



Overview of Cutoff Determinations and Risk Assessment Methods used in Dried Blood Spot Newborn Screening- Role of Cutoffs and Other Methods of Data Analysis

This publication was supported by Cooperative Agreement # 5NU600E000103 funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of CDC or the Department of Health and Human Services.

Contents



APHL ASSOCIATION OF PUBLIC HEALTH LABORATORIES 0

Acknowledgements..... 2

Acronyms 3

1. Purpose 4

 1.1 Limitations..... 4

2. Background 4

 2.1 Cutoffs..... 5

3. Overview of Cutoff Determination 7

 3.1 Testing Dried Blood Spots from Newborns..... 7

 3.2 Analyzing Data and Determining a Preliminary Cutoff 7

 3.3 Validating or Verifying the Cutoff 9

 3.4 Special Considerations 10

 3.5 Collaborative Laboratory Integrated Reports 12

 3.6 Summary of Methods 13

4. Disorder Specific Considerations 16

 4.1 Endocrine Disorders (Congenital Hypothyroidism, Congenital Adrenal Hyperplasia)..... 16

 4.2 Cystic Fibrosis..... 17

 4.3 Hemoglobinopathies..... 17

 4.4 Enzyme Deficiency Disorders..... 18

 4.5 Amino Acid, Fatty Acid Oxidation and Organic Acid disorders 18

 4.6 Severe Combined Immunodeficiency (SCID) 19

 4.7 X-Linked Adrenoleukodystrophy (X-ALD)..... 19

5. Monitoring and Ongoing Evaluation of Risk Assessment 20

6. References..... 21

Acknowledgements

APHL would like to recognize the following individuals that generously contributed to the development of this document and the members of the APHL Newborn Screening and Genetics in Public Health Committee and the APHL Newborn Screening Quality Assurance and Quality Control Subcommittee for their time and input.

Mei Baker, MD, FACMG Wisconsin State Laboratory of Hygiene
Joanne Bartkus, PhD, Minnesota Department of Health
Christianne Biggs, MS, Oregon State Public Health Laboratory
Natasha Bonhomme, Genetic Alliance
Michele Caggana, ScD, FACMG, New York State Department of Health- Wadsworth Center
Neena Champaigne, MD, Greenwood Genetics Center
Bob Currier, PhD, California Department of Public Health
Carla Cuthbert, PhD, FCCMG, FACMG, Newborn Screening Molecular Biology Branch, Centers for Disease Control and Prevention (CDC)
Tim Davis, Washington State Department of Health
George Dizikes, PhD, HCLD/CC (ABB), Tennessee Department of Health
Christine Dorley, MSP, MT (ASCP) Tennessee Department of Health
Amy Gaviglio, MS, CGC, Minnesota Department of Health
Fizza Gulamali-Majid, PhD, Maryland Department of Health and Mental Hygiene
Patrice Held, PhD, Wisconsin State Laboratory of Hygiene
Amy Hietala, MS, Minnesota Department of Health
Patrick Hopkins, Missouri Public Health Laboratory
Patricia Hunt, Texas Department of State Health Services
Rachel Lee, PhD, Texas Department of State Health Services
Adrienne Manning, Connecticut Department of Public Health Laboratory
Joanne Mei, PhD, Newborn Screening Molecular Biology Branch, CDC
Darren Michael, PhD, Colorado Department of Public Health and Environment
Christine Moore, Texas Department of State Health Services
Joseph Orsini, PhD, New York State Department of Health- Wadsworth Center
Hari Patel, MS, North Carolina Department of Health and Human Services
Kostas Petritis, PhD, Newborn Screening Molecular Biology Branch, CDC
Michael Ramirez, State Hygienic Laboratory at The University of Iowa
Deboshree Sarkar, MPH, Division of Services for Children with Special Healthcare Needs, Health Resources and Services Administration
Inderneel Sahai, MD, University of Massachusetts Medical School
Santosh Shaunak, Washington State Department of Health
Eleanor Stanley, MT (ASCP), Michigan Department of Health and Human Services
Bonita Taffe, PhD, Florida Department of Health
Susan Tanskley, PhD, Texas Department of State Health Services

Acronyms

The following is a list of acronyms used throughout this document:

AA	Amino Acid
AC	Acylcarnitine
APHL	Association of Public Health Laboratories
ANOVA	Analysis of Variance
CAH	Congenital Adrenal Hyperplasia
CCHD	Critical Congenital Heart Disease
CDC	Centers for Disease Control and Prevention
CH	Congenital Hypothyroidism
CF	Cystic Fibrosis
CLIA	Clinical Laboratory Improvement Amendments
CLIR	Collaborative Laboratory Integrated Reports
CLSI	Clinical and Laboratory Standards Institute
Cq	Cycle of quantitation
DBS	Dried Blood Spot
FIA	Flow Injection Analysis
Hb	Hemoglobinopathies
HPLC	High Performance Liquid Chromatography
IEF	Isoelectric Focusing
IRT	Immunoreactive Trypsinogen
LIMS	Laboratory Information Management System
LPCs	Lysophosphatidylcholines
LSDs	Lysosomal Storage Disorders
MoM	Multiple of the [Normal Population] Median
MS/MS	Tandem Mass Spectrometry
NBS	Newborn Screening
NewSTEPS	Newborn Screening Technical assistance and Evaluation Program
NICU	Neonatal Intensive Care Unit
NSQAP	Newborn Screening Quality Assurance Program
QA	Quality Assurance
QC	Quality Control
R4S	Region 4 Stork Laboratory Performance Database
RUSP	Recommended Uniform Screening Panel
SAS	Statistical Analysis Software
SCID	Severe Combined Immunodeficiency
SOPs	Standard Operating Procedures
TPN	Total Parenteral Nutrition
TRECs	T-cell receptor excision circles
X-ALD	X-Linked Adrenoleukodystrophy
µL	Microliter

1. Purpose

The purpose of this document is to provide a general overview of the variety of approaches utilized by newborn screening (NBS) laboratories and program staff to determine if a newborn is at risk for a screened disorder. An overview of historical and current approaches that laboratories rely on for risk assessment are described, as well as factors that should be considered when establishing and evaluating risk. This document is not intended to provide detailed instructions on performing risk assessment in NBS. The intended audience for this document is state NBS programs; it assumes that the reader has a solid understanding of NBS laboratory methodologies and risk determination. For a basic explanation of how risk is determined and why determinations may vary across states, refer to “Making Sense of Newborn Screening Cutoff Values,” available via the Association of Public Health Laboratories (APHL) Lab Blog.⁽¹⁾

1.1 Limitations

Newborn screening is performed as a means to assess risk in clinically asymptomatic newborns for a host of disorders⁽²⁾; it is not intended to establish diagnosis. Abnormal biomarker levels, identified through screening and evaluated using cutoffs, only indicate that a newborn may be at increased risk for a screened disorder. This information is communicated to the healthcare provider, who may request additional diagnostic testing to determine if the newborn has the disorder in question. Healthcare providers should understand that in instances where a newborn has a family history of a disease or is symptomatic, additional diagnostic testing is necessary regardless of the NBS result. Even though algorithms are used to determine infants at highest risk, NBS may not detect all affected newborns.⁽³⁻⁶⁾ A positive (i.e., abnormal) screening result is not a confirmed diagnosis of a disorder, and a negative (i.e., normal) screening result is not a confirmed exclusion of a disorder.

2. Background

Newborn Screening is a public health program that identifies infants at risk for certain conditions that may lead to death or disability if undetected or left untreated. NBS includes analysis of biochemical or molecular markers in dried blood spots (DBS) conducted at a laboratory in addition to point-of-care screening performed at birthing facilities (i.e., hearing and critical congenital heart disease (CCHD) screening). This document will focus on the analysis of biomarkers in DBS.

The purpose of screening is to identify individuals affected by a specific disorder from within an apparently healthy population prior to the appearance of symptoms, so that appropriate interventions can be initiated in time to minimize or avoid the harmful effects of the disorder.

The process of screening, as suggested by Sir Nicholas Wald, “is the systematic application of a test or inquiry, to identify individuals at sufficient risk of a specific disorder to benefit from further investigation or direct preventive action, among persons who have not sought medical attention on account of symptoms of that disorder.”⁽⁷⁾

The objective of screening is to determine the likelihood that a given specific analyte value obtained from a particular specimen came from an affected or unaffected patient; so that a determination of “sufficient risk” can be established for the specific condition being screened for (i.e., making a risk assessment); so that a patient determined to have an elevated risk can be referred for diagnostic testing

and/or interventions can be initiated. This objective is made more difficult because the distributions of analyte values from the unaffected and affected populations, which are used for risk assessment, typically overlap. The analyte value used to differentiate between high risk and low risk is referred to as the “Cutoff Value” for the specified condition.

Because the majority of disorders included on State newborn screening panels are rare (less than 1 in 3000 live births) and it requires a statistically significant number of affected cases to accurately identify the range of analyte values associated with high risk for a specific condition, there are significant barriers to set a cutoff value that has been mathematically calculated from the distributions of affected and unaffected infants. Thus, individual State NBS Programs must rely upon many of the methods described in this document to determine cut-offs for their State population.

The DBS specimen is an inherently complex matrix, with variables including hematocrit and leukocyte counts that can have broad effects on screening results and cutoff values.⁽⁸⁾ Additionally, for screened conditions, there is a clinical spectrum of phenotypes that can impact the biomarker levels and therefore the laboratory’s ability to detect the disorder in the newborn period. In general, marker levels of newborns with mild phenotypes are closer to the marker levels in the reference population than those with classic disorder phenotypes. Establishing an algorithm to detect mild cases may increase the false positive rate (decreasing the positive predictive value) of the screen – an important consideration when balancing the intent, benefits and harms of screens and the impact of an abundance of false positive results on the public health system.

Newborns with positive screening results are referred for diagnostic testing. In some cases, such as the screen resulting in a borderline result, an additional specimen is requested so that the screening test can be repeated. DBS screening, like any screening test, may not identify every instance of disease, as pre-analytical and analytical variables described in section 2.1 below may impact the results.

2.1 Cutoffs

Traditionally, cutoff values have been used to determine if a newborn is at risk for a disorder. To establish cutoffs, the screening laboratory must first establish a reference range for the disease marker(s) used. The reference range is the range of marker concentrations that would be expected in a healthy population of newborns measured by the laboratory for the specified method. A cutoff value may be at the low or the high end of the reference range, depending on the biomarker being evaluated and the disorder being screened. The cutoff may identify a reduction of analyte (i.e., biomarker is low) or an accumulation of analyte (i.e., biomarker is high).

In NBS, cutoff values may be fixed or floating. A fixed cutoff is a numerical value that is set in concentration units (for example 0.4 mg/L) , which do not change on a daily basis. Another example of a fixed cutoff is where the patient analyte value has been transformed into a unitless number by dividing the patient analyte value by the median analyte value previously determined for the normal (i.e., unaffected) population. The patient result is then reported as a multiple of median (MoM) and used as the screening variable with an associated fixed cutoff. In contrast, a floating cutoff will change on a daily or per batch basis and is typically based on a percentile of the marker, such as the top 5% of the biomarker level that is measured on a particular day. Other examples include the percent of daily mean or percent of daily median.

Due to analytical and other variables listed below, reference ranges and subsequently determined cutoffs will vary between NBS programs and must be established or verified by each individual laboratory for each test performed.^(9,10) The performance and reliability of the screen is not based on cutoff differences alone. Each screening laboratory should set cutoffs based on the expectations of the screening assay, while considering the phenotype of the disorder being screened and the sensitivity and selectivity of the screen. Programs should also weigh the impact of false positive and false negative rates on system capacity, in addition to the factors below:

- A. Age and health status of the newborn at time of specimen collection
- B. Gestational age, prematurity, and birth weight
- C. Environmental conditions during DBS transport
- D. The collective experience of NBS programs screening for the disorder, or if it is a disorder that is not yet widely screened nationally
- E. The specific method used for the analysis
- F. Instrument platform
- G. Sensitivity of the instrument
- H. Stability of the measured analytes
- I. Availability of reference materials
- J. Calibration material or the internal standard used
- K. Target condition (time-critical and severe to mild cases)
- L. Program policies with respect to the screened conditions
- M. The laboratory's screening algorithm including:
 - 1. Use of multiple markers and ratios to determine screen results
 - 2. Gender, race and ethnicities (for select conditions) within the target population
 - 3. Use of second-tier screening test(s)

Second-tier screening is necessary with some screening tests if the primary marker used for screening is not sufficiently specific, as is the case with cystic fibrosis screening using the primary biomarker immunoreactive trypsinogen (IRT). To decrease the false positive rate, NBS programs may implement second-tier testing,⁽¹¹⁻¹⁵⁾ which provides a more selective screen and may improve positive predictive values. Second-tier tests use the same DBS that was collected for the first-tier screen and are of value in reducing overall turnaround time, as no additional patient contact is necessary.

The factors listed above, and others will affect the reference range, which includes the calculated mean, median and other statistically derived values of biomarker concentrations measured in a population study. Additional variables to be considered when setting NBS cutoffs include:

- A. Feeding status (e.g., breast milk, soy-based formula, TPN)
- B. Transfusion status and time post-transfusion
- C. Whether the NBS program is a one-screen or two-screen program
- D. Analyte results from initial instrument validations, including true positive and false negative cases and quality assurance results from the platform to be used
- E. Instrument variability and whether instrument-specific cutoffs versus method-specific cutoffs will be utilized
- F. Nature of the screening test (i.e., enzyme or immunoassay, vs. direct measurement of biomarker concentration such as in tandem mass spectrometry (MS/MS) for amino acid disorders)

3. Overview of Cutoff Determination

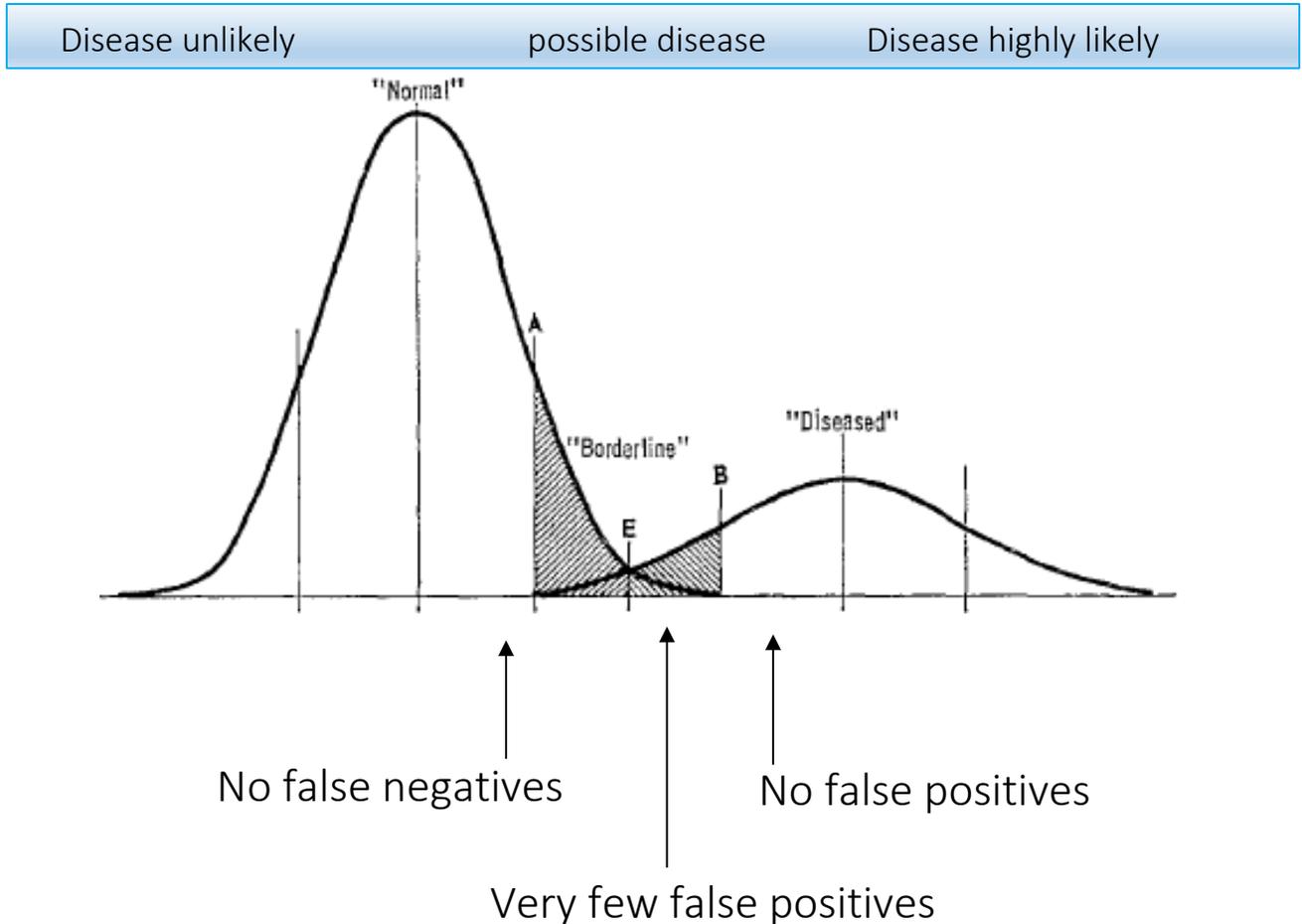
Generally, a cutoff value may be determined by: 1) testing DBS specimens from unaffected newborns and using manufactured controls (made in-house, purchased commercially, or obtained from the [Newborn Screening Quality Assurance Program \(NSQAP\)](#), Atlanta, GA⁽¹⁶⁾; 2) analyzing data to determine whether the screen has adequate precision and accuracy; 3) assigning a preliminary cutoff value; 4) validating or verifying the cutoff using known positive control specimens, obtained either from an external source (i.e., NSQAP) or preferably residual DBS from a confirmed patient, when available, and; 5) comparing cutoff values to other programs.⁽¹⁷⁾ The steps below summarize the process for determining cutoff values.

3.1 Testing Dried Blood Spots from Newborns

Perform a population study using manufactured controls and DBS specimens from presumed unaffected newborns. Obtain residual DBS that offer minimal limitations related to storage conditions and age (i.e., recently obtained residual DBS perform better than their older counterparts). Sample size for a population study will vary depending on several factors, including a laboratory's methodological guidelines, confidence that the data is representative of the population and/or the method's run-to-run variability, the state's birth rate, disease incidence and the screening implementation timeframe. The general practice is to analyze hundreds to thousands of specimens. Larger sample sizes may be needed to evaluate birth weight, age, seasonal effects and other variables. Specimens should be tested on all instruments that will be used to perform routine screening to understand and compensate for instrument-to-instrument variability prior to the start of screening (instrument matching protocols should be in place).^(10,18,19,20)

3.2 Analyzing Data and Determining a Preliminary Cutoff for Lab Developed Tests

- A.** Evaluate whether the analytical precision of the method is adequate to differentiate results close to the cutoff (see B.1. below). Consider running a higher sensitivity method or more replicates (punches from same DBS) to ensure precision of the final result. For qualitative tests, determine percentages of results identified as normal and abnormal before setting a preliminary cutoff. Given limited availability of true positive patient specimens, it is recommended to wait to test those specimens after the precision and accuracy of the methods have been evaluated and deemed acceptable.



- B.** For quantitative or semi-quantitative tests, use statistical software such as Excel, Statistical Analysis Software (SAS), Laboratory Information Management Systems (LIMS) vendor software or others to calculate the mean, median, standard deviation and possibly the percentiles of the data set. These functions can also be calculated using the log transformed data.
1. Remove outliers prior to the calculations.
 2. Using the manufactured controls, evaluate the precision of the assay (and the accuracy if possible). Generally, the precision is most important with an acceptable coefficient of variation (%CV) of $< \pm 20\%$ for the simulated positive controls.
 3. If using log transformed data, the antilog of the statistic can be used to characterize the reference range and cutoff.
 4. The mean or median plus a number of standard deviations can be calculated and used to represent the reference range (e.g., the median plus 3 standard deviations).
 5. View the data graphically by creating a frequency histogram or probability density function; for example, plot the x-axis with the range of analytical values and the y-axis with the number of specimens.
 6. Check population data to determine if it has a normal distribution (e.g., mean/median = 1.0).
 7. Determine if demographic factors (such as birth weight, gestational age, Neonatal Intensive Care Unit (NICU) status, transfusion status, feeding status or the infant's age at time of specimen collection) impact the distribution. For example, this evaluation can be done by

filtering out all specimens from infants with birthweights < 1,500 g, calculating the mean, median, standard deviation of the data set and ranges, and comparing to the normal birth weight population. Also compare histograms for each sub-population. Analysis of variance (ANOVA) methodology can be used to determine if there are statistically significant differences between sub-populations. If the data is statistically different, consider using the variable (e.g., birth weight) in setting the cutoff(s).

8. The preliminary cutoffs can be set using percentiles (e.g., > 99.9 or other percent), number of standard deviations above or below the mean or median as compared to other laboratory results, or by choosing a fixed or floating cutoff that is expected or demonstrated based on comparison of the results to the manufactured controls or other available information (e.g., from diagnostic laboratory results, where comparison of affected and unaffected ranges are already established).
9. A preliminary, conservative cutoff can be set by comparing the normal and positive specimen results (manufactured positive controls or other positives that were tested in the population study). A conservative cutoff should allow for detection of all true positives, should not create a high projected screen-positive rate relative to the expected incidence rate for the associated condition.

3.3 Validating or Verifying the Cutoff

A. Challenge the preliminary cutoff

1. When possible, test known true positive, false positive, carrier, pseudodeficiency and poor quality/diluted specimens from stored residual NBS specimens. True positive specimens or blinded patient panels may also be requested from other laboratories.
2. For some disorders, true positive specimens from older patients can be used (e.g., galactose-1-phosphate uridylyltransferase (GALT) or cystic fibrosis mutations), but only with a thorough understanding of the disorder and how the marker is affected by patient age. Specimens from older patients may not yield similar results to the values determined in newborns.
3. Test proficiency specimens from the CDC NSQAP⁽¹⁶⁾ and compare the results to those obtained by other NBS programs.

B. Compare the cutoff

1. Contact other laboratories that are using a similar method.
2. For laboratories using an FDA cleared or authorized kit, compare the calculated cutoff to the cutoff that is provided in the package insert.
3. The Region 4 Stork (R4S) Laboratory Performance Database and its successor Clinical Laboratory Integrated Reports (CLIR)^(17,21) can be used to compare population reference data with other programs for MS/MS, lysosomal storage disorders (LSDs) and adrenoleukodystrophy screening methods. Note that R4S is currently available as of July 2018, but this software will be phased out and replaced by the second generation of the software, CLIR.

C. Conduct a literature search to identify published cutoffs or control ranges, disease ranges, and the prevalence and incidence of the disorder. This information can be used to estimate the screen-positive rate and be compared to the incidence rate.

D. Evaluate results of the population study

1. Compare results from true positive specimens to results from the population study.

2. Determine the proportion of true positive specimens tested as part of the study that would be detected at the chosen cutoff. Ideally, all positive specimens will be detected. If the results are not acceptable, modify the cutoff so that all will be detected; or consider the quality of the true positive specimens, certainty of diagnosis and associated phenotype prior to adjusting the cutoff. True positive specimens with results that are close to the cutoff can be most informative and are useful to determine if the precision of the method is acceptable at the corresponding specimens marker level.
- E. Set the cutoff**
1. Determine whether the cutoff will be floating or fixed (see below section "[Special Considerations](#)"). If other laboratories are using the same assay, contact them to discuss their use of fixed or floating cutoffs.
 2. Consider the projected number of screen-positive newborns at risk for the disorder and consider consulting with your medical providers to ensure this number is acceptable to them.
 3. Consider setting conservative cutoffs initially to minimize the possibility of false negative test results. However, also consider that if a conservative cutoff is chosen, the screen may lead to the detection of newborns with mild or late-onset disease that may or may not develop until later in life, such as with LSDs. Determine if this is an acceptable component of reported results for your program.
 4. Working with clinical specialists, closely monitor the follow-up results of screen positive infants. Is there any correlation in clinical results with screen results (degree of out-of-normal range)? Is phenotype mild or severe? Are there other factors that may have affected the screen results such as prematurity or age of the newborn at the time of specimen collection? Continue to collect as much follow-up data on the referred infants and work with the clinicians and advisory board to adjust the cutoffs if necessary.

Cutoffs are not static; they should be monitored and adjusted based on feedback from short- and long- term follow up and clinical outcomes of infants, when new information about the disorder becomes available or when significant changes are made in the testing process that may affect the reference range. This includes events like revisions to analytical methods, instrument replacements and/or new substitutions to key consumables or chemicals used during testing.

3.4 Special Considerations

- A.** If the laboratory is the first to develop a method to screen for a new disorder, setting cutoffs can be particularly challenging.
1. The quality of the positive patient specimens is important. Use original newborn DBS specimens, if available.
 2. Consider the phenotype of the individuals used as positive patient control specimens. Usually, positive specimens from individuals with a classic or severe phenotype of the disease are available for testing. Specimens from patients with milder phenotypes may not have similar marker values. Prior to establishing a cutoff value, gather input from specialists to determine if mild or late-onset phenotypes should be identified.
 3. Perform a literature review to determine diagnostic methods currently used, normal and abnormal reference ranges and the specimen matrix used for diagnosis. Often, the

same biomarkers are used in diagnostic testing and NBS, and the associated reference ranges used in the diagnostic test can be a useful guide in setting the NBS cutoffs.

4. Determine where the marker resides in the blood (serum, red cells, or white cells). This will help predict the potential for variability on the measured values and differences in diagnostic versus screening reference ranges. For example, when testing for galactosemia using an assay that measures GALT enzyme activity, GALT is present only in red cells, so a negative screen may be a false negative if the newborn was transfused with red cells prior to obtaining the DBS specimen.
- B.** If the test has been well established by another laboratory, consider performing an exchange with a panel of normal, abnormal, and, if applicable, carrier or pseudodeficiency specimens. Reference ranges and results of the specimen exchange can be used to determine concordance between the two data sets. For example, if there is systematic bias observed in the comparison test, the results can be used to extrapolate reference ranges for the varied cohorts and in setting cutoffs, even if very few positive or other controls were used in the comparison.
- C.** If the condition is not imminently life threatening, NBS programs may set borderline cutoffs. A borderline cutoff is a range of values between normal and referral level marker concentrations. If a specimen is found to be in the borderline range, another DBS specimen is requested rather than referring the newborn for follow-up diagnostic testing. The borderline cutoffs are helpful when the marker being measured may increase or decrease depending on the age of the newborn at the time of specimen collection, or if the marker tends to be elevated at birth due to stress on the newborn resulting from the birthing process (such as 17-OHP). The CDC is in the early stages of developing quality assurance DBS materials that mimic historically confirmed cases that were close to NBS cutoff values. The intent of this initiative is to provide NBS programs with positive controls that produce results in the borderline range. Often, there are few true positive specimens that can be shared with other laboratories to challenge the cutoff, and there is limited residual specimen for those that do exist. These CDC-produced DBS will allow laboratories to challenge their cutoffs.
- D.** Determine whether to use fixed or floating cutoffs. Generally, fixed cutoffs are used for assays that directly measure marker concentrations (e.g., MS/MS analysis of phenylalanine). Floating cutoffs are used for indirect or functional assays, such as enzyme, or immunoassays where the concentration of the marker is not directly measured by the instrumentation and is dependent on an enzyme reaction or binding of an antibody to an antigen. With the latter assays, more uncontrolled variables can affect the day-to-day or lot-to-lot performance of the assay and the use of a floating cutoff compensates for these variables. With floating cutoffs, it is important to consider if variability of the assay is more day-to-day or lot-to-lot, and if the daily or population mean/median will be used. For example, using percent of daily mean or median would be more valuable if there is daily variability. The use of median may be preferred over mean due to it being less affected by outliers in the data set.

Some laboratories transform the patient analyte value into a Multiple of Median (MoM) by dividing the patient's analyte value by the statistically determined normal (i.e., unaffected) population median. This unitless value is then used as the screening variable and an appropriate cutoff is established as a multiple of the normal population median. As with other fixed cutoff algorithms, the establishment of the MoM cutoff assumes that the assay used to measure the biomarker is stable over time. In the same way a laboratory will monitor the performance of an

analyte cutoff value and make changes to the cutoff values when a significant shift is observed, laboratories using MoMs will monitor the population median and when a significant shift is observed will adjust the median value used for calculating a patient's MoM. However, unlike when using the analyte value as the screening variable which requires changing the cutoff value when a shift in assay performance is observed, the MoM cutoff value can stay constant over time; since, the median value for calculating the patient's MoM is adjusted, instead of the MoM cutoff value.

- E. Based on the information in this guideline and other resources, it is recommended that NBS programs have a general standard operating procedure (SOP) that reflects the laboratory's and regulatory requirements for establishing and for performing ongoing evaluation of method cutoffs. This SOP can then be referred to and cited in situations where a cutoff must be defended or changed.

3.5 Collaborative Laboratory Integrated Reports

Collaborative programs such as Collaborative Laboratory Integrated Reports (CLIR)⁽²¹⁾ can be used to provide additional risk assessment for a program's chosen cutoff values. CLIR was developed by the Biochemical Genetics Laboratory at the Mayo Clinic and uses a scoring system based on condition-specific degree of overlap between covariate-adjusted reference and disease ranges. Scores are integrated for all informative markers and calculated ratios by post-analytical interpretive tools based on multivariate pattern recognition software and provides an alternative to absolute cutoff values with a single threshold of likelihood of disease.⁽²²⁻²⁶⁾ Adjustments are calculated by regression models able to factor in two continuous covariates (typically, age at collection and birth weight), one categorical covariate (sex) and harmonization of data by location. The computational approach compares each marker to its reference range adjusted for the birth weight and age at collection (beginning in hours) of the patient, and then to calculate a score based on the degree that each result penetrates the disease range above or below the corresponding reference interval. The software relies on large amounts of data (millions of data points). Current data sets are adequate from one hour up to one year of age and a birth weight of >500 g.

Newborn screening programs that enter data may use CLIR in a variety of ways:

- A. **CLIR can be used to prospectively analyze screening data.** High throughput functionalities (tool runners) are available in CLIR to temporarily upload data (data are not retained when using post analytical tools) and return a report identifying patient specimens at variable levels of risk (possible, likely, very likely) for each of the screened conditions, allowing programs to focus their attention on fewer cases, allowing greater efficiency.
- B. **CLIR can be used to identify analytes and ratios that are informative for a disorder.** CLIR includes tools that identify analytes and ratios where the separation between reference and disease range is clinically significant, meaning a degree of overlap less than 50%. This tool allows the laboratory to review analyte ratios that separate normal and affected specimens.

- C. **CLIR can assist in reducing false positive results.** Tools are also available to provide a differential diagnosis between target conditions and particularly between true and potential false positive cases (for example, carriers manifesting a milder biochemical phenotype).
- D. **CLIR allows for laboratory comparisons between participating programs.** The interpretative tools for each condition plot the values of scores for the cases in three groups: the user's site, a choice of aggregator (e.g. the United States) and the entire project worldwide. This feature allows the laboratory to compare a particular specimen to other cases and allows visual comparison of the cases from one laboratory to others.
- E. **CLIR can provide indications for differential diagnosis.** CLIR tools allow laboratories and follow-up centers to compare the extent to which a particular profile matches any of a number of conditions characterized by particular analyte elevations.

CLIR offers users the following functionalities:

- A. **Covariate adjustment.** Analytes and ratios are adjusted for demographic variables such as birth weight and age at specimen collection. In addition, sex differences are assessed and, if needed, a further adjustment is applied.
- B. **Location harmonization.** Allows for direct comparison of values or cutoffs across contributing laboratories.
- C. **Global contribution of diagnosed case data.** Results in a compilation of positive cases to help determine the distribution of markers in the affected population.
- D. **Large database of normal profiles.** The database of normal profiles allows the incorporation of new tools and/or new analyte combinations so the normal distribution can be evaluated on existing data.

Users should also be aware of the following considerations in using CLIR:

- A. **Variability in case definitions.** The tool depends on data entered by individual users, and case definitions may vary across state programs and clinicians. This should be considered whenever states are sharing results or specimens in evaluation of risk determination.
- B. **Need to customize algorithm for accurate risk analysis.** Not all states screen for the same set of analytes; therefore users must customize the algorithm within CLIR for the disorder in question in order to obtain an accurate risk determination.
- C. **Access is conditional based on contribution of data.**
- D. **CLIR analysis is not available for all disorders.**

3.6 Summary of Methods

The following table summarizes functionalities and considerations for each method described in this document.

Method	Use for	Functionalities	Considerations
<p style="text-align: center;">Fixed cutoff</p>	<p>Assays that directly measure biomarker level</p>	<p>Simplest to establish and apply; simplest for reporting purposes.</p> <p>Fixed cutoffs work best for markers with little birthweight or age dependencies and minimal lot-to-lot or seasonal variability.</p>	<p>Requires regular monitoring to determine if cutoff adjustment or instrument adjustment is needed.</p> <p>Difficult to compare directly with other screening laboratories due to method or instrument differences</p>
<p style="text-align: center;">Floating cutoff (such as percent of daily mean or median, or top 10% of the specimens)</p>	<p>Assays that show daily variability, or assays that do not directly measure biomarker level (e.g., assay depends on binding of enzyme to substrate)</p>	<p>Can be used alone or in conjunction with a fixed cutoff.</p> <p>Produces relatively consistent screen-positive rates.</p> <p>Less prone to fluctuations; less frequent need to adjust cutoff based on assay variability and population mean/median.</p>	<p>Consider removing results from outlier specimens within batch to avoid skewing floating cutoff.</p> <p>Consider number of specimens within batch when using this method.</p> <p>Complicates STAT specimen analysis, as generally the specimens are treated as a batch.</p>

Method	Use for	Functionalities	Considerations
<p>Multiple of the Median (MoM)</p>	<p>Assays that are stable over time.</p>	<p>Functions as fixed cutoff expressed as a multiple of the normal population median.</p> <p>Flexible - can be used for any assay, (e.g., biomarker level, enzyme activity, Cq, etc.)</p> <p>MoM lends itself to inter-laboratory cutoff comparisons.</p> <p>Provides simple direct way of monitoring assay performance by monitoring the Median of the MoMs over time (which by definition should be 1).</p>	<p>Allows laboratories to use a constant multiple of the median cutoff value to analyze results; variations in measured median are from analytical or method variation, so population median is changed to effectively work with a fixed cutoff.</p> <p>Requires more computations steps than other methods to evaluate data.</p>

Method	Use for	Functionalities	Considerations
<p style="text-align: center;">Collaborative Laboratory Integrated Reports (CLIR)</p>	<p>Currently used for conditions screened for by MS/MS and LSD screening. Potentially has broad applicability and may be applied to more screening methods.</p>	<p>Easily compare data across varied NBS programs and data is normalized for easier interpretation and comparisons.</p> <p>Data from contributing programs allow for multi-variate profiling of marker concentrations to determine normal ranges as function of birthweight, age and sex.</p> <p>Tool uses statistical algorithm to identify infants at risk for the condition.</p>	<p>Need backup plan in case of database outage/network outages.</p> <p>Curation of data sets</p> <p>May need to customize algorithm and tools for different marker sets.</p> <p>Program may need to take responsibility to archive results.</p> <p>Integration of CLIR reports with NBS program reporting needs to be considered.</p> <p>Access is conditional.</p>

4. Disorder Specific Considerations

4.1 Endocrine Disorders (Congenital Hypothyroidism, Congenital Adrenal Hyperplasia)

Screening for endocrine disorders is challenging as there are significant dependencies affecting the marker concentrations for these disorders including birthweight, gestational age, and age at time of specimen collection. Consider these variables when setting cutoffs.

- A. Box plots may be useful if the cutoffs need to be tiered by birth weight or age at time of specimen collection (e.g., box plots created for weight ranges can show birth weights where values change significantly).
- B. 17-OHP cutoffs are typically tiered by gestational age or birth weight.⁽²⁴⁾
- C. Thyroid stimulating hormone (TSH) cutoffs are typically tiered by the age of the newborn at the time of specimen collection. This consideration in addition to a Low-Birth-Weight Protocol helps to identify delayed cases of congenital hypothyroidism in premature infants.⁽²⁸⁾

Since endocrine disorders have been on state NBS panels for a long time, it is beneficial to consult the literature and work with other programs to determine how they are setting cutoffs for these disorders.

4.2 Cystic Fibrosis

- A.** Screening for cystic fibrosis is primarily performed with immunoreactive trypsinogen (IRT) and second-tier DNA analysis for mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. DNA analysis is performed on specimens with an elevated IRT. NBS programs use these analyses in combination with testing algorithms such as IRT/IRT, IRT/IRT/DNA, or IRT/DNA. IRT is the marker used in the primary screen and because of the high false positive rate with this marker, second-tier testing is recommended.
- B.** The IRT cutoff can be either floating or fixed. IRT can be subject to seasonal variations and lot-to-lot variability of the reagents, and some laboratories have used floating cutoffs to account for this variability.
- C.** When a floating cutoff is used, laboratories have set this cutoff at a percentage between the 95th and 99th percentiles to identify the daily batch of specimens with high IRT values that require second-tier testing. Use at least 100 specimens for an appropriate percentile calculation of a floating cutoff.
- D.** A fixed cutoff can be set that corresponds to the 95-99th percentile if a floating cutoff is not preferred. A fixed cutoff can also be added as a failsafe, if the floating cutoff is higher than expected in that particular batch of specimens.⁽¹⁸⁾
- E.** A borderline cutoff can be set for ultra-high IRT values where no mutations are identified during DNA analysis. A borderline result initiates a repeat specimen collection for the infant. Two elevated IRT results from two different specimens from the same newborn indicate further diagnostic testing is required.
- F.** Neonatal intestinal obstruction due to meconium ileus is one of the signs of cystic fibrosis that occurs in approximately 20 percent of affected newborns. This situation can result in an IRT value within the normal range. In these instances, a sweat chloride test, which is typically reserved for diagnosis following a positive screen, is recommended despite an IRT value within the normal range.⁽¹⁸⁾
- G.** Blood-level IRT concentration is fairly stable until about 2 weeks of age when it begins to decrease. Age related cutoffs may be preferred for infants greater than 7 days of age.
- H.** IRT concentrations in premature, low birth weight, or sick infants may be high due to complications other than cystic fibrosis. Additionally, African American infants have higher mean IRT concentrations compared to white and non-Hispanic infants, which often results in more false positives in these populations.⁽¹⁸⁾

4.3 Hemoglobinopathies

Most hemoglobin disorders detected through NBS (e.g., sickle cell and sickle cell C diseases) are detected qualitatively and do not require cutoffs. Alpha thalassemia, however, is detected through Hemoglobin (Hb) Bart's quantitation as a percentage which is directly proportional to disease severity. Bart's, along with acetylated and glycosylated forms, is a gamma chain tetramer formed when there is a reduction in functional alpha chains. The best way to establish cutoffs to determine alpha thalassemia severity based on Bart's percentage is to perform a phenotype/genotype correlation. When that is not possible, cutoffs may be established by comparing Bart's percentages through a specimen exchange with another program with expertise in reporting alpha thalassemia.

Isoelectric focusing, high performance liquid chromatography (HPLC), or a combination of both may be used to detect Hb Bart's as described below.

- A. Isoelectric focusing (IEF) concentrates hemoglobin proteins based on their isoelectric charge through electrophoresis using an agarose gel medium infused with carrier ampholytes. Since Hb Bart's is detected visually, the quantitative interpretation is subjective, and differentiating the severity level of the alpha thalassemia (silent carrier, alpha thalassemia trait, and Hb H disease) is challenging using IEF alone. Some labs use a densitometer to quantify the IEF percentage which improves accuracy.
- B. HPLC is the preferred method for hemoglobin percent quantitation due to its precision and accuracy.^(29,30,31) It is important to note that different HPLC systems may have different Bart's percentages depending on resolution. Also, percentage variation can occur between column resin lots. The majority of NBS programs report Hb H disease (three affected genes) which corresponds to 15 – 30 percent Hb Bart's depending on HPLC resolution.

4.4 Enzyme Deficiency Disorders

Screening tests for galactosemia, biotinidase deficiency and lysosomal storage disorders (LSDs) can use fixed or floating cutoffs. Floating cutoffs can address any seasonal changes in enzyme activity, incubation temperature fluctuations, and reagent variations that affect the calculated values.

If using a floating cutoff, be aware that the results from low birthweight, early collection, and sick infants may skew the mean or median values for the day. Excluding these sample types in calculation of the mean or median used in the floating cutoffs will lead to stabilization of the results. This is especially true for states with low specimen volumes.

Since lysosomal storage and enzyme deficiency disorders are screened by measuring the product concentration near the low end of the reference range, it is important to consider the precision of the test near the cutoff. This level of performance can be accomplished by taking multiple measurements (more than in duplicate) and averaging the results for comparison to the cutoff.

Screening for some LSDs is highly dependent on the variables of the newborns birthweight and age at the time of specimen collection. Specimen storage and transport conditions can also affect measured enzyme activities.⁽²⁴⁾ These variables should be considered in evaluation of cutoffs. The CLIR tool provides data that may be helpful in evaluating the risk of the newborn for a screened LSD.⁽²³⁾

4.5 Amino Acid, Fatty Acid Oxidation and Organic Acid disorders

- A. Amino acid and acylcarnitines are measured using MS/MS. CLSI NBS04-04, *Newborn Screening by Tandem Mass Spectrometry* establishes best practices for analyzing markers and setting cutoffs using this method; this document is recommended as a reference.⁽¹⁰⁾
- B. Consult specific method and instrument parameters to optimize performance.

- C. If a metabolic profile will be used to interpret results rather than a single marker, assess cutoffs for all markers, including any ratios used.
- D. According to the CLSI NBS04-04 recommendations, if a program analyzes specimens collected outside of the newborn period (greater than 7 days of life), validate separate age-specific cutoffs.
- E. When calculating population study data statistics, consider developing two data sets to analyze and compare the statistical parameters. Include data points from all patients for one data set and exclude data from premature, low birth weight, or neonatal intensive care unit (NICU) infants for the other data set. The use of total parenteral nutrition (TPN) and medications in this special population can elevate some amino acid and acylcarnitine levels. If a significant difference is noted in the statistical parameters, setting separate cutoffs may help reduce the risk of false negatives in the normal population and false positives in the special population.
- F. Fixed cutoffs may be assigned using means and standard deviations, or percentile data.

4.6 Severe Combined Immunodeficiency (SCID)

SCID is screened for by quantitating T-cell receptor excision circles (TRECs) in extracted DNA from dried blood spots.

- A. SCID can use a cutoff such as the MoM approach to assess the amount of TREC and control reference gene in the DBS punch.
- B. A fixed cutoff can be determined using calibration standards to quantify the number of TREC and control reference gene copies per microliter (μL) of blood.
- C. A fixed cutoff can also be based on a cycle of quantitation (Cq) value for the polymerase chain reactin rather than a TREC copy number.
- D. Screening for SCID is challenging in sick and premature infants as they can have low levels of TRECs and not have SCID. Often borderline cutoffs are used for SCID. This screening also identifies other T-cell lymphopenias.⁽²⁸⁾

4.7 X-Linked Adrenoleukodystrophy (X-ALD)

X-ALD is screened for by quantitating lysophosphatidylcholines (LPCs) in DBS using either liquid chromatography (LC) MS/MS or flow injection analysis (FIA)-MS/MS.

- A. There can be fixed cutoffs for multiple LPCs such as C20:0-LPC, C22:0-LPC, C24:0-LPC, and C26:0-LPC.
- B. When using flow injection positive ion MS/MS analysis to screen for X-ALD, there is an unknown isobaric interference in determining the concentration of C26:0-LPC. Since the deuterated internal standard does not correct for effects of this isobaric interference, instrument-to-instrument and day-to-day concentrations of C26:0-LPC will vary more than is observed with other MS/MS assays (amino acids/acylcarnitines). Consider adjusting the daily median or mean for all results to match the population mean or median, so that the individual specimen results and the daily mean for this marker are consistent with the results measured during the validation study. The relative response factor can be adjusted as needed to achieve consistency in daily or instrument-to-instrument variability.
- C. Another option would be to consider a floating cutoff for this marker, but this would need to be done for each instrument, as instrument performance can vary day to day, and is common after performing routine maintenance.
- D. Additionally, fixed cutoffs can be established for ratios between the LPC markers, for example, C24:0 LPC over C20:0 LPC or C26:0 LPC over C22:0 LPC as well as other LPC ratios may be useful.

5. Monitoring and Ongoing Evaluation of Risk Assessment

- A.** Monitor cutoffs and/or other risk assessment methods closely for the first year by evaluating the test sensitivity and specificity when possible. Use the number of borderline and positive results, false positive rate and positive predictive value in the evaluation based on the follow-up outcomes.
- B.** Determine if the cutoffs/risk assessment identifying the expected number of positive cases for your state's population and if the false positive rate is reasonable or high based on the expected and detected incidence. Typically, once conducting routine screening, these adjustments will be made based on feedback from the follow-up program (e.g., if primarily mild phenotypes are detected at the present cutoff, etc.).
- C.** Meet with specialists, if possible, to evaluate population and screening data. Factors to evaluate include review of new true positive cases, false positive cases, false negative cases and any new peer-reviewed literature. Programs may also consider comparing their cutoff values and risk assessments to CLIR or other NBS programs.
- D.** Consider re-evaluating cutoffs/risk assessment following kit changes, changes in equipment or test modification or when new clinical information is received (e.g., from cases or knowledge of natural history of the disease).
- E.** Check the cutoffs/risk assessment method every six months when routine screening begins.⁽³²⁾
- F.** In the event that the screening laboratory has been informed of a false negative result, it is important to fully investigate the finding. To make this determination:
 1. Obtain the diagnostic results and consult with the specialist to understand the exact diagnosis, if the newborn is truly affected by the disorder being screened for and if it is a disorder that should be detected by the screen.
 2. Repeat the screen to see if the original result is obtained. If the second screen result is within experimental error of the original result (as defined by the NBS program), and the result is far removed from the current cutoff, check to see if any other extenuating circumstances may have affected the measured result such as specimen quality, birthweight, delays in specimen collection, transport or feeding issues, or if the specimen is from a different infant.
 3. If all other factors have been ruled out, consider adjusting the cutoff/risk assessment tool or testing the algorithm to reduce screen-negative results while not introducing an unacceptable number of false positive results.
 4. Consider if there are other markers or ratios of markers that can be used to detect the specimen as screen positive.
 5. Consider utilization of tools such as CLIR and peer information from neighboring states to evaluate the specimen and determine if it would have been flagged as at risk for the condition.
 6. Send the specimen to another screening laboratory to ensure consistency of results and for cutoff comparisons. Determine if other laboratories would have detected this specimen as a screen positive.

6. References

- ⁽¹⁾ Association of Public Health Laboratories (APHL), APHL Newborn Screening and Genetics in Public Health Committee, and APHL Newborn Screening Technical Assistance and Evaluation Program (NewSTEPS) Steering Committee. Making Sense of Newborn Screening Cutoff Values [Internet]. 2017 Feb [cited 2018 Jul 25]. Available from: <http://www.aphlblog.org/2017/02/making-sense-newborn-screening-cut-off-values/>
- ⁽²⁾ Advisory Committee on Heritable Disorders in Newborns and Children. Recommended Uniform Screening Panel [Internet]. 2016 Nov [cited 2018 Jul 25];1-2. Available from: <https://www.hrsa.gov/sites/default/files/hrsa/advisory-committees/heritable-disorders/rusp/uniformscreeningpanel.pdf>
- ⁽³⁾ Puckett RL, Lorey F, Rinaldo P, Lipson MH, Matern D, Sowa ME et al. Maple syrup urine disease: Further evidence that newborn screening may fail to identify variant forms. *Molecular Genetics and Metabolism*. 2010 Jun; 100(2):136-142.
- ⁽⁴⁾ Dunn CT, Skrypek MM, Powers AL, Laguna TA. The need for vigilance: the case of a false negative newborn screen for cystic fibrosis. *Pediatrics* [Epub]. 2011 Jul;128(2):446-449. doi: 10.1542/peds.2010-0286.
- ⁽⁵⁾ Sahai I, Bailey JC, Eaton RB, Zytkevich T, Harris DJ. A near-miss: very long chain acyl-CoA dehydrogenase deficiency with normal primary markers in the initial well-timed newborn screening specimen. *J Pediatr*[Epub].2010 Nov 12;158(1):172; author reply 172-3. doi: 10.1016/j.jpeds.2010.09.026.
- ⁽⁶⁾ Crushell E, Chukwu J, Mayne P, Blatny J, Treacy EP. Negative screening tests in classical galactosaemia caused by S135L homozygosity. *J Inherit Metab Dis*[Epub]. 2009 May 8;32(3):412-5. doi: 10.1007/s10545-009-1081-4.
- ⁽⁷⁾ Wald, NJ. (2008). Guidance on terminology. *Journal of Medical Screening*. 2008; 15(1), 50–50.
- ⁽⁸⁾ APHL Newborn Screening Quality Assurance/Quality Control Subcommittee. Newborn Screening Analyte Interference List [Internet]. 2014 Oct [cited 2018 Jul 25]. Available from: https://www.aphl.org/programs/newborn_screening/Pages/NBS%20Interference%20List.aspx
- ⁽⁹⁾ American Academy of Pediatrics. Update of newborn screening and therapy for congenital hypothyroidism. *Pediatric*. 2006; 11(6): 2290-2303.
- ⁽¹⁰⁾ NBS04 – Newborn Screening by Tandem Mass Spectrometry, 2nd Edition. Clinical and Laboratory Standards Institute.
- ⁽¹¹⁾ Kwon JM, Matern D, Kurtzberg J, Wrabetz L, Gelb MH, Wenger DA, et al. Consensus guidelines for newborn screening, diagnosis and treatment of infantile Krabbe disease. *Orphanet J Rare Dis*. 2018; 13:30. doi: 10.1186/s13023-018-0766-x.
- ⁽¹²⁾ Turgeon CT, Magera MJ, Cuthbert CD, Loken PR, Gavrilov DK, Tortorelli S, et al. Determination of total homocysteine, methylmalonic acid, and 2-methylcitric acid in dried blood spots by tandem mass spectrometry. *Clin Chem*. 2010 Nov; 56(11):1686-95. doi: 10.1373/clinchem.2010.148957.

- ⁽¹³⁾ Oglesbee D, Sanders KA, Lacey JM, Magera MJ, Casetta B, Strauss KA, et al. Second-tier test for quantification of alloisoleucine and branched-chain amino acids in dried blood spots to improve newborn screening for maple syrup urine disease (MSUD). *Clin Chem*. 2008; 54: 542-9. doi: 10.1373/clinchem.2007.098434.
- ⁽¹⁴⁾ Matern D, Tortorelli S, Oglesbee D, Gavrillov D, Rinaldo P. Reduction of the false positive rate in newborn screening by implementation of MS/MS-based second-tier tests: the Mayo Clinic experience (2004-2007). *J Inher Metab Dis*. 2007 Aug;30(4):585-92. doi: 10.1007/s10545-007-0691-y.
- ⁽¹⁵⁾ Minutti CZ, Lacey JM, Magera MJ, Hahn SH, McCann M, Schulze A, et al. Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab*. 2004 Aug;89(8):3687-93. doi: 10.1210/jc.2003-032235.
- ⁽¹⁶⁾ NSQAP: Newborn Screening Quality Assurance Program [Internet]. 31 Aug 2017 [cited 2018 Jul 25]; Available from: <https://www.cdc.gov/labstandards/nsqap.html>
- ⁽¹⁷⁾ Laboratory Performance Database (R4S) [Internet]. [cited 2018 Jul 25]; Available from: <https://www.nbstrn.org/research-tools/lab-performance-database>
- ⁽¹⁸⁾ NBS05-Newborn Screening for Cystic Fibrosis, Draft Edition, Clinical and Laboratory Standards Institute.
- ⁽¹⁹⁾ NBS07-Newborn blood spot screening for Pompe disease by lysosomal acid α -glucosidase activity assays, 1st Edition. Clinical and Laboratory Standards Institute.
- ⁽²⁰⁾ NBS06-Newborn Blood Spot Screening for Severe Combined Immunodeficiency by Measurement of T-cell Receptor Excision Circles, 1st Edition. Clinical and Laboratory Standards Institute.
- ⁽²¹⁾ Collaborative Laboratory Integrated Reports (CLIR) [Internet]. [cited 2018 Jul 25]; Available from: <https://clir.mayo.edu/Home/About>
- ⁽²²⁾ Mørkrid L, Rowe AD, Elgstoen KBP, Olesen JH, Ruijter G, Hall PL, Tortorelli S et al. Continuous age- and gender-adjusted reference intervals of urinary markers for cerebral creatine deficiency syndromes: A novel approach to the definition of reference intervals. *Clin. Chem*. 2015; 61: 760–768.
- ⁽²³⁾ Marquardt G, Currier R; McHugh DMS, Gavrillov D., Magera MJ, Matern D et al. Enhanced interpretation of newborn screening results without analyte cutoff values. *Genet. Med*. 2012; 14: 648–654.
- ⁽²⁴⁾ Hall PL, Marquardt G, McHugh DM, Currier RJ, Tang H, Stoway SD et al. Post-analytical tools improve performance of newborn screening by tandem mass spectrometry. *Genet. Med*. 2014; 16: 889–895.
- ⁽²⁵⁾ Minter Baerg MM, Stoway SD, Hart J, Mott L, Peck DS, Nett SL et al. Precision newborn screening for lysosomal disorders. *Genet Med* [Epub ahead of print]. 2017 Nov 9. doi: 10.1038/gim.2017.194.

⁽²⁶⁾ Tortorelli S, Eckerman JS, Orsini JJ, Stevens C, Hart J, Hall PL et al. Moonlighting newborn screening markers: the incidental discovery of a second-tier test for Pompe disease Genet Med [Epub]. 2017 Nov 2. doi: 10.1038/gim.2017.190.

⁽²⁷⁾ van der Kamp HJ, Oudshoorn, CGM, Elvers BH, van Baarke M, Otten BJ, Wit JM, Verkerk, PH. Cutoff levels of 17 α —hydroxyprogesterone in neonatal screening for congenital adrenal hyperplasia should be based on gestational age rather than on birth weight. J. Clin. Endocrinol. Meta. 2005; 90: 3904-3907.

⁽²⁸⁾ NBS03-Newborn Screening for Preterm, Low birth Weight, and Sick Newborns, 1st Edition. Clinical and Laboratory Standards Institute.

⁽²⁹⁾ APHL Hemoglobinopathy Laboratory Workgroup. Hemoglobinopathies: Current Practices for Screening, Confirmation, and Follow-up [Internet]. 2015 Dec [cited 2018 Jul 25]. Available from: https://www.aphl.org/aboutAPHL/publications/Documents/NBS_HemoglobinopathyTesting_122015.pdf.

⁽³⁰⁾ APHL Hemoglobinopathy Laboratory Workgroup. Hemoglobinopathy Screening Methods: Isoelectric Focusing (IEF), High Performance Liquid Chromatography (HPLC) webinar [Internet]. 2017 Apr [cited 2018 Jul 25]. Available from: <https://vimeo.com/163830104>.

⁽³¹⁾ APHL Hemoglobinopathy Laboratory Workgroup. Alpha Thalassemia: Newborn Screening in the United States webinar [Internet]. 2017 Jun [cited 2018 Jul 25]. Available from: <https://vimeo.com/224330050>.

⁽³²⁾ Department of Health and Human Services (HHS) Centers for Medicare and Medicaid Services (CMS). Medicare State Operations Manual Provider Certification: § 493.1281 Standard: Comparison of test results [Internet]. 2004 Feb [cited 2018 Jul 25]. Available from: <https://www.cms.gov/Regulations-and-Guidance/Guidance/Transmittals/downloads/r32som.pdf>.