isoenzymes in a set of unknown samples, and compared this method with a well-accepted electrophoretic method.

Materials and Methods

Samples

We obtained 155 samples with total alkaline phosphatase values ranging from 57 to 3358 U/L (median 185 U/L); these samples were routine plasma specimens from cancer patients (125) and routine pregnancy serum (30). We used 115 samples as the calibration set to estimate the inhibition coefficients of each isoenzyme; the remaining 40 (30 from cancer patients and 10 from pregnant women) served as a prediction set to validate the models. Procedures followed were in accordance with the Helsinki Declaration of 1975, as revised in 1983.

Alkaline Phosphatase Isoenzyme Models

Two models were used: a five-isoenzyme model [bone, liver, macromolecular (biliary), intestinal, and placental] and a four-isoenzyme model (bone, hepatic, intestinal, and placental). The latter model assumes there is no difference in inhibitor response between the liver and macromolecular isoenzymes; thus, the hepatic isoenzyme consisted of the pooled activities of the liver and macromolecular fractions on electrophoresis.

6 Nonstandard abbreviations: MLR, multiple linear regression; GHCl, guanidine hydrochloride; AMP, 2-amino-2-methyl-1-propanol; H53, heat inactivation for 10 min at 53°C; H56, heat inactivation for 5 min at 56°C; H65, heat inactivation for 10 min at 65°C; and RMSEP, root mean squared error of prediction.
Determination of Inhibition Coefficients

The inhibition coefficients were defined as the fractional activities of the pure isoenzymes remaining after each inhibitor treatment; they were obtained from the slope of the multiple regression equation for each inhibited activity (dependent variable) on each isoenzyme activity determined as the fraction of the total activity on electrophoresis (independent variable). All estimations were performed in duplicate and the mean values were used in the calculations. Alkaline phosphatase was measured on the Roche Cobas® Bio (Roche Diagnostics, Welwyn Garden City, UK) centrifugal analyzer with an optimized 2-amino-2-methyl-1-propanol (AMP) buffer assay (Boehringer Mannheim, Lewes, UK; BCL cat. no. 396 460) by the method of Bowers and McComb (12). The inhibitors used were guanidine hydrochloride, 0.3 mol/L (GHCl) (ultrapure, anhydrous; ICN Biochemicals, High Wycombe, UK); levanisole, 10 mmol/L (anhydrous; Sigma Chemical Co., Poole, Dorset, UK); 1-phenylalanine, 10 mmol/L (anhydrous; Sigma) (7); urea 1.3 mol/L (Analar®, BDH, Poole, Dorset, UK) (6). Thermal inactivation—for 10 min at 53°C (H53) (8), 5 min at 56°C (H56) (13, 14), or 10 min at 65°C (H65) (7)—was performed on the EquiBio Thermostat® thermocycler (Camlab, Cambridge, UK). Instrument programs for each inhibitor and total alkaline phosphatase activity were those used by Peake et al. (7, 9). We also considered total activity as a term to include for the solution of the system of equations (referred to as an inhibitor in the list of inhibitors below), with inhibition coefficients of 1 for each isoenzyme. The comparison method for isoenzyme quantification was the Paragon® IsoPal Plus (Beckman Instruments, Fullerton, CA) electrophoresis system with and without neuraminidase treatment of the sample. Gels were scanned with the Beckman Appraise® densitometer system. Samples with activities >400 U/L were prediluted for optimal densitometric results. Placental activity was not clearly separated from the bone fraction in this system, so we defined it as the heat-stable activity present after 10 min at 65°C. The inhibition coefficients for this particular thermal inactivation were therefore 1 for the placental isoenzyme and 0 for all other isoenzymes. If placental isoenzyme was present, its value was subtracted from the densitometric value for the bone or liver fractions, which overlap the placental fraction in this system.6

Statistical Analysis and Calculations

Multiple regression analysis was performed with Systat® for Windows® version 5.2 (Systat UK; West Wickham, Kent, UK) and BMIDP PC-90 program 9R (BMIDP Statistical Software, Cork, Ireland). Deming's regression was used for comparisons of multicomponent and electrophoretic methods, assuming equal variances in both methods (15). Prediction for all possible combinations of four- to eight-inhibitor sets and simplex optimization of inhibitor coefficients, by the method of Nelder

---


and Mead as implemented in the Basic program of Sunday (16), was performed on an Acorn Archimedes 420/1 Risc computer (Cambridge, UK). The criterion for ranking and optimizing each multicomponent system was the root mean squared error of prediction (RMSEP), defined as

$$\left[ \frac{1}{m} \sum_{i=1}^{m} \frac{1}{n} \sum_{j=1}^{n} (X_{\text{obs}}(ij) - X_{\text{est}}(ij))^2/(m \times n) \right]^{1/2}$$

where m is the number of samples in the prediction set (40), n is the number of isoenzymes (4 or 5), $X_{\text{obs}}(ij)$ are the observed activities (calculated with the multicomponent system), and $X_{\text{est}}(ij)$ are the expected activities (determined by electrophoresis).

Results and Discussion

Regression Estimates of Inhibition Coefficient Matrix

Electrophoresis of the calibration samples showed that a high proportion (97%) contained the macromolecular (biliary) isoenzyme fraction. MLR estimates of each inhibition coefficient for each set of isoenzymes were determined from the calibration set (Table 1). An MLR model with zero intercept was used, on the assumption that a sample with no isoenzymes should have zero activity. The inhibition coefficients were mostly highly significantly different from 1 (i.e., they inhibited the isoenzyme), except for the intestinal coefficients for GHCl and urea and the placental coefficients for H53, H54, and H56. These coefficients also had greater imprecision, perhaps due to the smaller number of patients with concentrations >0 U/L and the particularly narrow range of values for the placental and intestinal isoenzyme values (Table 1). The values obtained with GHCl were very different from those of Shephard and Peake (6). We could not explain this; repetition with different batches and makes of GHCl did not change the values we obtained for the inhibition coefficients. The inhibition coefficients for the liver and macromolecular isoenzymes were significantly different only for GHCl (P = 0.012 and H53 (P = 0.004). The heat lability of the macromolecular isoenzyme was less than that of the liver isoenzyme by the two treatments here, in contrast to the results of Crofton and Smith (13). The best discriminator between bone and liver forms was H56. The macromolecular inhibition coefficients had imprecisions about twice those of the bone, liver, and hepatic isoenzyme coefficients.

Selection of the Best Subset of Inhibitors

Four-isoenzyme model. Using the 40-sample prediction set, we investigated an exactly determined (four-inhibitor) and two overdetermined (five- and six-inhibitor) systems. All possible combinations in each subset were tested (i.e., 70, 56, and 28 combinations for four, five, and six inhibitors, respectively), and ranked according to their overall RMSEP. The five highest-ranked exactly determined systems are shown in Table 2. The most accurate inhibitor set (lowest RMSEP) was levamisole, Phe, H56, and H65. The most accurate over-
Table 1. Mean (SEE) inhibition coefficients for alkaline phosphatase isoenzymes from MLR analysis of calibration set (n = 115).

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Bone</th>
<th>Liver</th>
<th>Macromolec.</th>
<th>Hepatic*</th>
<th>Intestinal*</th>
<th>Placental*</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCl</td>
<td>0.454&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.846&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.662&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.793&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.456</td>
<td>1.043</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.025)</td>
<td>(0.056)</td>
<td>(0.012)</td>
<td>(0.417)</td>
<td>(0.230)</td>
</tr>
<tr>
<td>Levamisole</td>
<td>0.012&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.013&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.002&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.690&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.679&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.354&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.007)</td>
<td>(0.001)</td>
<td>(0.052)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.875&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.896&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.885&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.786&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.506&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.376&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.024)</td>
<td>(0.005)</td>
<td>(0.173)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>Urea</td>
<td>0.558&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.768&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.840&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.788&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.631</td>
<td>0.674&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.012)</td>
<td>(0.026)</td>
<td>(0.006)</td>
<td>(0.205)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>H53</td>
<td>0.477&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.765&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.917&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.809&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.486&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.033)</td>
<td>(0.007)</td>
<td>(0.254)</td>
<td>(0.140)</td>
</tr>
<tr>
<td>H56</td>
<td>0.284&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.641&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.719&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.683&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.159&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.896</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.034)</td>
<td>(0.007)</td>
<td>(0.253)</td>
<td>(0.140)</td>
</tr>
</tbody>
</table>

* Liver + macromolecular.  * 47 nonzero values, range 0–33 U/L;  * 27 nonzero values, range 0–72 U/L.  *–* Significant at:  * P < 0.0005,  * P < 0.005,  * P < 0.05. Significance refers to the null hypothesis that β (the regression or inhibition coefficient) = 1 (no inhibition).

determined system included five inhibitors (total, GHCl, levamisole, H56, H65) and had a RMSEP of 36 U/L; however, this system presented no practical advantage over the best four-inhibutor system, because it required an extra determination. Progressively more overdetermined systems showed no improvement.

**Five-isoenzyme model.** The most accurate prediction set was a six-inhibitor system (total, GHCl, levamisole, urea, H56, and H65) (Table 3). The RMSEP of this system was 96 U/L, much worse than the best four-component system. The increase in RMSEP was mostly in the liver and macromolecular isoenzymes. Seven- and eight-inhibitor systems showed the same trends in the best and worse cases as the four-isoenzyme model.

The overall best-performing system was the four-inhibitor (lev, Phe, H56, H65) four-isoenzyme (bone, hepatic, intestinal, placental) system.

Method Comparison and Performance of the Best Systems

An initial analysis of the regression of observed (calculated from best set of predictors) on expected (electrophoretic) results indicated that one very high result (bone 3078 U/L, hepatic 436 U/L) in the prediction set was exerting an undue influence on the bone regression (Cook distance = 152 (15)); this was removed from the data set and Deming regressions were performed on the best system for the four-isoenzyme model (Fig. 1 A–C and Table 4). There was a significant proportional (slope) bias in the bone prediction regression and a significant but small constant (intercept) bias in the intestinal prediction. The random component of the scatter of the bone and hepatic results was high, with residual SDs of 26 and 36 U/L, respectively (for reference, a comparison of two virtually identical routine AMP buffer alkaline phosphatase methods over a similar range of values on the Beckman Synchron systems had a residual SD of 6 U/L). Only 2 bone and 2 hepatic results were negative, but 22 of the intestinal results were (many of the expected intestinal values were at or near zero). Regression of the total activity, calculated from the sum of the predicted bone, hepatic, intestinal, and placental activities (observed), on the total as measured by the AMP method (expected) showed very close agreement (Fig. 1D; Table 4).

Regression analysis of the best system for the five-isoenzyme model gave similar results for the bone and intestinal isoenzymes. However, the liver isoenzyme showed a markedly greater residual SD than did the hepatic isoenzyme in the four-isoenzyme system (Fig. 2A and Table 4). The macromolecular enzyme showed a marked proportional bias and high degree of residual scatter, with 12 negative results (Fig. 2B). Again, the
predicted total activity showed very good agreement with the total activity measured directly by the AMP method (Fig. 2C; Table 4).

A plot of the differences (observed minus expected) for the four-isoenzyme system indicated that errors in the bone isoenzyme estimate were being matched by errors of equal magnitude but opposite sign in the liver isoenzyme estimate (Fig. 3). In the five-isoenzyme system, liver errors were balanced by equal and opposite macromolecular errors, and bone errors tended to be balanced by the combination of liver and macromolecular errors. This accounted for the far more accurate estimates of total activity by each system in comparison with the individual isoenzyme estimates.

Sources of Error in the Multicomponent System

Error propagation in the systems of linear equations used in multicomponent analysis has been extensively analyzed in the mathematical and statistical literature. In our multicomponent systems we had to consider the effect of errors in the inhibited activities and in the inhibitor coefficients on the error in the calculated result, as well as the influence of the inhibition coefficient matrix on error amplification in the system (see Appendix). Error in the inhibited activities is propagated as a simple linear compound into the error in the calculated isoenzyme activities (Appendix Eq. 4), and the variance of the result is given by the multivariate variance-covariance matrix (Appendix Eq. 6). Random error, measured as duplicate sample imprecision, was generally low (Table 5), the worst being in the heat and levamisole treatments, which were three- to sixfold greater than the others, but levamisole and H56 were still selected as components of the best systems. The calculated variance of the isoenzymes in the prediction set for each sample, was often much lower than that predicted on the assumption of statistically independent inhibited activities (Appendix Eq. 5), because of within-sample covariance. To test how much the effect of random errors in the inhibited activities affected the actual prediction performance of the best set, we re assayed in duplicate the 10 samples with the highest absolute and fractional differences, taking into account losses in storage, and recalculated the overall RMSEP. If these sample predictions were outlying because of random error, repetition should have shown a significant improvement in the RMSEP. The changes were mostly small com-

Table 4. Linear regression comparison of multicomponent predictions of isoenzymes with electrophoretic estimates.

<table>
<thead>
<tr>
<th>Isoenzyme</th>
<th>Slope*</th>
<th>Intercept*</th>
<th>Residual SD, U/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best four-isoenzyme system (Lev, Phe, H56, H65)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone</td>
<td>1.26^a (1.19–1.33)</td>
<td>-2.9 (-13.0 to 7.20)</td>
<td>26</td>
</tr>
<tr>
<td>Hepatic</td>
<td>0.97 (0.89–1.03)</td>
<td>-11.2 (-27.5 to 5.10)</td>
<td>36</td>
</tr>
<tr>
<td>Intestinal</td>
<td>1.26 (1.00–1.50)</td>
<td>-2.50^b (-4.30 to -0.80)</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>1.01 (1.00–1.03)</td>
<td>-4.9^c (-1.74 to -8.08)</td>
<td>6</td>
</tr>
<tr>
<td>Best five-isoenzyme system (total, GHCl, Lev, urea, H56, H65)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone</td>
<td>1.27^a (1.18–1.34)</td>
<td>-2.9 (-14.4 to 8.90)</td>
<td>30</td>
</tr>
<tr>
<td>Liver</td>
<td>1.11 (0.93–1.30)</td>
<td>1.5 (-25.5 to 28.6)</td>
<td>54</td>
</tr>
<tr>
<td>MM</td>
<td>0.58^a (0.38–0.78)</td>
<td>-8.2 (-25.4 to 9.01)</td>
<td>34</td>
</tr>
<tr>
<td>Intestinal</td>
<td>1.50^a (1.18–1.82)</td>
<td>-4.9^c (-7.11 to -2.71)</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>0.99^a (0.98–0.99)</td>
<td>1.96^c (0.17 to 3.78)</td>
<td>3</td>
</tr>
</tbody>
</table>

MM, macromolecular; other abbreviations as in Table 2.

* 95% confidence interval is given in parentheses.
^ Confidence interval does not encompass 1.
^ Confidence interval does not encompass 0.

Fig. 1. Four-isoenzyme model: Comparison of isoenzyme activities predicted by the best system (observed) with those obtained by the electrophoretic method (expected) for (A) bone, (B) hepatic, and (C) intestinal isoenzymes; (D) total activity compares the sum of the predicted bone, hepatic, intestinal, and placental activity (observed) with the total activity as determined by the AMP assay (expected). (—) line of identity; (—) Deming regression.
pared with the overall prediction error (Fig. 4) and resulted in only a small decrease, from 36 to 32.5 U/L, in the RMSEP, so we concluded that random error had a small effect. Though minimized, random error propagation is a problem, particularly for the bone and hepatic isoenzymes, as shown in Table 6, but it was not the major reason for the inaccuracy of prediction here. The degree of error amplification by the inverse coefficient matrix is a fixed property of the selected matrix; it depends on the relative selectivity of each inhibitor or treatment for each isoenzyme and is often measured by using the condition number of the matrix, derived from a matrix norm (Appendix Eqs. 8–10). The condition number of the matrix with the lowest RMSEP for the four-isoenzyme model was also the one with the lowest condition number, 12; this suggests that this factor is much more important than inhibited activity imprecision in determining the overall system imprecision.

The inhibition coefficients are subject to random error in their determination but are constants in the equation used to calculate isoenzyme activities; therefore, any error in their determination will contribute a systematic bias in the results. A fixed error or bias in the coefficients propagates as a proportional bias (Appendix Eq. 7); this perhaps accounts for the proportional biases encountered in the results for the bone isoenzyme in the four-isoenzyme system and for the bone and macromolecular isoenzymes in the five-isoenzyme system. Further examination of Eq. 7 with the coefficient values of the best four-isoenzyme system also showed that a small error in a single inhibition coefficient for (e.g.) the bone isoenzyme could produce an error that was proportional to the true bone isoenzyme but also an error of similar magnitude and opposite sign in the hepatic isoenzyme that was independent of the true hepatic isoenzyme concentration. A simple single coefficient error could therefore produce the pattern of equal and opposite errors seen in Fig. 3. Errors in the inhibition coefficients could originate from bias and imprecision in their determination by MLR analysis of the calibration data set. The MLR model assumed that: (a) there was no intercept term, (b) the variance in the dependent variable

Table 5. Imprecision in inhibitor activities (input variables).

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Median*</th>
<th>CV, %</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0.49</td>
<td>0–3.0</td>
<td></td>
</tr>
<tr>
<td>GHCl</td>
<td>0.98</td>
<td>0–5.7</td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td>3.2</td>
<td>0–142.0</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>0.54</td>
<td>0–5.3</td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td>0.73</td>
<td>0–3.4</td>
<td></td>
</tr>
<tr>
<td>H53</td>
<td>3.12</td>
<td>0–151.0</td>
<td></td>
</tr>
<tr>
<td>H56</td>
<td>3.96</td>
<td>0–18.8</td>
<td></td>
</tr>
<tr>
<td>H65</td>
<td>2.29</td>
<td>0–142.8</td>
<td></td>
</tr>
</tbody>
</table>

n = 155 for all variables except H65 (n = 32). Abbreviations as in Table 2. *Because CV was estimated from duplicate estimations, a highly asymmetric distribution was obtained; hence, the median was used as the best overall estimate rather than the mean.

The inhibition coefficients are subject to random error in their determination but are constants in the equation used to calculate isoenzyme activities; therefore, any error in their determination will contribute a systematic bias in the results. A fixed error or bias in the coefficients propagates as a proportional bias (Appendix Eq. 7); this perhaps accounts for the proportional biases encountered in the results for the bone isoenzyme in the four-isoenzyme system and for the bone and macromolecular isoenzymes in the five-isoenzyme system. Further examination of Eq. 7 with the coefficient values of the best four-isoenzyme system also showed that a small error in a single inhibition coefficient for (e.g.) the bone isoenzyme could produce an error that was proportional to the true bone isoenzyme but also an error of similar magnitude and opposite sign in the hepatic isoenzyme that was independent of the true hepatic isoenzyme concentration. A simple single coefficient error could therefore produce the pattern of equal and opposite errors seen in Fig. 3. Errors in the inhibition coefficients could originate from bias and imprecision in their determination by MLR analysis of the calibration data set. The MLR model assumed that: (a) there was no intercept term, (b) the variance in the dependent variable

CLINICAL CHEMISTRY, Vol. 40, No. 5, 1994 807
Fig. 4. Effect of reassaying the 10 samples in the prediction set with the highest absolute and proportional errors and analyzing with the best four-isoenzyme model for (A) bone and (B) hepatic isoenzymes.

The sample with a bone activity of 3078 U/L, which was excluded from the regression analyses (Figs. 1 and 2), is included here. (c) original result, (e) repeat.

Table 6. Imprecision propagation in the prediction set for the best overall system (four-isoenzyme, four-inhibitor).

<table>
<thead>
<tr>
<th>Input variables</th>
<th>Median CV, %</th>
<th>SD, U/L</th>
<th>Output variables</th>
<th>Median CV, %</th>
<th>SD, U/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lev</td>
<td>1.5</td>
<td>0.16</td>
<td>Bone</td>
<td>10.8</td>
<td>7.3</td>
</tr>
<tr>
<td>Phe</td>
<td>0.5</td>
<td>0.86</td>
<td>Hepatic</td>
<td>4.7</td>
<td>8.4</td>
</tr>
<tr>
<td>H56</td>
<td>2.9</td>
<td>3.74</td>
<td>Intestinal</td>
<td>2.2</td>
<td>0.19</td>
</tr>
<tr>
<td>H65</td>
<td>0.9</td>
<td>0.83</td>
<td>Placental</td>
<td>0.9</td>
<td>0.83</td>
</tr>
</tbody>
</table>

(by using the total sum of squared errors over the whole prediction set as the function to be minimized and the inhibition coefficient values for each inhibitor (except total and H65) as the input variables. The MLR estimates of each set of coefficients for each inhibitor were the initial estimates in the simplex algorithm. Although this resulted in some changes in the coefficient values for the four-isoenzyme system, the overall sum of squares was only marginally reduced (by 9.4%), suggesting that the MLR estimates were already quite close to the optimum and that the calculation of the isoenzyme content of the prediction samples was robust to inaccuracies in the coefficients. In contrast, stepwise simplex optimization of the best five-isoenzyme system produced a 65% decrease in the RMSEP, from 96 to 33 U/L. However, regression of observed (y) on expected (x) values resulted in a significant constant negative bias in the bone results (intercept = -19 U/L) and a significant constant positive bias in the macromolecular results (intercept = 15 U/L), giving nine negative bone results but only one negative macromolecular result.

The instability of the result suggests that the coefficients in the five-isoenzyme system were not optimal; however, we were unable to obtain a solution that adequately satisfied requirements for minimum bias and minimum imprecision in the results.

Sources of Error in the Comparison Method

Random errors in the electrophoretic method were not directly tested here, but Van Hoof et al. (17) showed that the repeatability of the method is comparable with that of other analytical methods, yielding between-batch SDs of 1–2.1 U/L (corresponding to CVs of 3–12%, the latter for very low concentrations of intestinal enzyme). However, a reexamination of the electrophoretic results in the prediction set revealed many possible sources of bias or inaccuracy. One specimen in the prediction set had a large unidentified band of fast mobility that was not bone, liver, macromolecular, placental, or intestinal; this band gave an unexpectedly high bone result on multicomponent analysis. Two results showed an apparent failure of neuraminidase treatment and two had poor band quality. A major feature of the electrophoreses was a substantial proportion of samples (19 of 40) in which the bone or liver isoenzyme was a shoulder on a larger liver peak; this must have led to inaccurate quantification of the isoenzymes, particularly the smaller ones. Many of the expected bone isoenzyme results were zero, an extremely unlikely result given the known normal ranges for bone and lack of evidence that bone isoenzyme is suppressed in conditions characterized by high concentrations of the liver enzyme; rather, small bone peaks were probably not discernible under a large liver peak. The error pattern seen in Fig. 3 could easily have been produced by a misallocation of the total activity among the isoenzymes by densitometry; because the total area of the densitometric scan is constant, an overestimation of one isoenzyme will result in the underestimation of the others if their bands overlap.

Another potential source of bias is the different methods.

(15) the activity in the presence of each particular inhibitor) was constant, and (c) there was no error in the independent variables (the electrophoretic estimates of each isoenzyme activity). The latter assumption is untrue, and the regression slope is slightly negatively biased (15); hence, estimates for the inhibitor coefficients may be slightly lower than expected. Also, variance in the dependent variable did increase with increasing values of the independent variables (i.e., the pattern was heteroscedastic). This increases the imprecision of the slope estimate but does not produce a biased estimate (15). To test these possibilities and examine whether the determined inhibition coefficients were not suboptimal for prediction, we used a simplex optimization technique. The coefficients for the best four- and five-isoenzyme systems were optimized for all the isoenzymes stepwise...
used for measuring the serum activity and the electrophoretic activity. Different substrates and buffers could produce markedly differing relative activities for the isoenzymes (18). Moreover, no account was taken of the intestinal variant enzyme (17); if it reacted to inhibitors the same way as the intestinal isoenzyme, we could have detected a positive bias in our results, but in fact the reverse was true (Fig. 1C). The presence of the variant, however, would have contributed to the overall scatter of results and inaccuracy in the determination of inhibition coefficients.

We have shown that it is feasible to use multicomponent systems for the determination of alkaline phosphatase isoenzymes, and that our optimum four-isoenzyme, four-component system gives moderately good quantitative estimates of bone, hepatic, intestinal, and placental fractions. This method could be used for the routine differentiation of bone and hepatic isoenzymes in clinical practice. It may be possible to estimate the macromolecular component separately by using an overdetermined system, but only two of the inhibitors tested, GHCl and HS3, gave significantly different values for the liver and macromolecular isoenzymes. The use of MLR for the determination of inhibition coefficients should give more accurate estimates than methods requiring enzyme isolation because matrix effects in plasma are included and a large calibration set reduces bias attributable to the variability between individual samples. Many of the published coefficients have been obtained for isoenzymes isolated from tissue rather than serum and often from only one or two individuals. Inaccurate inhibition coefficient determinations can theoretically be responsible for imprecision and inaccuracy in multicomponent analysis, but the simplex optimization gave no evidence that this factor was substantial for the best overall system. Imprecision amplification due to the condition of the inhibition coefficient matrix is still a problem, although our optimization has considerably reduced it for the inhibitors we used. Inhibitors that show more selectivity for the bone and hepatic isoenzymes would probably produce a significant improvement in imprecision. However, error analysis of our results suggests that much of the inaccuracy of the predictions could be explained by a misallocation of the total activity by the comparison electrophoretic method. A significant advantage of the multicomponent approach is that any treatment, not just chemical inhibition or denaturation, can be incorporated. More specific immune or lectin precipitation methods could be used to improve the selectivity between the bone and liver isoenzymes; any cross-reactivity inherent in such methods is taken into account in a multicomponent system. Multicomponent analysis also allows the possibility of developing partially or even fully automated quantitative assays for the isoenzymes with existing instrumentation.

We are grateful to Gerald Levin for the loan of the scanning facilities at the Department of Clinical Biochemistry, St. George’s Hospital, London.

References

Appendix
Multicomponent Systems

The fundamental equation for an exactly determined multicomponent system is

\[ y = Ax \] (1)
where \( y \) is the vector of enzyme activities in the presence of \( n \) inhibitors or treatments, \( x \) the vector of activities of each of \( n \) isoenzymes, and \( A \) is the \( n \times n \) matrix of activities of each pure component in the presence of inhibitor or treatment (the inhibition coefficient matrix). The solution vector of isoenzyme activities, \( x \), is given by

\[
x = A^{-1}y
\]

(2)

where \( A^{-1} \) is the inverse of the inhibition coefficient matrix. For overdetermined systems, where the number of inhibitors or treatments is greater than the number of isoenzymes, this becomes

\[
x = (A^TA)^{-1}A^Ty
\]

(3)

where \( A^T \) is the transpose of \( A \) and \( (A^TA)^{-1}A^T \) is the generalized inverse of \( A \).

Error Propagation in a Linear System

A fixed or constant error in the inhibited activities (\( y \)) in the exactly determined linear system (Eq. 1) is propagated as a fixed or constant error in the result (Eq. 2), assuming no error in \( A \), as

\[
e = A^{-1}h
\]

(4)

where \( e \) is the vector of errors in \( x \) (the isoenzyme activities), \( h \) the vector of errors in \( y \) (the inhibited activities), and \( A^{-1} \) the inverse of the inhibition coefficient matrix (19). The constant terms in \( A^{-1} \) can amplify the errors in \( y \) considerably, a property that is known as the condition of \( A \) (see below). If, in a given sample, the errors in \( y \) are random and the isoenzyme activities are statistically independent (they have zero covariance), the variance in the \( i \)th isoenzyme (\( \text{Var}_{\text{isoenzyme}_i} \)) is simply

\[
\text{Var}_{\text{isoenzyme}_i} = \sum_{j=1}^{n} a_{ij}^2 \text{Var}_{\text{inhibitor}_j}
\]

(5)

where the \( a_{ij} \) are the squared elements of the inverse of the \( A \) matrix, and \( \text{Var}_{\text{inhibitor}_j} \) is the variance of the \( j \)th inhibited activity (10). If they are not statistically independent (i.e., if the errors are correlated and have nonzero covariance), the isoenzyme variances are given by the diagonal terms of the transformed multivariate variance–covariance matrix

\[
(A^{-1})^T \Sigma A^{-1}
\]

(6)

where \( \Sigma \) is the variance–covariance matrix of the repeated inhibited activity values and \( (A^{-1})^T \) is the transpose of the inverse of the coefficient matrix (20). The exact variance of the results can therefore be calculated if repeat determinations of the inhibitor activities are made.

For fixed or constant errors in the inhibition coefficients \( (A) \), assuming no error in \( y \), the error vector in the solution \( x \) is given by

\[
e = -(I + A^{-1}F)^{-1} A^{-1}Fx
\]

(7)

where \( I \) is the identity matrix and \( F \) is the matrix of errors in the coefficients (19). This error depends not only on the magnitude of \( F \) and \( A^{-1} \) but also on the magnitude of the \( x \) vector, the (true) result, itself; therefore, a constant bias due to inaccurate determination of the coefficients produces a proportional bias in the results. Similar formulas will apply in an overdetermined system, where \( A^{-1} \) is replaced by \( (A^TA)^{-1}A^T \), the generalized inverse.

The overall imprecision in a sample is therefore determined by the imprecision in the isoenzyme activities, their covariance, and the condition of the coefficient matrix. The condition of a matrix (its ability to amplify or propagate errors) is conventionally measured by its norm or condition number.

The Matrix Norm and Condition Number

The norm of a matrix is a single measure of the magnitude of a matrix. The Euclidean norm of the inhibition coefficient matrix \( \|A\|_2 \) is defined as (21):

\[
\|A\|_2 = \left( \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}
\]

(8)

where the \( a_{ij} \) are the inhibition coefficients. The condition number of \( A \) is

\[
\text{Cond}(A) = \|A\|_2 \|A^{-1}\|_2
\]

(9)

where the norm of \( A \) can be the Euclidean or any other matrix norm. An inhibition coefficient matrix with a condition number of 1 is perfectly selective and shows no error propagation. Higher values indicate lower selectivity and higher degrees of error propagation. For an overdetermined system, the condition number (22) is

\[
\text{Cond}(A) = [\text{Cond}(A^TA)]^{1/2}
\]

(10)

The matrix norms can be used to derive upper and lower bounds on the error in these systems (19, 21), but the bounds can considerably exceed those actually found and they give no information on the error in each separate component in the system, which can vary widely. They are useful in selecting which of a number of different systems is likely to produce the least error propagation in practice, but will not substitute for actual calculations of inaccuracy and imprecision from a set of data for a given system.