

Introduction.....	1
Needs assessment.....	1
Methodology.....	2
1. Development and implementation in screening practice of clinically validated cut-off values and post-analytical tools	2
2. Training course at Mayo in small groups	6
3. Development of customized software to manage data collection, analysis, and reporting of NBS data.....	7
4. Collection, compilation and monitoring of performance metrics, with definition of targets of acceptable performance	8
5. Monthly conference calls and bi-annual face-to-face meetings	9
6. Continuing clinical validation of 2 nd tier tests	9
7. Round robin sample exchange	11
Resolution of challenges.....	12
References.....	13

Introduction

The goal of this renewal application is to continue and expand the laboratory quality improvement initiative funded previously as a regional collaborative project. In 2004, the newborn screening programs of all seven states included in Region 4 (IL, IN, KY, MI, MN, OH, and WI) agreed to participate in a project aimed to improve the analytical quality of newborn screening by tandem mass spectrometry (MS/MS). To date, active participation has expanded substantially beyond the boundaries of the region and even of the US, to include a total of 31 US states from all seven regions (Figure 1) and 29 international participants from 16 countries.

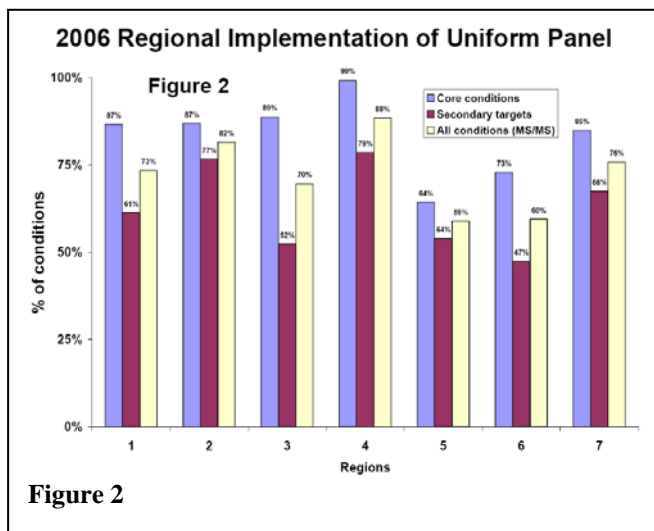


This level of participation has allowed a remarkable acceleration of the data collection phase (for example, we have accrued analytical data of more than 3,150 true positive cases), and expedited the creation, and frequent updating, of a series of tools that have contributed to the full achievement of the proposed goals of this project in the 2004-07 cycle. Namely, these goals were expansion and greater uniformity of testing, application of clinically defined cutoff ranges, and ultimately greater consideration and improvement of objective performance metrics. The successful launch of MS/MS testing in Kentucky [1] and the expansion of the Michigan program (testing all of the primary conditions and most of the secondary targets included in the ACMG/HRSA recommended panel) are two fitting examples of this project's positive impact.

This level of participation has allowed a remarkable acceleration of the data collection phase (for example, we have accrued analytical data of more than 3,150 true positive cases), and expedited the creation, and frequent updating, of a series of tools that have contributed to the full achievement of the proposed goals of this project in the 2004-07 cycle. Namely, these goals were expansion and greater uniformity of testing, application of clinically defined cutoff ranges, and ultimately greater consideration and improvement of objective performance metrics. The successful launch of MS/MS testing in Kentucky [1] and the expansion of the Michigan program (testing all of the primary conditions and most of the secondary targets included in the ACMG/HRSA recommended panel) are two fitting examples of this project's positive impact.

Needs assessment

As December 2006, the implementation in Region 4 of testing for the ACMG/HRSA recommended panel of conditions detectable by MS/MS [2-3] is the highest in the nation (99% and 79% for uniform panel conditions and secondary targets, respectively; see Figure 2).

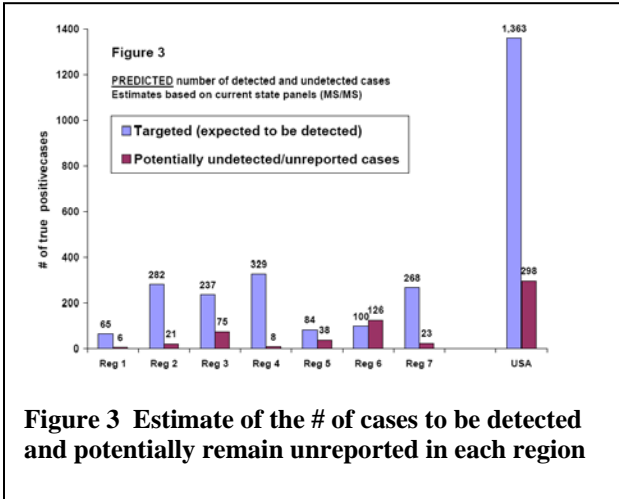


The only exception to a full regional implementation of the uniform panel [4] is likely to be resolved soon as the last state currently not screening for Tyrosinemia type I has agreed to a pilot implementation of the 2nd tier test for succinylacetone. The implementation of 2nd tier tests was also one of the stated objectives of the regional collaborative, and involves 12 US states (see table on p. 11).

While a full regional implementation of the uniform panel is encouraging, we recognize a continued need to underscore the clinical

importance of the differential diagnosis between primary and secondary conditions presenting with abnormal concentrations of the same marker(s), possibly leading to a greater consideration and eventual inclusion in state panels of additional secondary targets.

Based on the status of Region 4, we recognize a greater need to involve programs in other Regions. Limited implementation of the recommended panel results in significant variability in the rate of detection of affected cases. To place this issue in a practical perspective, an estimate of US regional detection rates based on 1) current state panels [4]; 2) incidence data from multiple sources [3, 5-9]; and 3) correction factors for selected ethnic variability (derived from



US 2002 census data) translates in a range of predicted detection rates between 1:2,301 (Region 4) and 1:5,982 (Region 6), with an overall average of 1:2,973 births (data not shown).

More relevant to the issue of maldistribution and existing gaps, the same analysis also predicts that approximately 300 cases per year are not identified at birth, 80% of them occurring in Regions 3, 5, and 6 (Figure 3). Our goal is to work more closely with already active participants in these Regions (Region 3: AL, FL, GA, NC, SC, and TN; Region 5: IA, MO, ND, and OK;

Region 6: CO, and UT) to mitigate the current inequity of services and to reduce the morbidity and mortality related to incomplete screening panels.

Methodology

1. Development and implementation in screening practice of clinically validated cut-off values and post-analytical tools

Over the last three years, this project has evolved from a rather improvised approach to an organized, web-based data collection system. National and international participants are provided with a user ID and password to gain access to a secure section of the Region 4 website (www.region4genetics.org). Once logged in, authorized participants have access to a folder unique to their state and to common folders defined for specific purposes (sample exchange, storage of templates, and posting of updated reports). The PI of the project and senior MPH personnel have administrative privileges and can access all folders. Using standardized templates (downloaded from the website), participants contribute on a monthly basis updates of the following data:

1. Complete set of data (amino acids, acylcarnitines) of confirmed true positive cases
2. Updated lists of selected percentiles (%iles) of normal population
3. Cutoff values
4. Performance metrics (period, volume, detection rate, false positive rate, positive predictive value)

These data are uploaded by the participants in dedicated subfolders of each state folder (true positives, score cards, cutoff tool) with a distinction between “active” and “archive” subfolders

for each category. Once a file has been posted, the PI is notified via e-mail that a new file has been added. In addition, a message is displayed after logging in that new files have been posted, with a direct link to the updated folders. Submitted data are added, or updated, in databases saved on a secure server with complete backup performed daily. These databases include:

1. True positive data (3,152 cases as December 31, 2006; same date applies to totals listed below)
2. Cutoff data (1,663 data points; submitted by all participants, and updated as necessary)
3. %iles of normal population (4,817 data points)

Each database is set to automatically perform basic calculation of descriptive statistics, particularly the calculation of pre-defined percentiles. These values (and the number of data points used to calculate them) are

automatically linked to a tool called the "Score card" (see Figure 4). This file offers a comparison between the distribution of values for each analyte and ratio in the normal population, the spread of cutoff values, and the %iles of the disease range (collection of data points of a given analyte or ratio in confirmed cases).

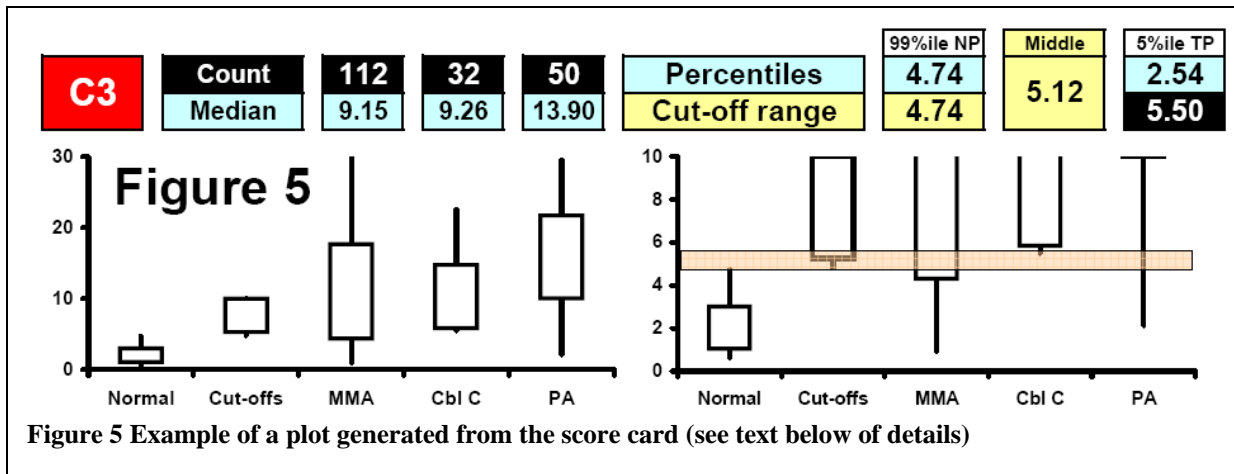
Region 4 - Laboratory Quality Improvement Collaborative Project (NBS by MS/MS)																										
Enter Name and 2-4 letter code		AMINO ACIDS										Figure 4														
State/Lab	Minnesota	MN																								
Updated 2/2006	NORMAL POPULATION					MN Cutoff	CUTOFF (CO)					CO RANGE		TRUE POSITIVES (DISEASE RANGE, DR)						Tot # cases 2,832						
	Cumulative Percentiles (N)						Cumulative Percentiles (N)					99%	5%	Cumulative Percentiles (N)												
ANALYTE		1%	5%	10%	50%	90%	95%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%
GLY	7	279	207	482	962	919	700	60	700	795	806	1045	1241	500 - 700	22	270	423	533	569	939	1,197	1,282	1,416	MSUD	Condition (WHO)	
VAL	10	58	77	102	141	186	200	22	169	277	298	316	420	186 - 189	60	111	189	216	283	377	467	643	600	MSUD	Condition (WHO)	
ILE/LEU	10	62	79	110	154	235	300	31	208	300	300	325	487	235 - 300	83	250	299	328	442	749	1,254	1,583	3,329	MSUD	Condition (WHO)	
MET	10	13	17	24	33	51	60	31	55	61	64	90	104	51 - 73	13	108	109	113	142	313	608	894	894	HCY	Condition (WHO)	
CIT	10	8	9	15	22	30	55	30	24	55	75	100	100	30 - 40	58	44	82	101	147	296	601	834	1,005	CIT-I	Condition (WHO)	
															28	29	33	45	58	138	247	382	887	CIT-II	Condition (WHO)	
PHE	10	39	44	58	77	103	130	37	110	130	139	163	242	103 - 131	459	148	242	271	328	433	574	801	1,070	PKU	Condition (WHO)	
															288	111	137	100	171	220	270	329	402	H-PHE	Condition (WHO)	
TYR	10	35	46	74	120	202	150	30	150	300	400	488	942	120 - 150	46	65	103	121	149	209	382	388	644	TYR-I	Condition (WHO)	
															8	228	240	278	324	430	500	749	1,117	TYR-II	Condition (WHO)	
ARG	9	3	4	12	21	35	50	25	43	60	80	140	220	35 - 110	4	96	110	128	180	268	330	338	338	ARG	Condition (WHO)	
															26	4	7	8	16	32	50	90	153	CIT-II	Condition (WHO)	
ASA	4	0.08	0.03	0.10	0.31	0.47	0.40	0.40	0.08	0.10	0.40	2.00	2.00	0.47 - 0.66	9	0.04	0.66	1.14	1.39	3.27	4.06	5.52	7.01	ASA	Condition (WHO)	

Figure 4 Score Card

determine the low end of the cutoff range. Without any data manipulation, the range of coefficient of variation (CV) is 10-56%, with a median value of 34%. However, the removal of the two most obvious outliers for each analyte improves the situation significantly, to a median value of 21%. This approach offers a practical mean to provide confidential feedback to the programs who filed the outlier values, so they can first exclude the possibility of a reporting error, and later assess the potential causes of the observed discrepancy in comparison to other participants. Furthermore, this kind of comparison validates the basic tenant of the collaborative project that in most cases data from different programs can indeed be merged. Assuming that a CV <20% is indicative of sufficient comparability, to date we have achieved this goal for at least one primary marker of 10 of the 20 conditions included in the uniform panel and 9 of the 22 secondary targets (six amino acids and six acylcarnitines).

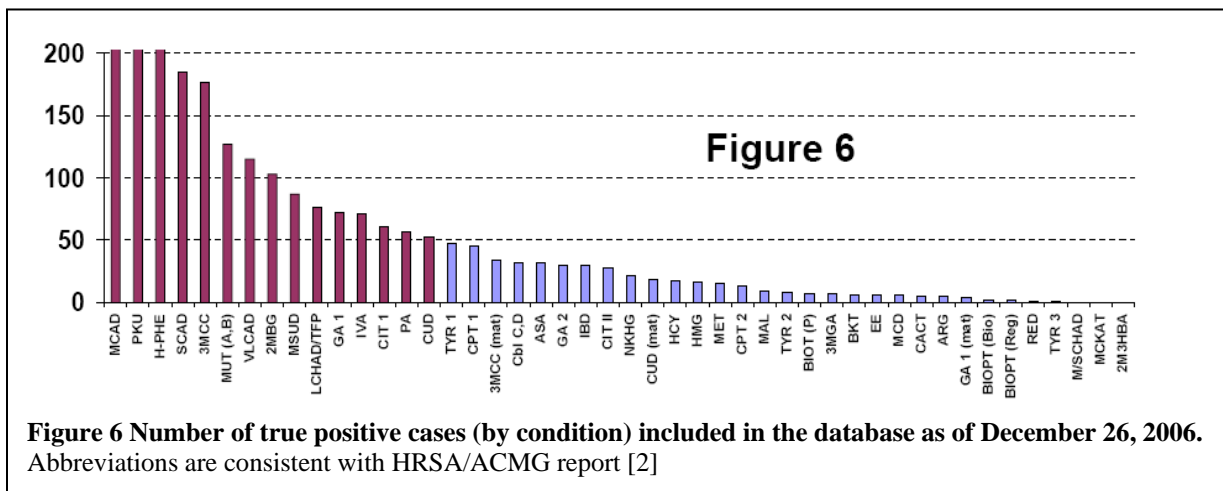
All the types of collected information drive the process of calculating a CUTOFF RANGE, which is defined as the interval between the 99%ile of the normal population (cumulative, calculated as the median of all submitted values) and the 5%ile of the disease range (also cumulative). Each set of data include an N value. When the degree of overlap between normal population and disease range makes inapplicable to use the rule stated above, one or both limits are modified to give priority to the disease range. In such instances, the value is shown with a black background. The table also includes a column where participants enter the actual cutoff

values used by their own program (see Figure 4), with embedded conditional formatting that highlight values that are within (green) or outside (orange) the cutoff range. To provide a more visual comparison between the three groups, the same data are linked to a set of plots (Figure 5).



Each group (normal population, cut-offs, and individual conditions, as applicable) are shown as a box encompassing the range between the 10%ile and the 90%ile of values included in the database. The vertical lines extend to the 1%ile (below) and 99%ile (above), respectively. The count of data points and median values are shown as numbers above the graph. For each analyte, the left plot shows a comparison over the whole range of values, the right plot highlights the clinically significant region where the cutoff range is located (shown as an orange shaded area). In the example shown in Figure 5 (propionylcarnitine, C3), the plot reveals a significant overlap between cutoff values (above the shaded area) and disease ranges of the three conditions linked to this analyte: Methylmalonic acidemia (MMA, includes three complementation groups: mut, CblA, and Cbl B), Methylmalonic acidemia (CblC) and Propionic acidemia (PA). Based on this evidence, one program has lowered its C3 cutoff by 35% and experienced a marked increase in their detection rate.

During the first cycle of the grant, the most pressing priority was to engage as many as possible participants. The rationale to expand participation beyond the boundaries of Region 4 relates to the reality of dealing with rare conditions, and our primary goal to collect data of at least 50 cases affected with each one the conditions included in the HRSA/ACMG recommendations. To date, 14 of the 20 primary targets listed in the uniform panel have reached, and in some cases



greatly exceeded (MCAD: N=586), the initial goal of 50 cases, 3 of the 22 secondary targets have reached the target as well (Figure 6). This has been possible only because of the expansion of the project beyond the regional boundaries: of the 10 US states which have contributed more than 100 cases, four are located outside Region 4. While this is evidence of steady progress, it also highlights the necessity to continue this effort and aggressively expand the number of national and international participants, so to achieve the goal of at least 50 cases for all conditions within the next five years. The benefits of a broad participation are not limited to data collection, but also include fostering a culture of data sharing and open communication, constructive inter-laboratory comparison, and ready exchange of methods, reagents, and specimens among participants.

The collection of multiple cases with the same condition allows the activation of a new phase of the project, namely the clinical validation and implementation into routine practice of interpretive post-analytical tools. These tools are directed to the resolution of the most challenging findings encountered in the interpretation of MS/MS results. Improved result interpretation is expected to enhance overall performance, the ultimate goal of this project. Initially, we have focused this effort on the biochemical characterization of carriers in conditions known to have an abnormal biochemical phenotype in heterozygote individuals as well as false negative outcomes [10-13].

The first product of this kind is shown in Figure 7. This tool is based on matching a patient's results against the %ile distribution of the disease range for the same analytes and ratios. In the case of VLCAD deficiency, the informative markers derived from the database of true positive cases are C14:2, C14:1, C14, and the C14:1/C16 ratio. These are the only data required to use the tool. A numerical score is assigned to each marker as it meets and exceeds the chosen %iles of the disease range.

The individual scores are added to others derived from additional interpretive rules (for example C14:1>C14 and C14:1>1 μM) and differential weighting between primary and secondary

markers (not shown). In the example shown in Figure 7, the calculated score correspond to the 11%ile of the scores obtained in confirmed VLCAD cases (N=51), and to the 73%ile of the carrier range (N=16), a result sufficiently abnormal to warrant reporting and confirmatory testing. This particular case (confirmed to have VLCAD deficiency by molecular analysis) was chosen to highlight the ongoing need of additional data to improve the discriminating power of this tool for score between 75 and 100, a range where overlap between the two groups is observed frequently.

On the other hand, it is worthwhile to mention that the quantitative data of only 51 of the 115 available cases of VLCAD deficiency were inclusive of all four necessary results. So, the distribution to participants of this tool also serves the purpose of motivating

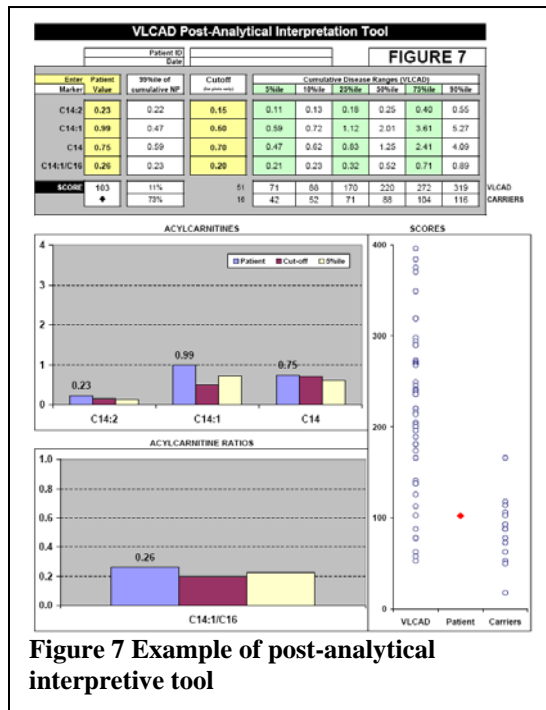


Figure 7 Example of post-analytical interpretive tool

programs to add incremental markers to their analytical and reporting panels, so to take full advantage of these tools. A similar instrument has been prepared to facilitate the segregation of MCAD carriers from affected cases (not shown).

2. Training course at Mayo in small groups

Several individuals from Region 4 and other states have visited the Biochemical Genetics Laboratory (BGL) at Mayo for training in various aspects of newborn screening by MS/MS, with the most emphasis being placed on post-analytical interpretation. In one occasion (April 2005), this was done in a structured format with three attendees (from MI and KY) supported by the Regional Collaborative. The curriculum was a combination of formal lectures, hands-on training in the laboratory, and review of daily batches of newborn screening results. As the demand for this activity has increased significantly since the inclusion in 2006 of additional states from other Regions, we propose to hold over the duration of the grant 5 to 10 one week-long sessions for small groups of project participants, the number of attendees not to exceed five individuals at one time. These individuals will join the PI of the project and BGL staff, when applicable, in performing all activities related to testing (including pre-analytical and analytical components), result interpretation, application of 2nd tier tests, and confirmatory testing. Training sessions will be held also to review processes directly related to the specific objectives of the project, namely data collection and submission through the website, calculation of %iles and performance metrics, and proficient use of the post-analytical tools like the ones described above.

2007 Training Schedule					
Funded through current funding	Feb. 5-9	Feb. 19-23	April 16-20		
Contingent on new funding	June 25-29	July 16-20	Aug. 20-24	Oct 1-5	Dec 3-7

Regional Genetics and Newborn Screening Collaborative (HRSA 04-085 & 07-016)
 REGION 4 COLLABORATIVE PROJECT (Principal Investigator: Cynthia Cameron, PhD)
 Laboratory Quality Improvement of NBS by MS/MS (www.Region4genetics.org)

APPLICATION FORM FOR TRAINING SESSION AT MAYO

Submission date (MM/DD/YYYY) _____ Received: (leave blank) **FIGURE 8**

NOTE: NO MORE THAN TWO (2) INDIVIDUALS FROM THE SAME PROGRAM WILL BE CONSIDERED FOR ATTENDANCE IN 2007

Name _____
 State _____ Region _____ Address: _____

Role in the program (check one only, use "X")
 Laboratory technician
 Technician, primary MS/MS operator
 Primary reviewer of MS/MS results
 Follow up Coordinator
 Supervisor, Laboratory Director
 Other (specify): _____ Phone #: _____
 Fax #: _____
 E-mail: _____

Years of NBS experience in this role: _____ yrs

SELECT 3 DATES (application for these dates will be accepted until January 15, 2007)
 Indicate 1st, 2nd, and 3rd choice by entering 1, 2, and 3 in the box next to the preferred dates

<input type="checkbox"/> February 5-5, 2007	<input type="checkbox"/> July 16-20, 2007
<input type="checkbox"/> February 19-23, 2007	<input type="checkbox"/> August 20-24, 2007
<input type="checkbox"/> April 16-20, 2007	<input type="checkbox"/> September 24-28, 2007
<input type="checkbox"/> June 25-29, 2007	<input type="checkbox"/> December 3-7, 2007

DATA SUBMISSION & PROJECT PARTICIPATION (required) (MM/DD/YYYY)

<input type="checkbox"/> Yes	<input type="checkbox"/> No	True positive cases	Most recent submission
<input type="checkbox"/> Yes	<input type="checkbox"/> No	%iles of normal population	Most recent submission
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Cutoff values	Most recent submission
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Performance metrics	Most recent submission
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Do you have access to the website?	Access to folder (state)
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Have you submitted data yourself?	What type? <input type="checkbox"/> TP <input type="checkbox"/> CO <input type="checkbox"/> %iles
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Have you attended face to face meetings?	(enter "X") <input type="checkbox"/> Performance metrics
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Have you participated in conference calls?	
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Does your program have funding to support your travel and lodging expenses?	
<input type="checkbox"/> Yes	<input type="checkbox"/> No	Did your Program Director approve the submission of this application?	

Date of approval _____ Name of Program Director _____
 Phone #: _____ Fax: _____
 E-mail: _____

Submit completed form to: Piero Rinaldo, MD, PhD E-mail (preferred) rinaldo@mayo.edu
 Biochemical Genetics Lab, Mayo Clinic, Rochester (MN) Fax # (507) 538-7060

Figure 8 Application form for training sessions at Mayo

The first three dates are based on current demand and resource made available to states in Region 3 and 5 through supplemental funding. Additional dates will be arranged in the following years based on the number of requests made by current and possible future participants. The application process for attendance at one of the sessions consists of three components:

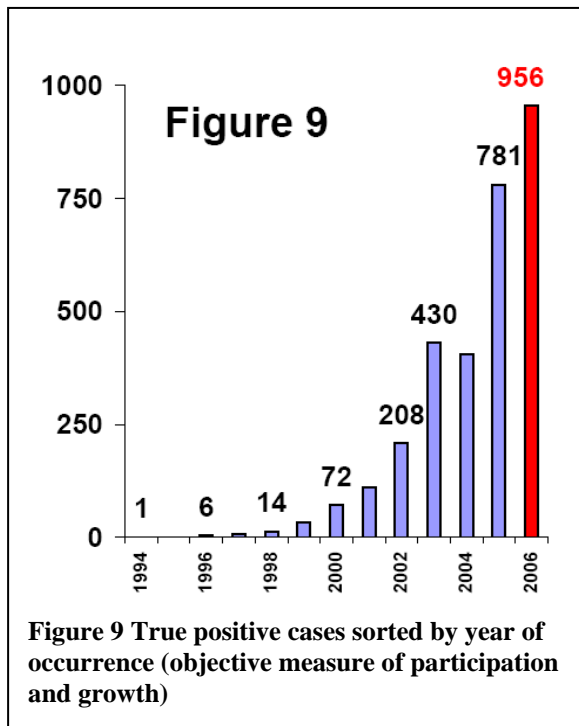
- Application form (Figure 8);
- Formal commitment from the leadership of the program of employment stating that funds are available to cover travel and lodging expenses;
- Evidence of good standing of the laboratory with regard of submission of all data related to the collaborative project (true positive cases, %iles of normal population, updated cutoff values, and performance metrics).

There is no competitive selection involved. Following submission of an application by a member of an active participant's program, all

requests from eligible individuals will be accommodated to the best of the ability of the Mayo staff. Additional dates could be arranged if the demand exceed current estimates. However, no more than two individuals from the same program will be considered for attendance in 2007 and following calendar years for the duration of the grant, and eligibility requirements (i.e., active and complete submission of data) will be strictly enforced.

3. Development of customized software to manage data collection, analysis, and reporting of NBS data

As December 31 2006, the true positive database includes 158,858 single data points. The cutoff repository includes 1,663 values, the database of %iles (currently 4,817 data points) is expected to grow to include up to 27,000 data points (based on the current number of participants, 83). Addition of a single new participant translates into a maximum of 325 more %ile data points, up to 68 more cutoff values, and up to 103 results for each true positive case (the magnitude of which could be extrapolated from the testing volume and an approximate detection rate of 1:3,000 live births). If participation of all US programs were to be achieved, 1,000 to 1,500 true positive cases could be added each year, corresponding to 100,000 to 150,000 data points assuming a close to complete compliance with the recommended panel of analytes and ratios. Finally, three performance metrics (plus period and volume definition) are required for full compliance with data submission of a given participant. So far, all of these data have been manually entered into the corresponding databases by Dr. Rinaldo, who has also been responsible for the creation and maintenance of the database framework using standard Microsoft® Excel format. The current setup is robust and sufficiently redundant to minimize the risk of data loss due to unexpected events. Links created between the databases and the different reports allow automated updates every time a file is opened and saved. All files reside on a secure Mayo server protected by daily backup.



In consideration of the rate of submissions (Figure 9; cases are sorted according to the reported year of testing; number of 2006 cases not final) and the established practice to provide monthly updates using an increasing number of tools, we propose two solutions to manage the situation. The short term solution is funding for 2 years a data entry position (1.0 FTE) who will work closely with Dr. Rinaldo and manage, after proper training, the centralized handling of all submissions and the timely creation and circulation of updated tools.

The request of 1.0 FTE is a conservative estimate if one considers the additional types of reports that could be generated if more time for data processing was available.

As a long term solution to the issue of data collection and analysis, also a mechanism to

ensure long-term sustainability, we propose to proceed with the custom creation of a dedicated computer program capable of meeting the needs of the project according to the following operating principles:

- a) web-based access, password protected, large (100-200) number of total users with the expectation of multiple simultaneous log-ins
- b) peripheral data submission (participants enter own data, not done centrally)
- c) automation of administrative functions (e-mail reminders, monthly posting of general updates)
- d) on demand, user-driven production of project tools (score cards, cutoff tool, plots)
- e) easy generation of customize reports (comparison of own data vs. cumulative data)
- f) flexibility to add new conditions and markers (with potential applicability beyond MS/MS panel), and to query the database to generate novel reports.

An estimate of the cost of developing this software and subsequent maintenance has been derived from separate quotes by the Mayo Research Computing Facility (Steven Monk) and the Mayo Bioinformatics unit (Dr. Jean-Pierre Kocher) and is included in the budget justification section. The definition of coding requirements, programming, and testing of the final product is expected to take up to two years (June 2007 to May 2009). In the third year, we anticipate the need to make adjustments and improvements based on user feedback.

While Region 4 will provide the majority of the funding for the software development and hardware purchase, Regions 1, 2, 3, 5 and 6 have agreed to provide \$10,000 each in Year 1 and \$1,000 each in Years 2-5 to offset the costs.

4. Collection, compilation and monitoring of performance metrics, with definition of targets of acceptable performance

The progress achieved in data collection during the first cycle was not adequately matched by collection of performance metric data, which has been very limited. Improvement of this aspect of the project will become a focus of the next phase. We are confident that the large number of participants, and the increasing familiarity with the objective of the project, could provide a sufficient level of anonymity so that participants will agree to share “sensitive” information about the performance of their program. In particular, during 1Q of 2007 participants will be asked to provide a summary of their 2006 performance metrics, which will be used to generate a baseline of reference for the second cycle of the project. We propose a routine calculation, monitoring, and objective inter-laboratory comparison of three key performance metrics: detection rate (prevalence; collective and for individual condition), false positive rate, and positive predictive value [14].

The detection rate of a newborn screening program is expressed as the number of neonates that on average needs to be tested to detect one affected patient. Published data relative to the outcome of MS/MS screening are sparse, and often widely variable. Until an even greater uniformity of testing is accomplished, the collective detection rate is likely to represent the least informative of available metrics, so we anticipate the need to express the detection rate separately for each condition. This process will be implemented by generating program-specific reports of 2006 data, and each year afterward. The prevalence figures will be compared and monitored, with feedback provided to the group as well as to individual participants. The false

positive rate of a newborn screening program is expressed here as the proportion of positive tests in subjects proven by follow up evaluation not to have one of the conditions targeted by a given screening program. For the purpose of objective comparison, we believe that all cases requiring follow-up testing beyond the analysis of the initial blood spot should be included in the determination of the false positive rate over a defined period of time, with the exception of specimens considered unsatisfactory because they were collected at less than 24 hours of age. The positive predictive value of a test is the probability that the patient has the disease when restricted to those patients who test positive. This metric is less dependent of the extent of the panel of conditions, and should not be arbitrarily restricted to subsets of cases. These metrics, of course, have been applied to newborn screening for decades, but apparently they have been kept local, restricted to own longitudinal assessment (variations over time in a single lab), with limited interest to seek inter-laboratory comparison as an objective mean to define what constitutes an acceptable, or at least average, performance. Currently, the targets we considered indicative of adequate performance are as follows: 1) Detection rate <1:3,000 births 2) False positive rate <0.3% 3) Positive predictive value >20%. A simple tool for calculation in a standardized format and reporting will be circulated to all participants in early 2007.

5. Monthly conference calls and bi-annual face-to-face meetings

One of the important outcomes of the first three years of this project is the ongoing communication among lab personnel in Region 4. This was brought about through monthly conference calls and face-to-face meetings, which provided a venue for participants to get to know each other. Region 4 participants now use each other as resources and call and email each other on a regular basis with questions and/or problems (personal communication, Nancy Breen, Indiana NBS lab). In the new grant cycle, participants from all regions will be invited to join the conference calls. Conference calls will be held every other month and participants will be polled to identify agenda items to ensure that issues relevant to lab personnel are addressed. Region 4 will provide the funding for these calls.

Face-to-face meetings have been held on a bi-annual basis and are “. . . the most useful tools for information gather and brainstorming” (personal communication, Nancy Breen). In October, 2006 participation in the Region 4 meeting was expanded to include representatives from Regions 3 and 5. A meeting is currently being scheduled in conjunction with the Newborn Screening and Genetics Testing Symposium in Minneapolis in May 2007. To date, representatives from 14 states in addition to the 7 Region 4 state have indicated they will attend. Beginning in 2007-08, bi-annual one-day meetings will be held with participants from all Regions invited. Region 4 will provide funding for the meeting room, food and travel for Region 4 representatives. Regions 1, 2, 3, 5, and 6 have confirmed that they will pay for travel and lodging for participants from their regions to attend the face-to-face meetings. The agenda will be set by Dr. Rinaldo and participants in the project. The meetings will be facilitated by the Region 4 Director, Cynthia Cameron, PhD.

6. Continuing clinical validation of 2nd tier tests

Several participants have embraced the strategy to reduce their false positive rate by the implementation of 2nd tier tests, which are defined as reflex tests performed on the same dried blood spot used for the primary screening, without additional patient contact. The trigger to perform the 2nd tier test could be a result in the range of values corresponding to the overlap

between normal (unaffected) population and true positives, or an intrinsic poor specificity of the primary screening (for example, 17-OH progesterone measured by fluoroimmuno assays). Most of the analytical development and clinical validation of these tests has been performed in the Biochemical Genetics Laboratory at Mayo [15-19], however both on site testing and outsourcing are available to all participants. During the first phase of the grant (2004-07) the tests sent to Mayo were performed on a research basis, partially supported by the grant. Recently, however, three states (KY, OK, MI) have elected to transition to a routine clinical utilization, on a fee for service basis.

The table below summarizes the level of participation as December 2006:

Region	State	CAH	MMA (*)	HCY (*)	MCA (*)	ALLO-ILE	SUAC
2	New York	+ (I)	-	-	-	-	-
3	Georgia	+ (I)	-	-	-	-	-
4	Illinois	-	-	-	-	-	-
4	Indiana	-	±	-	-	-	±
4	Kentucky	+ (F)	+	+	+ (D)	+	+
4	Michigan	+ (F)	±	-	-	-	+
4	Minnesota	+	+	+	+ (D)	+	+
4	Ohio	-	-	-	-	-	+
4	Wisconsin	-	±	-	-	-	-
5	Missouri	± (F)					
5	Oklahoma	+ (F)	-	-	-	-	-
6	Utah	+ (I)	-	-	-	-	-
7	California	+ (I)	±	±	-	-	-

Table legend (abbreviations listed in alphabetical order): **+**, test implemented (outsourced to Mayo from states other than MN); **+** (D), test still under development but available; **+** (F), test performed on a fee for service basis; **+** (I), test implemented locally; **±**, sporadic utilization; **-**, test not utilized; **ALLO-ILE**, LC-MS/MS analysis of allo-isoleucine for the evaluation of newborns with mild elevation of isoleucine/leucine and/or valine in the amino acid profiles; **CAH**, LC-MS/MS steroid profiling (17-OHP, androstenedione, cortisol, and 11-deoxy cortisol) for the evaluation of newborns with abnormal CAH primary screening by immunoassay; **HCY**, LC-MS/MS analysis of total homocysteine for the evaluation of newborns with elevation of either propionylcarnitine (C3) or methionine (see below for additional information); **MCA**, LC-MS/MS analysis of methylcitric acid for the evaluation of newborns with elevation of C3; **MMA**, LC-MS/MS analysis of methylmalonic acid for the evaluation of newborns with elevation of C3; (*), test included in a common testing algorithm described below; **SUAC**, LC-MS/MS analysis of succinylacetone for the evaluation of newborns with elevation of tyrosine to rule out a possible diagnosis of Tyrosinemia type I.

As previously shown in Figure 5 (page 4), the frequent occurrence of false positive results has driven several programs to raise their cutoff values for propionylcarnitine to levels significantly overlapping with the disease ranges of all conditions linked to this analyte. Even when the cutoff is between 8 and 10 µM, it is common to encounter markedly elevated concentrations in non affected subjects, particularly in newborn with hyperbilirubinemia. This situation has stimulated us to develop multiple 2nd tier test that have the potential to correctly identify each and every one of the possible underlying conditions, with the expectations that negative results of all 2nd tier tests would permit the overruling of the initial result of the primary screening and ultimately a

constant improvement of the positive predictive value for conditions known to have propionylcarnitine (C3) as primary marker.

A critical concept behind this strategy is to report these cases as normal, without delays. The 2nd tier tests are performed in batches, once or twice a week, with an overall turn-around-time of 4-5 days. In the event of a positive test, a revised report is issued and communicated promptly to the primary care provider. In our experience, this approach has been extremely effective in reducing the false positive rate for C3 in the two states (KY and MN) which have implemented the protocol in full.

The table below shows how the 2nd tier tests could lead to the differential diagnosis of as many as 7 conditions (5 primary and 2 secondary targets).

Condition	1st tier			2nd tier		
	C3	C5-OH	Met	MMA	HCY	MCA
Methylmalonic acidemia (mut)	+	-	-	+	-	+
Methylmalonic acidemia (Cbl A,B)	+	-	-	+	-	+
Methylmalonic acidemia (Cbl C,D)	+	-	-	+	+	+
Maternal vitamin B12 deficiency	+	-	-	+	±	±
Propionic acidemia	+	-	-	-	-	+
Multiple carboxylase deficiency	+	+	+	-	-	+
Homocystinuria	-	-	+	-	+	-
Hypermethioninemia	-	-	+	-	-	-
False positive case (C3)	+	-	-	-	-	-

Table legend: +, abnormal result; ±, borderline result; -, normal result; **C3**, propionylcarnitine. Also considered are the C3/C2 ratio and the C3/C16 ratio (not shown); **C5-OH**, hydroxyl isovalerylcarnitine; **HCY**, total homocysteine (abnormal >15 µM); **MCA**, methylcitric acid (definition of cutoff in progress; almost undetectable in normal newborns); **Met**, methionine; **MMA**, methylmalonic acid (abnormal >5 µM). 1st tier results are deemed abnormal on the basis of the cutoff ranges discussed elsewhere.

7. Round robin sample exchange

The validation of the process behind the definition of cutoff ranges is based on objective comparison of diverse data obtained by different laboratories. In addition to the comparison (and assessment of variability) of %iles values in the normal population, we have sought the implementation of a sample exchange initiative, a process where the same specimen is analyzed by the laboratory where the initial diagnosis was made (submitter) and another participant (recipient) chosen in alphabetical round robin order. For example, if Illinois has 6 samples to distribute, the first will go to Indiana, second will go to Kentucky, third will go to Michigan, 4th will go to Minnesota, 5th will go to Ohio, and 6th will go to Wisconsin. Both laboratories submit a copy of their results to Dr. Stephanie Mayfield (KY) and her associates, who compile a summary file where the identity of the recipient is not disclosed. Results are posted quarterly on the Region 4 Collaborative website, sample exchange folder in the password protected section under Newborn screening by MS/MS.

The results are posted showing the submitter, the submitter diagnosis and informative results, and the recipient's diagnosis and values. As December 2006, 76 exchanges have taken place, involving all states in Region 4. The diagnosis concordance so far is 74/76 (97.4%) the

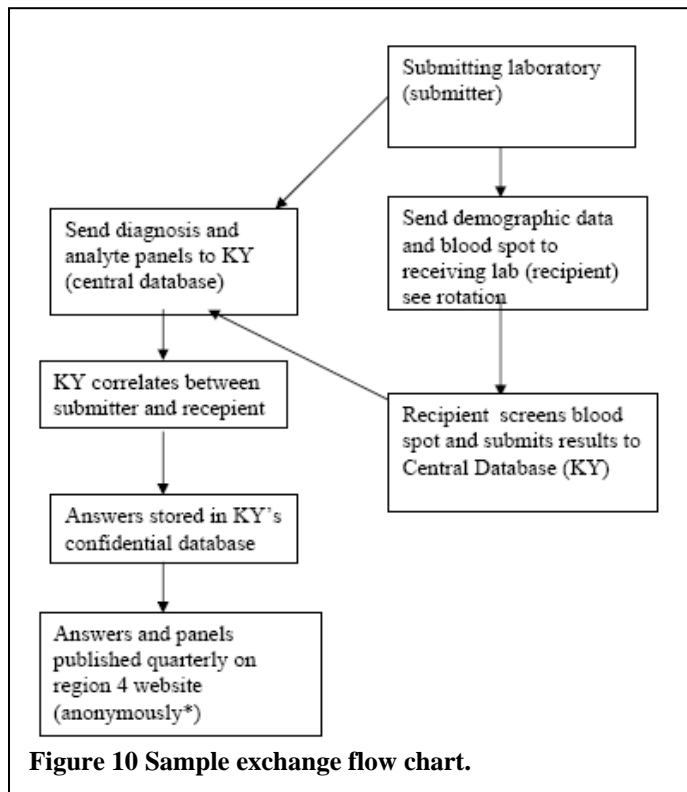


Figure 10 Sample exchange flow chart.

variability between primary markers is below 20%. We have already received requests to join the sample exchange program from additional states outside Region 4, including international participants.

To monitor the impact of the sample exchange initiative, targets have been set as follows:

1. 100% correlation of TP cases with threshold of 95% for current year (2006)
2. 90% of primary analyte(s) values within 20% of submitter's corresponding values
3. 100% active participation from all states within Region 4 in the exchange process
4. Increase participation to include at least 3 additional states outside of Region 4 in 2007.

As December 2006, all the goals listed above have been achieved.

Resolution of challenges

There were many challenges at the beginning of the project, mostly driven by a diffuse skepticism (inside and outside Region 4) that comparison and harmonization could be achieved in a constructive and collegial manner. Furthermore, there were substantial concerns about confidentiality, especially when the collection of data was driven by a private, non public health entity. As the first cycle of the grant is close to completion, the situation has improved dramatically, with a pervasive impression of voluntary desire to actively participate. The critical role of the program director (Cynthia Cameron, PhD) as facilitator of the face-to-face meetings and conference calls cannot be overstated, as she has been able to diffuse tension between participants and at the same time stand firm and curtail inappropriate behaviors. Difficult situations were also addressed in a very effective way by the Regional advisory group (for membership see advisory group section of the project website, <http://www.region4genetics.org>), which provided mediation and sound advice on how to reach resolution of conflicts. State leads have been very effective, too, as they reached out to participants in their state and facilitated their active involvement. To date, the working group is a large, diverse, and vibrant assembly of newborn screening professionals with a wide range of professional backgrounds and variable levels of technical expertise (Attachment 1). Newcomers and less experienced participants are encouraged to interact with all members, and we have seen a growing number of personal

interactions and exchanges of information taking place between programs which had little or no contacts before getting involved in this project.

References

1. Mayfield S, Lim J, Foree V et al. (2006) Expanded NBS in Kentucky: Performance metrics in the first three months. Proceedings of the 6th Meeting of the International Society for Newborn Screening, Awaji, Hyogo & Tokushima, Japan, p. 177 (P133).
2. Watson MS, Mann MY, Lloyd-Puryear MA, et al. [editors]. (2006) Newborn screening: Toward a uniform screening panel and system [Executive summary]. *Genet Med* 8(Supplement):1S-11S.
3. Watson MS, Lloyd-Puryear MA, Mann MY, et al. [editors]. (2006) Newborn screening: Toward a uniform screening panel and system [Main report]. *Genet Med* 8(Supplement):12S-252S.
4. National Newborn Screening and Genetics Resource Center (NNSGRC). (2006) US national newborn screening information system. Available at <http://genes-r-us.uthscsa.edu>.
5. Zytkevich TH, Fitzgerald EF, Marsden D, et al. (2001) Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. *Clin Chem* 47:1945-1955.
6. Wilcken B, Wiley V, Hammond J, et al. (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* 348:2304-2312.
7. Schulze A, Lindner M, Kohlmuller D, et al. (2003) Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics*. 111:1399-406.
8. Hoffmann GF, von Kries R, Klose D, et al. (2004) Frequencies of inherited organic acidurias and disorders of mitochondrial fatty acid transport and oxidation in Germany. *Eur J Pediatr* 163:76-80.
9. Frazier DM, Millington DS, McCandless SE, et al. (2006) The tandem mass spectrometry newborn screening experience in North Carolina: 1997-2005. *J Inherit Metab Dis* 29:76-85.
10. Spiekerkoetter U, Sun B, Zytkevich T, et al. (2003) MS/MS-based newborn and family screening detects asymptomatic patients with very-long-chain acyl-CoA dehydrogenase deficiency. *J Pediatr* 143:335-342.
11. Liebig M, Schymik I, Mueller M, et al. (2006) Neonatal screening for very long-chain acyl-coA dehydrogenase deficiency: enzymatic and molecular evaluation of neonates with elevated C14:1-carnitine levels. *Pediatrics* 118:1065-1069.
12. Schymik I, Liebig M, Mueller M, et al. (2006) Pitfalls of neonatal screening for very-long-chain acyl-CoA dehydrogenase deficiency using tandem mass spectrometry. *J Pediatr* 149:128-1230.
13. Boneh A, Andresen BS, Gregersen N, et al. (2006) VLCAD deficiency: pitfalls in newborn screening and confirmation of diagnosis by mutation analysis. *Mol Genet Metab* 88:166-170.
14. Rinaldo P, Zafari S, Tortorelli S, Matern D. (2006) Making the case for objective performance metrics in newborn screening by tandem mass spectrometry. *MRDD Research Reviews* 12:255-261.
15. Lacey JM, Minutti CZ, Magera MJ, et al. (2004) Improved specificity of newborn screening for congenital adrenal hyperplasia by second tier steroid profiling using tandem mass spectrometry. *Clin Chem* 50:621-625.
16. Minutti CZ, Lacey JM, Magera MJ, et al. (2004) Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol & Metab* 89:3687-3693.
17. Cuthbert CD, Magera MJ, Hahn SH, et al. (2005) Analysis of homocysteine and methylmalonic acid in blood spots by tandem mass spectrometry as a second tier newborn screening test for homocystinuria and methylmalonic acidemias. Proceedings of the Annual Meeting of the American College of Medical Genetics, Dallas, TX .
18. Magera MJ, Gunawardena ND, Hahn SH, et al. (2006) Rapid quantitative determination of succinylacetone in dried blood spots by liquid chromatography tandem mass spectrometry. *Mol Genet Metab* 88:16-21.
19. Oglesbee D, Lacey JM, Spolar CD, et al. (2006) Newborn screening for MSUD: increased specificity by addition of 2ND-tier assay for allosoleucine by LC-MS/MS (Abstract O-14-4). *J Inherit Metab Dis* 29(Suppl 1):55.