Preliminary Proficiency Testing Results for Succinylacetone in Dried Blood Spots for Newborn Screening for Tyrosinemia Type I

Barbara W. Adam,1* Timothy H. Lim,1 Elizabeth M. Hall,2 and W. Harry Hannon1

BACKGROUND: Succinylacetone (SUAC) is the primary metabolite accumulated in tyrosinemia type I—an inborn error of metabolism that, if untreated, can cause death from liver failure during the first months of life. Newborn screening laboratories measure SUAC in dried blood spot (DBS) samples to detect asymptomatic tyrosinemia type I. We used panels of SUAC-enriched DBSs to compare and evaluate the performance of these screening tests.

METHODS: We prepared sets of DBS materials enriched with predetermined SUAC concentrations and distributed samples of these materials, along with a screening practices questionnaire, to laboratories that perform SUAC tests. We compared their reported SUAC concentrations and questionnaire responses to identify screening practices that affect SUAC test outcomes.

RESULTS: Data from 2 pilot surveys showed large differences among laboratories in SUAC recoveries, reproducible within-laboratory recoveries, and stable performance of the DBS materials. Results from 257 proficiency test analyses contained a total of 6 false-negative misclassifications. Reported recoveries of added SUAC ranged from 0 to $\geq 100\%$. Low-biased SUAC recoveries were associated with 1 method used by 5 laboratories. All laboratories that reported SUAC recoveries $\geq 100\%$ used DBS matrix calibrators.

CONCLUSIONS: The wide ranges of SUAC concentrations reported for pilot and proficiency testing specimens demonstrate a need to harmonize quantitative results among laboratories. Although DBS matrix calibrators are important for optimizing SUAC recoveries, the preparation of these calibrators is not standardized among laboratories. Certified DBS-based SUAC calibrators are needed for accuracy and harmonization.

© 2009 American Association for Clinical Chemistry

Tyrosinemia is an inborn error of metabolism that, if untreated, can cause death in the early years of life. Tyrosinemia type I (hepatorenal tyrosinemia) is a good candidate for newborn screening (NBS) because its untreated acute form results in death from liver failure during the first year of life (1). Early treatment with 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione is effective in preventing the adverse effects of this disorder (2–4) in $>90\%$ of patients. Patients who are not responsive to the pharmaceutical intervention can be rescued by liver transplantation (1).

Both tyrosine and succinylacetone (SUAC) accumulate in the blood of patients with tyrosinemia type I; however, measuring tyrosine alone is neither diagnostically specific nor sensitive enough to serve as an NBS test for tyrosinemia type I (5). SUAC is the primary biochemical marker accumulated in the blood of newborns with tyrosinemia type I and is specific to that disorder (6).

Recent publications of tandem mass spectrometry (MS/MS) methods for measuring SUAC concentrations in dried blood spot (DBS) specimens (7–11) have presented practical protocols for NBS for tyrosinemia type I, and NBS laboratories worldwide are implementing these and other protocols (12, 13).

For $>30$ years, the Newborn Screening Quality Assurance Program (NSQAP) has provided quality assurance (QA) materials and services to support NBS laboratories worldwide. In 2008, NSQAP began development of a SUAC QA program to support laboratories that perform NBS tests for tyrosinemia type I.
In this document, we report results from 2 distributions of the SUAC pilot blood spot materials and the first 2 SUAC proficiency testing (PT) challenge panels. In addition, we summarize participants’ clinical classifications of the PT specimens and their responses to a screening practices questionnaire that was distributed with the PT surveys.

Materials and Methods

All DBS materials for this study were made from whole blood units collected in citrate phosphate dextrose adenine (CPDA)-1 and purchased from a regional blood bank. Before use, each blood unit was adjusted to 55% hematocrit by plasma removal.

The SUAC (4,6-dioxoheptanoic acid, 99.5% purity) used to prepare stock solutions for enriching whole blood was purchased from Sigma-Aldrich. We used deionized water to prepare a SUAC stock solution (10,000 μmol/L) and a working solution (1000 μmol/L).

We used a Rainin EDP 2 multidispensing pipettor (Rainin Instruments) to dispense the blood in 75-μL aliquots onto Whatman 903 filter paper printed with dashed-line circles 13 mm in diameter. During dispensing, the filter paper cards were suspended horizontally on racks that prevented them from touching any surface. After overnight drying under controlled ambient conditions, we separated the stacked DBS cards with sheets of weighing paper (Fisher Scientific) and packaged them for storage at −20 °C in zip-closure Bitran Series S liquid-tight specimen bags (Com-Pac International). The bags contained desiccant packets (Poly Lam Products) and humidity indicator cards (Desiccare, Inc.) to ensure that humidity remained <30%. The stored DBS materials were removed from −20 °C storage, allowed to acclimate to room temperature, and transferred to zip-closure MylarFoil bags (International Molded Packaging Corp.) before addition of fresh desiccant packets and distribution to participating laboratories.

We prepared a set of pilot SUAC DBS calibrators by dividing 1 unit of hematocrit-adjusted blood into 7 portions for SUAC enrichment at 0, 1.5, 3.0, 5.0, 10.0, 20.0, and 50.0 μmol/L and validated the SUAC concentrations by performing duplicate MS/MS analyses (11) of every calibrator in each of 5 runs (Table 1). Representative spots from all 7 pools in the set were distributed by courier to 8 laboratories for measurement of the SUAC concentrations. After approximately 9 months, these pilot DBS calibrators were distributed again to monitor the reproducibility of results. We used information gathered from the pilot studies to guide the development of NSQAP’s first SUAC PT challenges.

<table>
<thead>
<tr>
<th>Table 1. Certification of SUAC DBS calibrators.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibrator A B C D E F G</td>
<td></td>
</tr>
<tr>
<td>Enrichment, μmol/L</td>
<td>0 1.5 3.0 5.0 10.0 20.0 50.0</td>
</tr>
<tr>
<td>Observations, n</td>
<td>10 10 10 10 10 10 10</td>
</tr>
<tr>
<td>Mean, μmol/L</td>
<td>0.07 1.75 3.27 4.98 9.87 19.5 48.88</td>
</tr>
<tr>
<td>SD, μmol/L</td>
<td>0.09 0.20 0.17 0.43 1.11 1.66 3.99</td>
</tr>
<tr>
<td>CV, %</td>
<td>132.09 11.33 5.13 8.61 11.25 8.50 8.17</td>
</tr>
</tbody>
</table>

*a Results from MS/MS analyses performed at CDC.

The challenge panel for the first SUAC PT event contained 4 members of the pilot set of SUAC calibrators (specimens 3831–4) and 1 specimen from a DBS pool made of nonenriched hematocrit-adjusted blood from a single donor (specimen 3835). The 5-specimen panel for the second PT event contained 3 members from the set of SUAC DBS calibrators (specimens 4832–4) and 1 each of DBS specimens from nonenriched and SUAC-enriched portions of a single unit of hematocrit-adjusted blood (Specimens 4831 and 4835, respectively).

The 5-specimen DBS PT panels were packaged for distribution in the same manner as the pilot panels and sent at ambient temperature by courier to every active PT program participant, all of which received identical specimen panels. An instruction page, a data report form, and a questionnaire pertaining to SUAC screening practices were enclosed with every PT specimen panel.

Participants were asked to report each PT specimen’s SUAC concentration and its presumptive clinical classification (within or outside normal limits) and to respond to the screening practices questionnaire. We sorted participants’ reported quantitative results according to their reported screening practice variables to evaluate the effects of these variables on SUAC measurements.

Results

In general, each pilot-study laboratory recovered a consistent fraction of added SUAC across the concentration range tested (0–50 μmol/L), but among the laboratories, recoveries of added SUAC ranged from approximately 25% to 200%. The overall means of reported SUAC concentrations were comparable for participants that submitted reports for both surveys; results from the second survey were consistent with those from the first.

We used paired data sets from the 6 laboratories that participated in both pilot surveys to compare, for
each data set, the reported SUAC concentrations with the enriched SUAC concentrations of the pilot DBS materials. We then subjected the resulting sets of \( xy \)-coordinates to linear regression analysis (Fig. 1). The regression line slope of 0.8 generated from the first survey data was essentially the same as that from the second survey data; the \( y \)-intercept was 0.5 for the first survey and 0.6 for the second.

Reported quantitative and qualitative results from the 2 2008 quarterly PT events (14, 15) are summarized in Table 2. Quantitative data values outside the 99% CI around each specimen’s mean of all reported values were not included in the statistical analyses.

For the first PT event, 5 presumptive clinical assessments differed from our expected assessment because of participants’ specific clinical assessment practices. In those cases, we applied the participants’ reported cutoff values in a previously described grading algorithm (16) used to evaluate test performance. The 5 reported clinical assessments, which were otherwise incorrect, were judged correct by this procedure; thus, results from the first PT event contained no false-negative reports. Reported results from the second PT event contained 6 false-negative reports, 4 of which were associated with low quantitative values.

Reported SUAC cutoff values ranged from 0.5 to 10.0 \( \mu \text{mol/L} \). Participants’ reported quantitative results for PT specimen 4834 (enriched with 50 \( \mu \text{mol/L} \) blood) were used to evaluate the relationship between measured SUAC concentrations and cutoff values (Table 3). We used reported cutoff values for the first event for the 2 laboratories that did not report their cutoff values for the second PT event.

Thirty-one participants reported quantitative SUAC analytical results (\( \mu \text{mol/L} \)) for the second PT event and responded to the screening practices questionnaire. For these laboratories, the reported quantitative data for specimen 4834 were grouped by specific screening practice variables shown in Fig. 2.

Fig. 2A shows that the SUAC results from 23 non-kit MS/MS methods ranged from 3.3 to 124.9 \( \mu \text{mol/L} \); results from the 2 non-MS/MS methods were 0 and 30.6 \( \mu \text{mol/L} \), and results from a kit MS/MS method ranged from 1.3 to 15.3 \( \mu \text{mol/L} \).

Thirty laboratories that reported quantitative results for specimen 4834 also reported the strategies they used to extract SUAC from the disks punched from the DBS samples. The medians of reported concentrations from 12 laboratories that extracted SUAC from the residual blood disks (“ghosts”) left after amino acid and acylcarnitine extraction and 12 laboratories that used freshly punched dried blood disks to perform stand-alone SUAC tests were 36.2 \( \mu \text{mol/L} \) [interquartile range (IQR) 32.5 \( \mu \text{mol/L} \) and 34.4 \( \mu \text{mol/L} \) (IQR 34.3 \( \mu \text{mol/L} \)], respectively. These median values were higher than the median from the 6 laboratories that reported performing SUAC tests on the extracts prepared for analyzing other amino acids and acylcarnitines [8.9 \( \mu \text{mol/L} \) (IQR 5.9 \( \mu \text{mol/L} \))] (Fig. 2B).

Twenty-eight participants that reported using MS/MS tests also reported the internal standards they used. For specimen 4834, the medians of reported SUAC concentrations from 14 laboratories using 5,7-dioxooctanoic acid [41.2 \( \mu \text{mol/L} \) (IQR 41.6 \( \mu \text{mol/L} \)] and 9 laboratories using \( ^{13} \text{C}_5 \)-SUAC [28.2 \( \mu \text{mol/L} \) (IQR 30.4 \( \mu \text{mol/L} \)] internal standards were higher than the median of reported concentrations from the 5 laboratories that used the internal standard supplied with the MS/MS kit [9.1 \( \mu \text{mol/L} \) (IQR 4.9 \( \mu \text{mol/L} \)] (Fig. 2C).

Thirty participants that reported quantitative results for specimen 4834 also reported their calibration schemes (Table 4). The median of SUAC concentrations reported by 17 participants using DBSs for calibration was 48.0 \( \mu \text{mol/L} \) (IQR 30.2 \( \mu \text{mol/L} \); the median
of concentrations reported by the 8 participants that did not use DBSs for calibrating nonkit SUAC tests was 17.7 μmol/L (IQR 12.1 μmol/L); and the median of concentrations reported by the 5 participants that used only the internal standards supplied with the MS/MS kit was 9.1 μmol/L (IQR 4.9 μmol/L) (Fig. 2D). All laboratories that reported SUAC recoveries for specimen 4834 equal to or higher than the SUAC enrichment used DBS calibrators, and 11 of the 17 laboratories that used DBS calibrators reported quantitative SUAC values for specimen 4834 above the median of all values reported by the participants included in the data sorting [24.5 μmol/L (IQR 37.7 μmol/L)]. Only 2 of the laboratories that did not use DBS calibrators reported values that were above the median.

For PT specimen 4834, results reported by the 5 laboratories that used response factors (numerical factors used to correct for extraction losses or other known sources of analytic biases) ranged from 4.8 to 55.7 μmol/L, whereas results from the 23 laboratories that did not use response factors ranged from 3.3 to 124.9 μmol/L.

**Discussion**

In each of the 2 2008 pilot surveys, the between-laboratory differences in measured SUAC concentrations ranged from approximately 25% to approximately 200% of the SUAC enrichments. However, the within-laboratory and overall reproducibility of results between surveys indicated sustained long-term performance of the laboratories’ analytic methods and the SUAC pilot DBS materials—some of which were subsequently included in the first PT specimen panel.

Most laboratories correctly assigned the presumptive clinical classifications in the PT events, although their reported quantitative results ranged from 0% to 200% of the specimens’ SUAC enrichment values. This outcome confirms that their cutoff decision points are based on the recovery performance of their SUAC assay. Quantitative PT results are not used directly to judge the quality of screening performance; however, these results are reported to the participants for self-assessment and comparative purposes. Laboratories are evaluated on how successfully they use their

### Table 2. Reported proficiency testing results: SUAC concentrations.

| Specimen | Survey 1, quarter 3, 2008 | | Survey 2, quarter 4, 2008 |
|-----------|--------------------------|--------------------------|
|           | 3831 | 3832 | 3833 | 3834 | 3835 | 4831 | 4832 | 4833 | 4834 | 4835 |
| Enrichment, μmol/L | 0 | 1.5 | 10 | 20 | 0 | 0 | 3.0 | 5.0 | 50.0 | 40.0 |
| Observations, n | 17 | 18 | 18 | 18 | 17 | 27 | 29 | 30 | 30 | 30 |
| Outliers, n | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 |
| Mean, μmol/L | 0.5 | 1.8 | 8.4 | 17.0 | 0.6 | 0.4 | 2.1 | 4.0 | 28.8 | 21.3 |
| 95% upper limit* | 1.5 | 4.2 | 20.6 | 42.5 | 1.5 | 1.0 | 4.9 | 9.2 | 70.4 | 52.6 |
| 95% lower limit* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Clinical assessments, n | | | | | | | | | | |
| Within limits | 18 | 15 | 2 | 0 | 18 | 32 | 24 | 12 | 24 | 12 |
| Above limits | 0 | 3 | 16 | 18 | 0 | 1 | 9 | 21 | 2 | 0 |

a Values reported as inequalities were not included in the data set.

b (16).

c (15).

d Not included in number of observations.

* Confidence limit around the mean.

f Not evaluated: SUAC enrichment value was too close to cutoff values to expect classification consensus among laboratories.

### Table 3. Reported SUAC cutoff values compared with reported SUAC concentrations for proficiency testing specimen 4834.

<table>
<thead>
<tr>
<th>Range of reported cutoff values, μmol/L</th>
<th>n</th>
<th>SUAC concentration, μmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Median</td>
</tr>
<tr>
<td>0.05–1.50</td>
<td>11</td>
<td>8.89</td>
</tr>
<tr>
<td>1.51–4.00</td>
<td>11</td>
<td>28.24</td>
</tr>
<tr>
<td>4.01–5.55</td>
<td>9</td>
<td>44.20</td>
</tr>
</tbody>
</table>

* Specimen 4834 was chosen for examination because it was a SUAC DBS calibrator that contained an elevated succinylacetone concentration.
quantitative results to classify PT specimens as “within limits” and “outside limits.” Three of the 33 participating laboratories submitted a total of 4 false-negative misclassifications for tyrosinemia type I that were associated with reported low quantitative SUAC values. An examination of the screening practices of these 3 laboratories showed no commonalities that could be associated with the reported low values.

All laboratories that participated in the pilot analyses of SUAC-enriched DBS materials and all but 3 of the laboratories that participated in the PT survey challenges used MS/MS methods. With the exception of laboratories that used an MS/MS kit method, the MS/MS test protocols differed in their test strategies. These differences included the size of the DBS punched-disk used for analysis, the extraction of SUAC from DBS punched disks, the composition of the SUAC extraction cocktail, the internal standard used, the calibration scheme (internal standard alone or internal standard plus DBS calibrators), the use of response factors to correct for analytic biases, and other variables such as reagent composition and instrument settings. We grouped the reported quantitative data for 1 PT specimen (4834) according to some of the reported screening practice variables—test method, SUAC extraction strategy, internal standard, or calibration strategy—to examine the ways in which these variables affected SUAC recovery.

**Table 4.** Reported results for proficiency testing specimen 4834* sorted by participants’ calibration schemes.

<table>
<thead>
<tr>
<th>Calibration scheme</th>
<th>n</th>
<th>Median, (\mu \text{mol/L})</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS/MS nonkit method with internal standard and blood spot calibrators</td>
<td>16</td>
<td>49.2</td>
</tr>
<tr>
<td>Non-MS/MS method with blood spot calibrators</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>MS/MS nonkit method with internal standards only</td>
<td>7</td>
<td>16.2</td>
</tr>
<tr>
<td>Non-MS/MS method with liquid calibrators only</td>
<td>1</td>
<td>30.6</td>
</tr>
<tr>
<td>MS/MS kit method with internal standards only</td>
<td>5</td>
<td>9.1</td>
</tr>
</tbody>
</table>

* Specimen 4834 was chosen for examination because it was a SUAC DBS calibrator that contained an elevated succinylacetone concentration.
Laboratories that do not use DBS calibrators in their data reduction schemes should include the estimated blood volume used per dried blood disk for patient samples. Most laboratories reported using 3.2-mm (1/8-inch) disks; a few used 3/16-inch or 1/4-inch disks. Because several laboratories did not report the blood volume they equate with their punched disks, we were unable to group the reported results by blood volume per disk.

SUAC recoveries from 12 laboratories that extracted SUAC from ghosts were comparable to those from 12 laboratories that extracted it from freshly punched disks. This finding indicates that the extraction cocktail used for routine extraction of other amino acids does not extract SUAC equivalently. The SUAC recoveries reported by laboratories that coextracted SUAC with other amino acids and acylcarnitines reflect the lower SUAC extraction efficiency from that extraction cocktail. Other low-biased clusters of results were associated with using a commercial MS/MS kit method, coextracting SUAC and other amino acids and acylcarnitines from a single DBS disk, using an internal standard other than $^{13}$C$_5$-succinylacetone or 5,7-dioxooctanoic acid, or using the kit’s internal standard alone for calibration. In all cases, the cluster of low-biased SUAC-recoveries contained all the results reported by MS/MS kit users.

Although the kit method results were low relative to both SUAC enrichment and the overall mean of all reported results, this data grouping demonstrated the important harmonizing contribution of using a common test method calibrated with a single-source internal standard prepared according to a common protocol. In addition, this grouping showed that protocol harmonization can produce sets of results that are consistent without achieving test accuracy. This bias problem could be challenged if an external certified SUAC calibrator in the appropriate matrix was available. For each screening practice variable charted, the broad distributions of values from the nonkit methods reflected the between-laboratory diversity of test protocols and the use of calibrators from different sources prepared according to various protocols. Grouping data according to calibration strategies illustrated that more than half of the data reports came from laboratories that used DBS matrix calibrators for their SUAC tests, which establishes the need for DBS calibrators to enhance the performance of these SUAC screening tests.

The 5 laboratories that used response factors to correct for analytic bias did not report their uncorrected results, so we were unable to evaluate the direct effect of these response factors. For 1 PT specimen (4834), the distribution of quantitative results from the laboratories that used response factors fell within the range of those reported by the laboratories that did not use such factors. Response factors can be useful if methodologies are well established and well controlled, but they may be misleading if applied while the method is still in development flux or before an adequate database has been accumulated to focus method enhancements.

Almost every laboratory that used a nonkit method reported a unique set of screening practice variables. Grouping their results by screening practice variable pairs (such as analytic method and calibration strategy) produced no conclusive results. Some variables, such as composition of SUAC-extraction cocktails and instrument settings, were too complex to address adequately in the screening practices questionnaire. However, these variables probably contributed to the differences and biases observed in reported quantitative results. Additionally, it should be noted that some of the data reported for the 2 PT events came from laboratories using SUAC methods that were in development or pilot study status. The variability in quantitative results may, in part, reflect the relative inexperience of these laboratories. This question will be resolved by future PT challenges.

The reported presumptive clinical classifications of the PT specimens were in agreement more than an examination of the large variance in the quantitative data would suggest. However, the basis for each participant’s clinical classification decision is reflected in its assigned cutoff value, which is derived from using the test assay to analyze a sizeable unaffected group of patient specimens. Screening data reflect analytic bias; therefore, laboratories that reported low quantitative results also used and reported lower cutoff values.

The widespread use of DBS calibrators prepared in-house from SUAC standards of undeclared origin, coupled with the wide range of SUAC concentrations reported for PT specimens, suggests a need for certified, high-accuracy SUAC DBS calibrators to improve global performance of these assays. Such calibrators would help to harmonize among laboratories the accuracy of measured SUAC concentrations. To meet this need, NSQAP will make available sets of certified DBS SUAC calibrators that laboratories can use for establishing and evaluating their SUAC methods.

Currently, 49 laboratories have enrolled in the program for the next PT challenge. Through QA services and technical consultation, methods for detecting asymptomatic newborns with tyrosinemia type I should continue to improve. Improved methods will help reduce follow-up workload and parental stress by minimizing the number of false-positive reports (improving the specificity and positive predictive value) and by preventing delayed diagnosis of babies that is caused by undetected diagnosis in NBS.
Succinylacetone Measurements in Newborn Screening Tests

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures of Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: W.H. Hannon, Vice-President, ISNS.

Consultant or Advisory Role: W.H. Hannon, consultant, CDC management and laboratory.

Stock Ownership: None declared.

Honoraria: None declared.

Research Funding: None declared.

Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: We thank Thomas M. Hickey, PhD, DABCC, for an early (2007) alert to the emerging need for an external SUAC QA program.

References