Validation of Accuracy-based Amino Acid Reference Materials in Dried-Blood Spots by Tandem Mass Spectrometry for Newborn Screening Assays

DONALD H. CHACE,1 BARBARA W. ADAM,2 S. JAY SMITH,2 J. RICHARD ALEXANDER,2 STEVEN L. HILLMAN,3 and W. HARRY HANNON2*

Background: Advances in technology and the earlier release of newborns from hospitals have pressed the demand for accurate calibration and improved interlaboratory performance for newborn screening tests. As a first step toward standardization of newborn screening aminoacidopathy tests, we have produced six-pool sets of multianalyte dried-blood-spot amino acid reference materials (AARMs) containing predetermined quantities of five amino acids. We describe here the production of the AARMs, validation of their amino acid contents, and characterization of their homogeneity and their stability in storage.

Methods: To each of six portions of a pool of washed erythrocytes suspended in serum we added Phe (0–200 mg/L), Leu (0–200 mg/L), Met (0–125 mg/L), Tyr (0–125 mg/L), and Val (0–125 mg/L). Six-pool sets (1300) were prepared, dried, and packaged. We used isotope-dilution mass spectrometry to estimate the endogenous amino acid concentrations of the AARMs and validate their final amino acid concentrations. We used additional tandem mass spectrometry analyses to examine the homogeneity of amino acid distribution in each AARM, and HPLC analyses to evaluate the stability of the amino acid contents of the AARMs.

Results: The absolute mean biases across the analytic range for five amino acids were 2.8–9.4%. One-way ANOVAs of the homogeneity results predicted no statistically significant differences in amino acid concentrations within the blood spots or within the pools (P >0.05). Regression slopes (0 ± 0.01) for amino acid concentrations vs storage times and their P values (>0.05) showed no evidence of amino acid degradation at ambient temperatures, 4 °C, or −20 °C during the intervals tested.

Conclusion: The validation, homogeneity, and stability of these blood spots support their use as a candidate national reference material for calibration of assays that measure amino acids in dried-blood spots.

© 1999 American Association for Clinical Chemistry

Isotope-dilution mass spectrometry (IDMS)4 has been used as a definitive or reference method for the accurate quantification of a variety of compounds, including specific proteins (1) and peptides (2), and amino acids in dried-blood-spot (DBS) samples collected for newborn screening tests (3–5). The unequal specificity provided by IDMS is an important factor in the production of reference materials and has contributed substantially to the development of reference materials for clinically important serum analytes (6–9). In this study, we used IDMS as a reference method to accurately measure the endogenous concentrations and to certify the enriched concentrations of five amino acids in the dried-blood matrix of candidate reference materials for the calibration of analytic methods used to measure amino acids.

Effective screening of newborns with the use of blood-spot samples collected from newborns soon after birth, combined with diagnostic studies and treatment, helps

1 Neo Gen Screening, 110 Roessler Rd., Pittsburgh, PA 15220.
2 Centers for Disease Control and Prevention (CDC), Newborn Screening Quality Assurance Program, Atlanta, GA 30341-3724.
3 Duke University Medical Center, Mass Spectrometry Facility, Research Triangle Park, NC 27709.
*Author for correspondence. Fax 770-488-4831; e-mail whh1@cdc.gov.
Received March 3, 1999; accepted May 4, 1999.

4 Nonstandard abbreviations: IDMS, isotope-dilution mass spectrometry; DBS, dried-blood spot; AARM, amino acid reference material; NMR, nuclear magnetic resonance; S&S, Schleicher & Schuell; MS, mass spectrometry; and MS/MS, tandem mass spectrometry.
essential for assay calibration of newborn screening methods because matrix characteristics are not adequately managed during the standardization process. Primary standards, purified amino acids, are not suitable in most newborn screening DBS methods. AARMs are attainable among laboratories through the standardization process. The accuracy of measured values is substantially reduced interlaboratory variability (10). These factors support the position that accuracy-based DBS calibrators for amino acid DBS tests are needed. To meet this identified need, the Newborn Screening Quality Assurance Program operated at the CDC in Atlanta, GA produced 1300 six-pool sets of multianalyte DBS candidate amino acid reference materials (AARMs) (11) by dividing a single batch of mixed-donor blood of measured hematocrit into six pools; enriching the pools with pure amino acids to cover the usual analytic ranges of Phe, Leu, Met, Tyr, and Val; and dispensing 100-mL spots of the blood pools onto Schleicher & Schuell (S&S) Grade 903 blood-collection paper.

The development of certified reference materials is the first element in any comprehensive effort to standardize an analytic assay system. The variety of assays for amino acids must be standardized to ensure the comparability of data and the accurate interpretation of presumptive clinical assessments from laboratory measurements, especially when small changes from physiologic cutoff points may be important predictors of disease risk. When the concentrations of the assay calibrators are certified by use of an IDMS method, the accuracy of measured values is attainable among laboratories through the standardization process. Primary standards, purified amino acids, are not suitable in most newborn screening DBS methods because matrix characteristics are not adequately managed (10); therefore, a certified blood-spot product is essential for assay calibration of newborn screening methods. Here we provide the data for such a candidate DBS reference material.

**Materials and Methods**

**Materials for the AARMs**

Packed red blood cells were obtained from the International Red Cross, frozen human serum was obtained from Worldwide Biologicals, and saline, 0.9% Injection USP, solution was obtained from Baxter Healthcare. Amino acid calibrator solutions were prepared with Phe, Leu, Tyr, and Met from Sigma Chemical and Val from Eastman Kodak. The purities of the Phe, Met, Tyr, and Val, determined by proton nuclear magnetic resonance (NMR) analyses, were >99%; the NMR spectrum for Leu contained 1–2% impurities. Concentrated hydrochloric acid was obtained from J.T. Baker. S&S Grade 903 (lot no. W941) paper, printed with dashed-line circles of 13-mm diameter, was used as the support matrix for the blood-spot materials. All DBS materials were stored in metalized zip-closure bags from LPS Industries that contained desiccant packets from Multiform Desiccant Products.

**Stock Calibrator Solutions for AARM Whole-Blood Pools**

The candidate reference materials were enriched with two stock calibrator solutions: solution A, containing Met, Tyr, and Val, each at a concentration of 5 g/L in dilute HCl; and solution B, containing Phe and Leu, each at 5 g/L in deionized water.

**Materials for the Tandem MS Analyses**

Stable isotopes of [2H5]Phe, [2H3]Leu, [2H3]Met, [2H4]Tyr, and [2H8]Val were purchased from Cambridge Isotopes as mass spectrometry (MS) internal standards. These internal standards were of both high chemical and isotopic purity, >99% as certified by the isotope manufacturers. Stock solutions of the internal standards were prepared in deionized water and HCl from J.T. Baker, and aliquots were added to whole-blood samples for IDMS analyses or to a methanol-deionized water (1:1, by volume) DBS sample-extraction solution for other tandem MS (MS/MS) analyses. High-purity methanol from Burdick and Jackson was used for sample extraction, butanolic-HCl (3 mol/L) from Regis was used for derivatizing the amino acids, and glycerol and sodium octyl sulfate from Sigma were used to reconstitute the samples for MS analyses according to methods published previously (3–5).

**Materials for HPLC Analyses**

3,5-Dibromo-L-tyrosine [(Br)2 Tyr], ornithine (both from Sigma), and 200 proof ethanol (Pharmco) were used to prepare the blood-spot extraction solution for HPLC analyses; Amino Acid Standard H (lot no. 891207087) from Pierce was used to calibrate the HPLC analyses. To perform the HPLC analyses, we used a Waters™ AccQFluor™ Reagent kit (cat. no. 52880) and Eluant Concentrate (cat no. 52890), reagent-grade water and acetonitrile (Sigma), and Nova Pak column guards and an AccQ. Tag™ analytical column (cat. no. 52885; Waters).

**Preparation of AARM Dried-Blood Spots**

Recently collected blood units of insufficient quantity for transfusion were washed with three portions of saline to remove anticoagulants and the buffy coat. After centrifugation and removal of the last saline wash, the combined red cells were reconstituted to a hematocrit of 57% ± 0.5% with pooled human serum that had been clarified by
serial filtration through filters with pore diameters of 1.2, 0.8, 0.65, 0.45, and 0.22 μm. This blood batch was divided into portions for enrichment with Phe and Leu at 0, 40, 80, 120, 160, and 200 mg/L blood and Met, Tyr, and Val at 0, 25, 50, 75, 100, and 125 mg/L blood. The liquid added during enrichment was sufficient to reduce the hematocrit to 53% ± 0.5%.

After enriching the whole-blood pools with the amino acids, adding sufficient blood to bring them to volume, and mixing them thoroughly, we refrigerated them overnight. To minimize the effects of fluctuations in room temperature and heat transfer from the magnetic stirrer on which the blood pools were placed, we brought the pools to the temperature of a running water bath and left them in the bath during the dispensing process. Micromedic (ICN Biomedicals) Model 25,000 piston-driven automatic pipettes were used to dispense the blood. These pipettes were gravimetrically calibrated to dispense the six pools of whole blood with intact red cells in 100-μL portions onto S&S Grade 903 (lot no. W941) filter paper placed horizontally on special racks. We dried all blood spots overnight under ambient conditions before transferring them to zip-close metalized plastic bags containing desiccant packets and storing them at −20 °C. Humidity within the containers was maintained at <30% during storage.

**PREPARATION OF ISOTOPE-ENRICHED AARMS FOR MEASURING ENDOGENOUS AMINO ACID CONCENTRATIONS**

A portion of each amino acid-enriched whole-blood pool used to prepare the AARM blood spots was combined with 40 μL of a solution containing deuterium-labeled amino acid calibrators to make a final volume of 5 mL per pool. These whole-blood materials for IDMS analysis were dispensed onto S&S Grade 903 (lot no. W941) paper, dried overnight under ambient conditions, and stored at −20 °C in zip-close metalized plastic bags containing desiccant packets.

**MS/MS ANALYSES**

We extracted amino acids from the DBS disks, using a published procedure (3) modified for use with one 3/16-inch (~4.8-mm) DBS disk per analysis.

Isotopically labeled internal standards for the IDMS samples were incorporated into the dried-blood samples during preparation; thus, no additional internal standard was added during IDMS sample preparation for MS/MS analysis. However, the AARM blood spots used to assess the homogeneity of amino acid distributions within the DBSs and within the DBS pools did not contain isotopically labeled amino acid internal standards. Internal standards for these MS/MS analyses were added to the methanol extraction solution as described previously (3–5) and thus did not go through the same extraction process as the amino acids contained in the blood spots.

A Quattro I triple-quadrupole tandem mass spectrometer from Micromass Instruments (formerly Fisons, VG) with a Lab-base data system (Micromass) was operated in the liquid secondary ionization mode [fast ion bombardment, as described previously (4, 5)]. We introduced samples using a static probe at a frequency of ~1 sample/3 min. Neutral loss of 102-Da scan functions were used with a mass range of 125–300 to accumulate data for 1 min as described previously (3–5). We postprocessed data using a program that assigned quantitative values for each mass of interest and then calculated ion abundance ratios of the masses for metabolites of interest. In this paper, we define ion abundance ratios as the ratio of the ion abundance for the native amino acid divided by the ion abundance for the deuterated amino acid. For example, the ratio of ion intensity of Phe (m/z 222) to the respective internal standard ion intensity of d5-Phe (m/z 227) was determined for each analysis. These ratios can be converted to concentrations by applying correction factors and constants derived from known calibrators and/or calibration curves.

**HPLC ANALYSES**

We used previously described sample extraction (12) and derivatization (13) procedures, with modified DBS disk size and sample elution time, to prepare samples for HPLC analysis. Elution of amino acids from 1/4-inch (6.4 mm) DBS samples into a 700 mL/L ethanol solution containing (Br)2Tyr internal standard was facilitated by a 10-min sonication. A 20-μL aliquot of each eluate was combined with 60 μL of AccQ.Flour borate buffer and 20 μL of AccQ.Flour derivatizing reagent, and the mixture was heated for 10 min at 55 °C.

We used protocols provided by the manufacturer to perform amino acid analyses with a Waters AccQ.Tag HPLC system. The HPLC system consisted of an autosampler and LC Module I pump, a control system with a column heater maintained at 37 °C, an inline degasser, and a 474 scanning fluorescence detector (Waters) set for an excitation wavelength of 250 nm and an emission wavelength of 395 nm. A reversed-phase AccQ.Tag 150 × 3.9 mm C18 column with a solvent system consisting of a three-eluent gradient (AccQ.Tag eluent, acetonitrile, and water) was used at a flow rate of 1.5 mL/min for a sample run time of 19 min. Separation and data reduction were controlled by Millinium® 2010 Chromatography Manager (Waters) software.

**DETERMINING THE ENDOGENOUS AMINO ACID CONCENTRATIONS OF THE BLOOD USED FOR THE AARMS**

From duplicate IDMS analyses in each of 16 analytical runs, we determined the mean ion abundance ratio of each of the five amino acids for every pool in the AARM series. The means of the ion abundance ratios for each amino acid were plotted as ordinate or dependent-variable y values against the amino acid enrichments of the pools (abscissa or independent-variable x values), and a
weighted linear regression analysis of the resulting plot was used to determine the $y$-intercept and slope of the regression line. We calculated the regression analysis weights as the inverse variances of the MS response measurements at each concentration. The $y$-intercepts determined from the Phe, Met, Tyr, and Val regression analyses were equivalent to the endogenous concentrations of the blood batch from which the DBS reference materials were made. The $y$-intercept determined from the Leu regression analysis represented the sum of (Leu + Ile) because, as described previously (5), Leu and Ile have identical masses and, therefore, are not separated by this MS/MS method. However, certain assumptions let us closely approximate the sum of the concentrations of Leu and Ile, and using HPLC we were able to quantify each of these amino acids (5). We used HPLC analyses to measure the Leu and Ile fractions in the nonenriched AARM pool; we then derived the endogenous Leu concentration by multiplying the $y$-intercept, determined from IDMS analyses, by the Leu fraction (0.66), determined from HPLC analyses. We recognize that Leu values are derived and that this derivation may not necessarily be applied to every newborn sample. It was important to include these data for the following reasons: (a) These calculations can be replicated by other users of MS/MS as described in the validation article on Leu, Ile, and Val by MS/MS (5), and (b) it is a reference point for other technologies that allows comparison between other methods that measure either Leu and Ile individually or as a total (Leu + Ile) in control or nondiseased blood spots. We determined the square of the coefficient of correlation ($r^2$) from each linear regression to indicate the strength of the among-pool linear relationships.

ASSIGNING THE TARGET CONCENTRATIONS OF THE AARM MATERIALS

We defined the target concentration of each amino acid in an AARM pool as its endogenous concentration plus its enrichment (its added concentration). We chose this method of setting the target concentrations on the assumption that amino acid calibrator solutions could be prepared and added to the blood pools more accurately than the amino acid content of the DBSs could be extracted and measured.

VALIDATING THE TARGET CONCENTRATIONS OF THE AARMS

The weighted linear regression plots used to establish endogenous concentrations of the amino acids in the AARMs were also used to derive their measured concentrations. To correct for the concentration bias that would be introduced by leaving the endogenous concentration ($y$-intercept or $b$ value) on the plot, we subtracted it from every point along the regression line so that the linear equation ($y = mx + b$) became $y = mx$. We then divided the measured ion abundance ratios ($y$-axis values) by the regression line slope ($m$) to derive the measured concentrations ($x$-axis values).

To derive the endogenous Ile concentration, we multiplied the IDMS-measured (Leu + Ile) concentration in the nonenriched AARM pool (pool A) by the Ile fraction (0.34), determined from HPLC analyses. We then subtracted the endogenous Ile from the (Leu + Ile) concentrations of all six AARM pools to yield their measured Leu concentrations. Finally, we validated the target concentrations by comparing the assigned target concentrations of the five amino acids with their measured concentrations.

EVALUATING THE HOMOGENEITY OF ANALYTE DISTRIBUTION WITHIN THE AARM BLOOD SPOTS

In a series of 16 MS/MS analytical runs, we measured the ion abundance ratios of Phe, Leu, Met, Tyr, and Val in north, east, south, and west disks punched from DBSs on the first and last cards of each AARM pool and used these ratios in a one-way ANOVA to determine the probability that disks taken from different locations within the DBSs contain the same mean amino acid concentrations.

EVALUATING THE HOMOGENEITY OF ANALYTE DISTRIBUTION WITHIN THE AARM PRODUCTION BATCH

From the 16-run series of MS/MS analyses, we used the amino acid ion abundance ratios of DBS disks taken from the north position of DBSs on the first, last, and four intermediate cards of pools C and F in a one-way ANOVA to determine the probability that disks punched from different cards within a pool contain the same amino acid concentrations.

ESTIMATING AMINO ACID STABILITIES IN AARMS STORED AT VARIOUS TEMPERATURES

Strips of blood spots from AARM pool D were stored in zip-closure plastic bags with desiccant packets and humidity indicator cards in the dark at 37 °C, ambient temperature, 4 °C, and −20 °C. At predetermined intervals, strips of DBSs were transferred from increased, ambient, and refrigerated temperatures to storage at −20 °C, and strips of DBSs initially stored at −20 °C were transferred to storage at −70 °C. The humidity of all samples was controlled to <30% throughout the storage studies. All samples in each complete time and temperature set were analyzed in duplicate in each of two HPLC analyses. We used simple linear regression analysis of measured amino acid concentrations vs storage time to assess the stability of the amino acids at each storage temperature, and we tested each regression slope with a statistical $t$-test to evaluate whether it was significantly different from zero.
Results

**Endogenous amino acid concentrations of the blood used to prepared AARMs**

The endogenous amino acid concentrations of the AARM pools, as determined from the y-intercepts of the regression analyses, and the square of the coefficient of correlation ($r^2$) from each linear regression are shown in Table 1. By-analyte linear regression analyses of ion abundance ratios determined by IDMS vs enriched amino acid concentrations yielded $r^2$ values >0.98 for all analytes except Tyr. Visual inspection of ion abundance ratios plotted vs enriched amino acid concentrations showed an ion abundance ratio for Tyr in pool E that was much lower than expected. For this reason, we deleted the analytic results from Tyr pool E and recalculated the Tyr linear regression to derive the Tyr endogenous concentration and $r^2$ value that are shown in Table 1.

**Validation of the target concentrations of the AARM materials**

We compared the target amino acid concentrations with concentrations measured by IDMS to assess the degree of concordance between the two sets of concentration values. The absolute mean biases across the analytic ranges of the amino acids were 2.8% for Phe, 6.4% for Leu, 4.8% for Met, 3.1% for Tyr, and 9.4% for Val. The Tyr ion abundance ratio in pool E was much lower than expected. The results of MS/MS analyses used to assess homogeneity of amino acid distribution within the DBSs and within the production lots and HPLC analyses (data not shown) indicated that the spurious Tyr result from IDMS was related to an error in internal standard addition. The results from the secondary MS/MS and HPLC analyses indicated a linear relationship among all Tyr pools and supported the validity of the Tyr target concentration of pool E.

**Within-spot homogeneity of analyte distribution in the AARM blood-spot pools**

The $P$ values determined from one-way ANOVAs express the probability that amino acid concentrations are the same throughout the blood spot. $P = 0.95$ indicates a 95% probability that amino acid concentrations are identical throughout the blood spot; $P = 0.05$ indicates a probability of only 5% that amino acid concentrations are identical throughout the blood spot. Table 2 shows the $P$ values of within-spot distributions for the five amino acids with which the AARM pools were enriched.

**Homogeneity of analyte distribution within the AARM pools**

The $P$ values from the one-way ANOVAs, shown in Table 3A, are derived from one measurement per card in each of 16 analytical runs. Sixteen measurements were available from each card included in these MS/MS analyses. The mean ion abundance ratios for Phe are typical of the card-to-card variability and are shown in Table 3B.

---

**Table 1. Endogenous amino acid concentrations of AARMs (y-intercepts of weighted regression lines).**

<table>
<thead>
<tr>
<th>Analyte</th>
<th>No. of observations$^a$</th>
<th>$y$-intercepts, mg/L blood</th>
<th>SD</th>
<th>Lower</th>
<th>Upper</th>
<th>Mean $r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>16</td>
<td>12</td>
<td>0.67</td>
<td>10.69</td>
<td>13.31</td>
<td>0.9964</td>
</tr>
<tr>
<td>Leu</td>
<td>16</td>
<td>11</td>
<td>2.52</td>
<td>12.06</td>
<td>21.94</td>
<td>0.9862</td>
</tr>
<tr>
<td>Met</td>
<td>16</td>
<td>1</td>
<td>0.86</td>
<td>-0.69(0)</td>
<td>02.69</td>
<td>0.9810</td>
</tr>
<tr>
<td>Tyr</td>
<td>16</td>
<td>9</td>
<td>0.43</td>
<td>8.16</td>
<td>9.84</td>
<td>0.9929</td>
</tr>
<tr>
<td>Val</td>
<td>16</td>
<td>28</td>
<td>3.05</td>
<td>22.02</td>
<td>33.98</td>
<td>0.9935</td>
</tr>
</tbody>
</table>

$^a$ The observations are the means of duplicate analyses in each of 16 analytical runs.

**Table 2. Test of homogeneity of analyte distribution within AARM blood spots.**

$P$ values$^a$ from one-way ANOVAs (north, east, south, and west punches)

<table>
<thead>
<tr>
<th>Analyte</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>Mean$^b$ $P$ values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>0.4366</td>
<td>0.7850</td>
<td>0.3328</td>
<td>0.6426</td>
<td>0.9268</td>
<td>0.1390</td>
<td>0.5438</td>
</tr>
<tr>
<td>Leu$^c$</td>
<td>0.9173</td>
<td>0.7591</td>
<td>0.0652</td>
<td>0.9429</td>
<td>0.2851</td>
<td>0.2880</td>
<td>0.5429</td>
</tr>
<tr>
<td>Met</td>
<td>0.7462</td>
<td>0.9460</td>
<td>0.6766</td>
<td>0.9412</td>
<td>0.9928</td>
<td>0.7135</td>
<td>0.8360</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.8169</td>
<td>0.7785</td>
<td>0.5326</td>
<td>0.3916</td>
<td>0.1809</td>
<td>0.2899</td>
<td>0.4984</td>
</tr>
<tr>
<td>Val</td>
<td>0.7501</td>
<td>0.3497</td>
<td>0.9879</td>
<td>0.9733</td>
<td>0.4168</td>
<td>0.3455</td>
<td>0.6372</td>
</tr>
</tbody>
</table>

$^a$ $P$ values for pools A, B, C, E, and F were derived from 128 observations; $P$ values for pool D were derived from 127 observations.

$^b$ No statistical interpretation of these values is implied; they are shown for comparison only.

$^c$ Leu $P$ values were computed from (Leu + Ile).
ESTIMATING THE AMINO ACID STABILITIES AT 37 °C, AMBIENT TEMPERATURE, 4 °C, AND −20 °C
The P values derived from the linear regression analysis of amino acid concentration vs storage time predict no statistically significant amino acid degradation during storage at −20 °C for 1 year and no statistically significant degradation during the intervals at which the samples were stored at ambient or refrigerated temperatures with humidity controlled at <30%. (If P values were >0.05, we concluded that the regression slopes were not significantly different from zero and that we therefore had no evidence of amino acid degradation.) All five amino acids showed significant degradation after storage for 28 days at 37 °C. The analytic results for Phe shown in Fig. 1 are representative of the thermal stability data observed for the other amino acids.

Discussion
Reference materials are major contributors to the harmonizing of results among different methods and laboratories. A reference material is a material whose property values are sufficiently well established to be used for

![Figure 1](image-url)

Fig. 1. Stability of Phe in AARM blood spots stored at 37 °C (A), ambient temperatures (B), 4 °C (C), and −20 °C (D). The solid lines indicate regression slopes.
calibrating an apparatus, assessing a measurement method, or assigning values to other materials (14). A certified reference material is a reference material whose property values are certified by a procedure that establishes its traceability to an accurate realization of the unit in which the property values are expressed and for which each certified value is accompanied by an uncertainty level or a level of confidence (14). Thus, the development of reference materials is a formidable task for which the ultimate certification depends on the final acceptance of the materials by the end-users. The development of dried-blood reference materials is an especially difficult task relative to the development of liquid materials because of the many additional variables that must be controlled and measured in the production of dried-blood materials.

Because all pools in the AARM series were made from a single batch of whole blood, all contained identical endogenous amino acid concentrations. At low concentrations, such as the endogenous amino acid concentrations in the AARMs, baseline noise can contribute substantially to assay imprecision. To derive more precise estimates of the endogenous concentrations, we used ion abundance ratios of the analytes, determined at multiple concentrations, in linear regression analyses.

Because the molecular ion and its fragment ions for Leu and Ile generated by MS/MS methods are identical in mass, we used HPLC analyses to determine the relative endogenous concentrations of these two isomers and multiplied the MS/MS-generated combined total of (Leu + Ile) by the Leu fraction to obtain the endogenous Leu concentration in the AARM pools. The derived concentrations of Leu, from the total of Leu and Ile concentrations, is essential for newborn screening purposes, particularly for MS/MS, which uses a derivation (5) to estimate the total concentration of Leu and Ile. These derived concentrations may be important for other methods as well; therefore, the derived values were computed for this analyte in the DBS reference materials to benefit and inform all users. The endogenous concentrations of the other four amino acids were directly equivalent to the y-intercepts determined from the regression analyses of mean ion abundance ratios vs enriched concentrations. In contrast, the endogenous Phe value for the base-blood pool for production of the European Phe DBS reference material was based on HPLC measurements obtained from multiple laboratories (15).

Because the ion abundance ratio for Tyr in pool E was much lower than expected, and subsequent analyses by other methods indicated an accurate Tyr enrichment of pool E, we omitted the pool E Tyr ion abundance ratio from the linear regression analysis used to determine the endogenous Tyr concentration of the AARM pools and the calculation of the $r^2$ value for that regression line. After the Tyr $r^2$ value was corrected, all $r^2$ values were $>0.98$, indicating strong among-pool linear relationships for all amino acids.

The endogenous concentrations of amino acids are reflected in the differences between the enriched and target concentrations of the AARM pools. Comparison of MS/MS-measured amino acid concentrations with target concentrations showed concordance between the two sets of values, thus validating the target concentrations. $P$ values $≤0.05$ would indicate a statistically significant probability, at the 95% confidence level, of nonhomogeneous analyte distribution within the blood spots. The ANOVAs of ratios of signals from amino acids in the DBSs to signals from the deuterated amino acid internal standards added to the extraction solution predicted no statistically significant differences in amino acid distributions within the blood spots. In examining the homogeneity of amino acid distributions throughout the production lots, we found that the $P$ values showed no statistically significant differences in amino acid concentrations among cards within a pool. Variability of amino acid recovery from the AARM pools was not related to card number.

In a recent review of DBS storage stability studies, Therrell et al. (16) reported that, for maximum stability of most analytes, DBSs should be stored at low temperatures and controlled low humidity. We conducted the thermal stability evaluations reported here to test the stability of the AARM materials that were maintained in a dark environment at $<30\%$ humidity while stored at temperatures that might be encountered in transit (increased and ambient temperatures), in routine storage for daily laboratory use (refrigerated at 4 °C), and in long-term frozen storage ($−20\,\circ \text{C}$). Only after prolonged exposure to increased temperature ($37\,\circ \text{C}$) did the AARM materials show significant loss of amino acid concentrations. Using regression slopes and y-intercepts from analyses of the $37\,\circ \text{C}$ storage data, we estimated that Phe, Leu, Tyr, and Val would retain 83–85% of their initial concentrations after storage for 30 days at $37\,\circ \text{C}$, whereas Met would retain $−76\%$ of its initial concentration after 30 days at $37\,\circ \text{C}$. Our observation that Met is less stable in storage than Phe, Leu, and Tyr is consistent with previously reported studies on the stabilities of newborn screening blood-spot samples stored at room temperature for periods $≥2$ years (17). Because the $P$ values derived from analyses of samples stored at ambient room temperatures and under refrigeration were $>0.05$, we concluded that no statistically significant loss of amino acid concentration occurred during the time intervals tested. We detected no loss of analyte concentration from the materials stored for 1 year in the dark at $−20\,\circ \text{C}$ and $<30\%$ humidity; longer-term evaluation of the stability of the AARMs at $−20\,\circ \text{C}$ is in progress.

Only two sources of filter paper have been approved by the Food and Drug Administration for blood collection in the United States. These papers, Grade 903 (Schleicher & Schuell) and BFC 180 (Whatman), must meet the national criteria defined for acceptable performance (18). Because performance can differ among filter papers (19, 20), DBS reference materials for use in the US must be
prepared on an approved paper source. Hematocrits and
the size of the blood spot are also important performance
variables, and they should simulate the routine specimen
being collected from the newborns. Each of these vari-
ables affects the retained serum volume in an aliquoted
punch from the blood spot. The Phe DBS reference
material produced in Europe (15) was not prepared on a
US-approved paper source, and the blood-spot volumes
are much smaller (50% or more) than those routinely used
in the US for newborn screening. The amount of serum in
a given paper-punch-sized blood aliquot increases with
the volume of blood applied to the paper, but this increase
may not be directly proportional to the volume applied.
The exact nature of this relationship is difficult to predict
and may be influenced by several variables, such as red
blood cell lysis, room humidity, force of application, point
or style of application (chromatographic effects), and type
of filter paper. Our multi-amino acid DBS materials were
prepared under monitored environmental conditions and
on US-approved paper with precise blood volumes sim-
ulating the collected volume for screening laboratories in
the US and several other national screening programs.

Most screening programs analyze for several amino
acids in addition to Phe. Furthermore, the multi-amino
acid calibrator allows laboratories to use amino acid ratio
comparisons (21) for improved quality-assurance efforts.
Because mothers and babies are being discharged from
birthing centers earlier, screening programs need more
sensitive and specific methods of testing to reduce testing
redundancy and improve follow-up operations for babies
with metabolic disorders. These changes in newborn
screening have driven the improvement and shift in
testing technologies for laboratories. The new technolo-
gies need high accuracy-based calibration materials for
multiple amino acids in a single calibration material to
complement the testing technologies (e.g., HPLC and
MS/MS).

In an effort to demonstrate and quantify the impact
that standardized calibration of DBS methods has on the
measurement of amino acids, we will distribute sets of
AARMs, along with a panel of test samples, to a group of
manufacturers and screening laboratories that use differ-
cent analytical methods. This standardization process ini-
tiated by the availability of common reference materials
will focus efforts that can lead to enhanced overall inter-
laboratory comparability. In Europe, the recently released
Phe reference material has reportedly demonstrated im-
proved interlaboratory comparability (15), and similar
improvements should occur for the five amino acids with
the use of the AARMs in US newborn screening labora-
tories and other laboratories using the same type of
blood-collection paper and larger blood-spot collections.

These candidate DBS reference materials for method
 calibration will be primarily directed toward producers of
screening-test products; however, requests from screen-
ing laboratories will be considered on the basis of their
justified need. Initially, the distribution will not be very
restrictive because of the large quantity of materials
available (1300 sets) and because several screening labo-
atories use noncommercial methods. Notification of the
availability of materials will be made by an announce-
ment in appropriate newsletters. The AARMs will be
stored at CD and distributed from CDC. On the basis of
the production quantity of these materials and the docu-
mented stability of the amino acids, we anticipate the
AARMs to be available for at least 5 years, especially
because they are kept in continuous storage under opti-
mal conditions. We will monitor the stability of the
materials during the storage intervals and take appropri-
ate actions to produce new materials when a significant
change in analyte concentration is observed.

References

1. Barr JR, Maggio VL, Patterson DG, Cooper GR, Henderson LO,
   Turner WE, et al. Isotope dilution-mass spectrometric quantifica-
   tion of specific proteins: model application with apolipoprotein
   Characterization of β-endorphin in human pituitary by fast atom
   bombardment mass spectrometry of trypsin-generated fragments.
3. Chace DH, Hillman SL, Millington DS, Kahler SG, Adam BW, Levy
   HL. Rapid diagnosis of homocystinuria and other methioninemias
   from newborns’ blood spots by tandem mass spectrometry. Clin
4. Chace DH, Millington DS, Terada N, Kahler SG, Roe CR, Hofman
   LF. Rapid diagnosis of phenylketonuria by quantitative analysis for
   phenylalanine and tyrosine in neonatal blood spots by tandem
5. Chace DH, Hillman SL, Millington DS, Kahler CR, Naylor EW. Rapid
diagnosis of maple syrup urine disease in blood spots from
newborns by tandem mass spectrometry. Clin Chem 1995;41:
62–8.
et al. Total serum cholesterol by isotope dilution/mass spectrom-
7. Bernert JT, Bell CJ, McGuffey JE, Waymack PP. Determination of
   free glycerol in human serum reference materials by isotopoe-
dilation gas chromatography-mass spectrometry. J Chromatogr
8. Patterson DG, Patterson MB, Culbreth PH, Fast DM, Holler JS,


