SEEKER™
Clinical Study Report

baebies

Table of Contents

1. STUDY SYNOPSIS

Sponsor:
Baebies, Inc. 615 Davis Drive, Suite 800 Durham, NC 27560
Study Title:
Evaluation of the SEEKER™ System for Quantitation of Lysosomal Enzyme Activity for IDUA, GAA, GBA, and GLA in Newborn Screening.
Indications/Condition:
Low activity for the tested enzymes may be indicative of the following disease states, respectively: Mucopolysaccharidosis Type I (MPS I), Pompe, Gaucher, and Fabry.
Investigation Site:
Missouri State Public Health Laboratory (MSPHL)
Principal Investigator:
Patrick Hopkins Chief, Newborn Screening Unit Missouri State Public Health Laboratory
Phases/Study Periods:
Pre-pilot: Test period used to establish baseline cutoff values. (Approximately 2 months or 13,000 de-identified samples) Pilot: January 15, 2013-August 26, 2013 Pivotal : August 27, 2013-January 14, 2015
Objective:
The purpose of this prospective clinical study was to demonstrate the clinical performance of the SEEKER™ System to quantitatively measure α-L-iduronidase [IDUA], α-D-glucosidase [GAA], β-glucocerebrosidase [GBA] and α-D-galactosidase A [GLA] on newborn dried blood spots. The primary end points of the study are the following: • Identification of true positives • False negatives • False positive rate • Retest rate

Design:

Endpoint Classification: Safety and Efficacy Study
Intervention Model: Single Group Assignment
Masking: None
Primary Purpose: Screening
Time Perspective: Prospective

Methodology:

As a part of Missouri's ongoing newborn screening program, dried blood spots (DBS) were prospectively collected from hospitals, birthing centers, and/or clinics throughout the state of Missouri and shipped to the central state newborn screening laboratory (Missouri State Public Health Laboratory, MSPHL). These same DBS were tested for reduced enzymatic activity associated with certain lysosomal storage disorders (LSDs) using SEEKER™. Those presumed affected newborns were referred for confirmatory testing. The results of the confirmatory tests were provided by the referral center to MSPHL.

Intervention:

Newborns with low enzymatic activity that were considered to be at significant risk received confirmatory testing and diagnosis by an independent laboratory(s). The results of confirmatory testing were reported to the MSPHL.

Number of Subjects:

Total: 154,412 newborns born during the study
Pilot: 48,813
Pivotal: 105,599

Subject Inclusion/Exclusion Criteria:

Inclusion Criteria:
-Newborns receiving newborn screening in the state of Missouri.
Exclusion Criteria:
-No exclusion criteria of newborns (subjects).

Sample Exclusion (for Analysis):

Certain test samples from newborns that did not meet minimum criteria for quality were excluded from the analysis:

1) Samples taken at <24 hours
2) Samples with no recorded age at collection
3) Samples that, despite re-testing, produced no valid test results.
4) Samples that were designated poor quality upon receipt by MSPHL

If all samples from a newborn were excluded, then the newborn was excluded from the analysis. The final number of newborns that were included in analysis was 153,697.

Duration of Treatment:

Not applicable

Follow Up:

Subjects identified by MSPHL as presumed affected were referred by MSPHL to a treating physician for confirmatory testing. If the newborn was seen by a treating physician, the results of confirmatory testing were provided back to MSPHL. Based on the results of confirmatory testing, the newborns were either 'true positive' or 'false positive'.

For those subjects deemed as 'screen negative' by MSPHL, there was no additional follow up.

In the state of Missouri, the Missouri Department of Health and Senior Services – Bureau for Genetics and Health Childhood is the central reporting function for genetic disorders. To determine if any new or previously unreported cases of LSDs occurred, i.e. 'screen negatives' that might actually have been positive and therefore 'false negatives', the MSPHL NBS lab continually followed up with the Bureau, which in turn followed up with the four contracted genetic referral centers. This follow up was utilized as the determinant for the 'false negative' rate. As of 3/30/2016, 15 months after completion of the two-year study, no false negatives have been reported.

Test Product:

The SEEKER™ System is intended for quantitative measurement of the activity of multiple lysosomal enzymes from newborn dried blood spot samples. Reduced activity of these enzymes may be indicative of a lysosomal storage disorder. The enzymes measured using the SEEKER™ Reagent Kit and their associated lysosomal storage disorder are indicated in the table below.

Enzyme (abbreviation)	*Disorder*
α-L-iduronidase (IDUA)	MPS I
α-D-glucosidase (GAA)	Pompe
β-glucocerebrosidase (GBA)	Gaucher
α-D-galactosidase A (GLA)	Fabry

Reduced activity for any of the four enzymes must be confirmed by other confirmatory diagnostic methods.

Statistical Methods:

The following statistical methods were used during the study:

- Calculation of the false positive rate percentage
- Calculation of the retest rate
- Calculation of the reference intervals

2. DEFINITIONS AND ABBREVIATIONS

Assay: Quantitative measurement of the activity of a single analyte (enzyme). Each test has four assays.

Borderline cutoff: Borderline cutoff is set above the high risk cutoff to account for imprecision around the high risk cutoff. Samples above the borderline are presumed normal. Samples below the borderline cutoff are typically repeated to obtain a final average result, based on which further disposition occurs.

Card: A small sheet of filter paper for collection of dried blood spots from a newborn. Each card should contain five (5) full dried blood spots from a single newborn, collected during a standard time frame (24 to 48 hours postnatal age).

Carrier: Indicates the presence of an inherited recessive allele for a genetic trait or a mutation for which there is no clinical presentation of symptoms. Carriers are, however, able to pass the allele onto their offspring, who may then manifest the disorder if they inherit the recessive allele from both parents. Carriers can exhibit lower enzymatic activity than normal population. If identified through newborn screening, carriers are considered as false positives.

Cut-offs: A lab-established quantitative enzymatic activity value below which samples are flagged for further resolution. The common practice in newborn screening programs is to have two cutoffs – a high risk cutoff and a borderline cutoff.

Dried Blood Spot (DBS): One of the five circular dried blood specimens on a newborn screen collection card.

False Negative: A newborn who was presumed normal through the newborn screening process and who later on is reported with a diagnosis of the disease.

False Positive: A newborn who was presumed affected and referred for diagnostic testing, which revealed the absence of disease.

False Positive Rate: False positive rate is calculated by dividing the total number of true positive samples by the total number of newborns screened (minus true positive samples and samples lost to follow up).

GAA: Abbreviation for α-D-glucosidase (EC 3.2.1.20). Pompe disorder is associated with deficient GAA activity.

GBA: Abbreviation for β-glucocerebrosidase (EC 3.2.1.45). Gaucher disorder is associated with deficient GBA activity.

Genotype of Unknown Significance: Presence of a genotype combination that is not currently found in the registry of mutations for a particular disorder.

GLA: Abbreviation for α -D-galactosidase A (EC 3.2.1.22). Fabry disorder is associated with deficient GLA activity.

High-risk cutoff: High risk cutoff is the clinical decision making level. Samples below the high risk cutoff are at a higher risk of having the disorder.

IDUA: Abbreviation for α -L-iduronidase (EC 3.2.1.76). Mucopolysaccharidosis Type I disorder is associated with deficient IDUA activity.

Infantile Onset: Onset of disorder symptoms within the first few months of life (infancy).

Late Onset: Onset of disorder symptoms from early childhood to adulthood.

NBS: Newborn screening.

Presumed Affected: A newborn who is identified in screening as high-risk for disease and thus requires confirmatory diagnosis.

Presumed Normal: A newborn who is identified in screening as low risk for disease.

Pseudodeficiency: Indicates the presence of a known mutation for a disorder which causes reduced or no enzymatic activity in *in vitro* testing without the presentation of clinical symptoms. If identified through newborn screening, pseudodeficient newborns are considered false positives.

Punch: A single 3.2 mm circle removed from a spot. Up to eight (8) punches can be extracted from a single dried blood spot; given five full spots, up to 40 punches can be extracted from a single screen. Each "test" requires extract from a single punch.

Sample: Can be used interchangeably with "screen" and "specimen"; denotes a single specimen card with five (5) full dried blood spots from a single newborn, collected at the same time. Multiple samples can be collected from a single newborn at different points in time. Missouri guidelines require multiple screens be collected from a newborn in a number of circumstances, including blood transfusion and/or premature birth.

Screen: Can be used interchangeably with "sample" and "specimen"; card for collection of dried blood spots from a newborn. Each card should contain five (5) full dried blood spots from a single newborn, collected at the same time.

Screen Negative: See "presumed normal".

Screen Positive: See "presumed affected".

Specimen: Can be used interchangeably with "sample" and "screen"; card for collection of dried blood spots from a newborn. Each card should contain five (5) full dried blood spots from a single newborn, collected at the same time.

Spot: One of the five circular dried blood specimens on a screen card.

Test: One set of multiplexed assays performed on blood spot extract. Each test requires extract from a single DBS punch. Multiple tests may be performed on a single screen, but each test requires a separate punch.

True Positive: A presumed affected newborn who is confirmed to have the disorder by diagnostic testing.

Unclassified Onset: Confirmed disorder was labeled unclassified if the diagnostic test center did not provide an onset while reporting back to MSPHL.

Unknown Onset: Confirmed disorder by enzymatic and confirmatory diagnosis of which the type of onset (infantile onset, late onset) is unknown.

3. OBJECTIVE

The purpose of this prospective clinical study was to demonstrate the clinical performance of the SEEKER™ System to quantitatively measure α-L-iduronidase [IDUA], α-D-glucosidase [GAA], β-glucocerebrosidase [GBA] and α-D-galactosidase A [GLA], on newborn dried blood spots. Low activity for these tested enzymes may be indicative of the following lysosomal storage disorders, respectively: Mucopolysaccharidosis Type I (MPS I), Pompe, Gaucher, and Fabry.

To meet this objective a single-site clinical study using the sponsor's SEEKER™ System was conducted in the Missouri State Public Health Laboratory (MSPHL). All newborn dried blood spot samples received at the MSPHL for mandatory routine newborn screening were tested for the 4 lysosomal storage disorder (LSD) enzymes during the study.

The end points of the study were the following:

- False positive rate
- False negative rate
- Retest rate
- Identification of true positives

4. INTRODUCTION

Lysosomal storage disorders (LSD) are a group of approximately 50 rare, inherited metabolic disorders that are caused by lysosomal dysfunction, usually as a consequence of a deficiency in a single enzyme required for the metabolism of lipids, glycoproteins or mucopolysaccharides. FDA-approved enzyme replacement therapy is now available for the following lysosomal storage disorders: Pompe, Fabry, Mucopolysaccharidosis Type I (MPS I), Mucopolysaccharidosis Type II (MPS II), Mucopolysaccharidosis Type VI (MPS VI) and Gaucher[1]. Currently, LSDs are diagnosed through a combination of enzymatic assays, chest x-rays, analysis of glycosaminoglycans, and molecular analysis only after the presentation of symptoms. Pre-symptomatic newborn screening for LSDs and subsequent early treatment prevents the onset of irreversible organ damage and death. There are no FDA approved methods for testing for LSDs.

The SEEKER™ System is used to measure lysosomal enzymes for 4 LSDs – MPS I, Pompe, Gaucher, and Fabry.

4.1. MPS I

Mucopolysaccharidosis Type I (MPS I) is an autosomal recessive disorder that is caused by a deficiency in the enzyme α-L-iduronidase (IDUA). Clinical symptoms, if not treated early, include hydrocephalus, corneal clouding, hepatosplenomegaly, and cardiomyopathy. MPS I typically leads to death by age 10 if not detected and treated prior to the onset of symptoms. For treatment, enzyme replacement therapy (ERT) or hematopoietic stem cell transplantation (HSCT) are indicated.

4.2. Pompe Disorder

Pompe disorder (glycogen storage disorder type II) is an autosomal recessive disorder. A deficiency in the enzyme acid α-glucosidase (GAA), which prevents the degradation of lysosomal glycogen, leads to Pompe. Pompe disorder presents as a continuum of disorder severity – at one end of the spectrum is classic infantile Pompe and at the other end is late-onset Pompe. Glycogen accumulation can lead to cardiomyopathy, respiratory and muscle weakness (hypotonia). For newborns diagnosed with Pompe disorder, early initiation of enzyme replacement therapy (ERT) using recombinant human GAA (rhGAA) has resulted in significantly improved outcomes.

4.3. Gaucher disorder

Gaucher disorder is an autosomal recessive disorder that results from a deficiency in the enzyme glucocerebrosidase (GBA), or β-glucosidase, in which cerebrosides accumulate. If not detected early, Gaucher disorder may result in anemia, thrombocytopenia, massive hepatosplenomegaly and bone marrow involvement. Enzyme replacement therapy and substrate reduction therapy are effective treatments.

4.4. Fabry disorder

Fabry disorder is an X-linked disorder characterized by a deficiency in α-galactosidase A (GLA), resulting in glycosphingolipid ceramide trihexoside (GL-3) accumulation. For female patients, due to random X-chromosome inactivation, enzyme activity is highly variable and for some, it could overlap with the normal range. Therefore, enzyme activity for females should be interpreted with caution, as some female carriers can have enzyme activity in the normal range despite clinical manifestations[2]. The accumulation of GL-3 with Fabry disorder may cause cardiomyopathy, hearing loss, cardiac hypertrophy, rhythm disturbances, renal insufficiency, GI involvement, and joint swelling. ERT is life-saving in Fabry and has improved outcomes[3].

5. STUDY DESIGN

5.1. Study Type

The clinical study in MSPHL was an investigational prospective study. The standard process for collection and initial testing of newborn screening samples at the Missouri State Public Health Laboratory was utilized. All screen positive specimens were referred for confirmatory testing and confirmed as affected (true positive) or unaffected (carriers, pseudodeficient, or normal confirmatory enzyme level) using other currently available diagnostic testing methods prior to treatment. All screen negative specimens were presumed to be normal with no further action. Monitoring for false negatives was accomplished through MSPHL's network of metabolic centers.

5.2. Study Blinding

This study was not blinded. Test results were evaluated using a screening algorithm developed by the MSPHL. This algorithm required decisions to retest based on initial test results and decisions to refer for confirmatory testing based on a set of Risk Assessment criteria.

5.3. Subject Selection

There were no subject inclusion or exclusion criteria for this study since the intended purpose of this test is population newborn screening. All newborn dried blood spots received at the MSPHL for routine newborn screening in the study period were screened for the 4 LSD enzymes. No newborns were omitted from this study, unless their parents had opted-out of routine newborn screening on religious grounds as per Missouri state law[4].

6. STUDY PROTOCOL

The product was used in accordance with the SEEKER™ 4-Plex Assay Instructions for Use.

6.1. Study Period

The study was conducted using specimens received at MSPHL between January 11, 2013 and January 14, 2015. The testing of these specimens occurred from January 15, 2013 and January 14, 2015, inclusive.

The study period was retrospectively separated into "pilot" and "pivotal" phases. The pilot phase included newborns born on or before 8/26/2013. The pivotal phase included newborns born on or after 8/27/2013.The pilot phase began on January 15, 2013. Pre-pilot testing was completed at MSPHL using ~13000 de-identified newborn DBS and 29 known affected samples. The results of the pre-pilot testing were used to set initial cutoff values for the pilot study. During the pilot phase, several adjustments were made to the cutoff values as the laboratory acquired additional knowledge about the assay performance for different subgroups (for example, cutoffs based on sample age at collection) as well as confirmatory diagnosis on screen positive newborns that were referred. Five minor and one major change to the SEEKER™ System also occurred during the pilot period. These device changes are listed in Appendix C.

The pivotal phase began with newborns born on or after 8/27/2013. At this time, MSPHL began using the final version of the SEEKER™ Cartridge and software, which were used through the conclusion of the study period.

6.2. Specimen Handling

6.2.1. Specimen Collection and Shipping

MSPHL provided blank newborn screening dried blood cards to all birthing facilities in the state of Missouri. The kits were requested by the hospital (or other birthing facility) and contained blank dried blood spot cards and supplies needed to mail the specimens to MSPHL. MSPHL used Ahlstrom 226 filter paper as the matrix for dried blood spot specimens (K062932). Specimen collection instructions were provided on the specimen collection form. Each dried blood spot card had 5 dried blood spots.

The timing of specimen collection followed Missouri's guidelines for collection of newborn screening samples which is summarized below. These guidelines are consistent with approved CLSI guidelines for dried blood spot collection[5] and screening of preterm, low birth weight and sick newborns[6].

1. A specimen collected between 24-48 hours of life was considered optimum for newborn screening.
2. If the initial specimen was collected prior to 24 hours of life, then a second (repeat) specimen/screen was required within 14 days of life.
3. Initial specimens from ill or premature newborns were collected prior to blood transfusion or between 24-48 hours of life. All ill or premature newborns required a repeat specimen/screen collected between 7-14 days of life.
4. A newborn specimen (repeat specimen) collected at 28 days of life was recommended for all infants who were less than or equal to 34 weeks gestational age or less than 2,000 grams at birth.

6.2.2. Specimen Accessioning

All specimens received at the laboratory were examined for specimen acceptability according to MSPHL guidelines, and identified as initial, repeat or poor quality specimens. Poor quality specimens may not have enough blood to perform all the testing, may have been collected improperly, and or may have been delayed in the mail.

6.3. Analytical Procedure

Table 1 describes the high level steps involved in the analytical procedure used during the study.

Table 1: Enzymatic assay analysis procedure

Step #	Step	Description
1	Analyzer & software Preparation	Power up analyzer and thermal system, prepare instrument, start software.
2	Reagent & Cartridge Preparation	Thaw reagents, vortex and spin. Load filler fluid into cartridges.
3	Sample Setup	Obtain samples. Extract DBS punches in extraction buffer. Transfer calibrators to extraction plate.
4	Load and check Cartridge	Insert cartridge into analyzer and engage with instrument. Check connectivity of cartridge to analyzer
5	Enter run info	Enter run information into Spot Logic®
6	Load samples/reagents	Load stop buffer, reagents and samples into cartridge.
7	Start & Run Assay	Execute assay protocol. Perform calculations.
8	Unload Cartridge	Remove cartridge from analyzer and dispose.

6.4. Run and Individual Sample Acceptance

6.4.1. Calibration failure

No results are generated if calibration routine fails on the cartridge. In this case, all the samples from that cartridge are repeated using a single re-punch and retest from the same dried blood spot as the initial run.

6.4.2. QC failure

MSPHL included 2 levels of Quality Control material (QCL and QCM) in duplicate per run. If each of the duplicate QCMs or QCLs were flagged to be outside of ±3SD, the newborn sample results were reviewed and samples from that cartridge repeated based on their activity level.

6.4.3. Individual Sample Acceptance

An invalid data points triggers a single re-punch and retest from the same dried blood spot. Invalid data points fall into the following categories:

- Test result reported as “n/a”
- Test result below a lower threshold : IDUA < -2µmol/L/hr (or) GAA < 0µmol/L/hr (or) GBA < 0µmol/L/hr (or) GLA < 1µmol/L/hr

- Test result above an upper threshold (set at 3SD[1] above normal median)

Due to the multiplexed nature of the system, a retest generates results for all enzymes even if a single enzyme triggered the retesting.

6.5. Test Result Interpretation

6.5.1. Cutoffs

The common practice in newborn screening programs is to develop and apply two cutoffs in the screening process – a high risk cutoff and a borderline cutoff[7]. This approach is especially useful for LSDs, where newborns at risk of having a disorder have very low enzymatic activity. The high risk cutoff which is the primary clinical decision making level therefore lies in the lower end of the measurement range of the system. This naturally results in higher imprecision around the high risk cutoff. To mitigate this risk a borderline cutoff is set above the high risk cutoff to allow for imprecision around the clinical decision making level. Samples below the borderline cutoff have additional testing (2 minimum for a total of 3 tests per sample) performed on the same specimen.

Samples with enzymatic activity below the borderline cutoff are additionally tested in duplicate using 2 additional punches, and the average of all 3 values is used for further disposition. This effectively results in a standard error of the mean equal to standard deviation divided by $\sqrt{3}$. Replication of tests improves the confidence in the enzymatic activity values near the cutoff values and effectively reduces the imprecision.

6.5.2. Result Interpretation

1. Samples with results above the borderline cutoff for all enzymes were considered low risk and were presumed normal. No additional action was taken.
2. Samples with at least one enzyme below borderline cutoff were retested in duplicate, with two new punches from the same dried blood spot sample using different blood spot circles on the card as the initial test. If the variability of the replicates was high, MSPHL standard protocol was to perform additional testing of the sample, from additional dried blood spots on the card (to evaluate spot-to-spot variability). The average of all test values, excluding outliers, was used to make the final disposition on the newborn.
3. If the average of all samples, excluding outliers, was above the high risk cutoff, the sample was presumed normal.
4. If the average of all tests, excluding outliers, was below the high risk cutoff, a Risk Assessment was performed. This involved reviewing prior LSD results from the newborn (if available), other enzyme results from the multiplex assay, other newborn screening results, infant's gestational age, age at specimen collection, and health status. The Risk Assessment is detailed in 6.5.3. If the newborn was considered high risk after this risk assessment process, it was referred to 1 of 4 contracted genetic referral centers for further evaluation,

[1] The standard deviation is calculated as a geometric standard deviation since the distribution is lognormal.

confirmatory testing and diagnosis. Genetic center referrals were based on geographic regional coverage and the location of the newborns' residence.

6.5.3. Risk Assessment

To make a determination regarding referral and confirmatory diagnosis, MSPHL applied a Risk Assessment process to the evaluation of final results.

MSPHL applies a number of criteria when assigning risk to samples with test results below the high risk cutoff. Some criteria are reflected in the clinical information collected on the screening card by the clinicians who acquire the blood sample i.e. 'sick', 'transfused', 'premature, i.e. gestational age', age at collection, example of a DBS screening card). This process also includes transit time which can be determined from the information on the card. Table 2 below provides details on the Risk Assessment criteria and a description of their use during the study.

Table 2: Risk Assessment Details

Risk Assessment Criteria		Description
Additional Samples		Other DBS screen(s) (other DBS specimen cards) from the same newborn that were tested and determined to be presumed normal for LSDs.
Newborn Status	Transfusion status	Samples from transfused babies can have unreliable results and the lab obtains at least one repeat screen; so they did not refer transfused newborns based on the first screen.
	Gestational Age	At least one repeat screen is required from babies that are born significantly premature (<= 34 weeks gestational age). In cases where the birth was premature and another sample was expected, the disposition may be postponed until the repeat sample is received.
	Other Altered Health Status	If the baby is indicated as "Sick" (represented by a category of that designation on the screening card) at the time of collection, the lab requires another screen from the newborn.
	Age at Sample Collection	Initially MSPHL used the newborn age at collection as criterion for Risk Assessment. Later during the study MSPHL instituted age specific cutoffs for samples collected between 7-13 days of life and greater than 14 days of life.
Suspected Poor Quality	Transit Time / Birthplace	MSPHL utilizes a courier system that transports > 95% of newborn samples from the birthing centers to the laboratory. Samples are mailed to the lab in the cases of home births or deliveries at birthing centers that do not participate in the courier program. They have found that in cases where the sample spent significant time in transit, the activity values may be reduced. For samples accessioned 15+ days after collection, the lab will classify the sample as poor quality and request a new screen.
	Other LSD results	MSPHL considered samples where an LSD enzyme is below the high risk cutoff and at least one other enzyme is below the borderline cutoff to be a potentially poor quality sample. Given the population distribution of all four assays and assuming that the activity values of the four assays are expected to be biochemically independent, the estimated likelihood of one assay below high risk and another below borderline is between 1 in 125,000 and 1 in 1,400,000.

6.5.4. Clinical Diagnosis

True clinical status of the referred newborns was determined by the methods summarized in Table 3. These follow up tests are the current standard of care for lysosomal storage disease diagnosis.

Table 3: Methods for determining true clinical status of referred newborns

Disorder (Enzyme)	Tests	Possible Diagnosis	
		Affected (True Positive)	**Normal (False Positive)**
MPS I (IDUA)	IDUA assay on leukocytes Mutation analysis	• Attenuated • Severe • Genotype of unknown significance	• Normal • Carrier • Pseudodeficiency
Pompe (GAA)	GAA assay on leukocytes Urine HEX4 assay Creatine kinase	• Classical Infantile Onset • Nonclassical Infantile Onset • Late Onset • Unknown Onset • Genotype of Unknown Significance	• Normal • Carrier • Pseudodeficiency
Gaucher (GBA)	GBA assay on leukocytes Mutation analysis.	• Neuronopathic • Non-neuronopathic • Unknown Onset • Genotype of Unknown Significance	• Normal • Carrier
Fabry (GLA)	Male: GLA assay in leukocytes Mutation analysis Female: Mutation analysis	• Classical • Late Onset • Genotype of Unknown Significance	• Normal • Pseudodeficiency

The entire newborn screening workflow used during the clinical study is summarized in Figure 1 below.

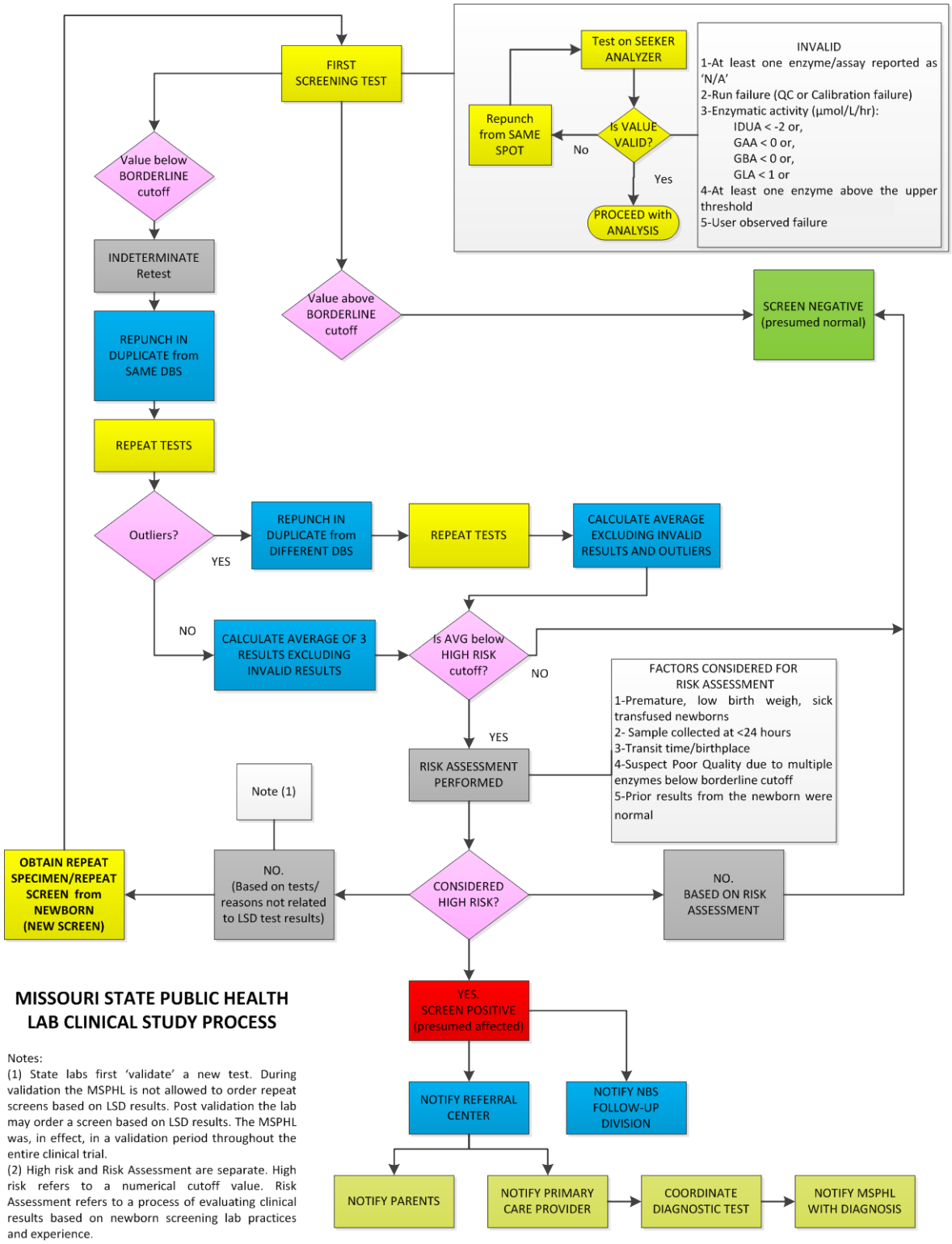


Figure 1 - MSPHL LSD Screening Workflow

7. MSPHL CUTOFF IMPLEMENTATION

7.1. Initial Cutoffs

MSPHL set initial high risk cutoff values for each of the 4 LSDs by analyzing ~13,000 presumed normal de-identified specimens and 29 known affected specimens (three newborn specimens and 26 specimens from patients ranging from three months to 70 years of age) on the SEEKER™. This testing was done over the course of two months prior to the start of the pilot phase. The cutoffs were chosen to detect all known affected specimens, while minimizing the false positive rate.

MSPHL set their borderline enzymatic cutoff at a slightly higher level than the high risk cutoff values. Initial results below the borderline cutoff were flagged by the system for re-punching (from same dried blood spot sample, but different blood spot circles on the card) and re-testing in duplicate.

Based on the pre-pilot testing, the high risk and borderline cutoffs were set at the levels indicated in Table 4.

Table 4: Initial cutoffs for clinical study

Enzyme	High Risk (µmol/L/hour)	Borderline (µmol/L/hour)
IDUA	4.0	5.0
GAA	8.0	10.0
GBA	4.5	7.0
GLA	5.5	7.0

7.2. Cutoff Changes

To make a change to the high risk cutoff MSPHL required the approval of their Lysosomal Storage Disorders (LSD) Task Force. The LSD Task Force consisted of geneticists, genetic counselors, newborn screening laboratory staff, newborn screening follow-up staff, a chemist, and an adult with Fabry disease. The LSD Task Force met quarterly (or as needed to meet immediate concerns) to review results from the previous testing period and to decide if any recommended cutoff changes needed to be implemented.

During the course of this clinical study, MSPHL made changes to the cutoffs for the following reasons.

1. Reduce false negatives: If no or very few false positives (all or most screened positives are diagnosed with the condition) are found, then there is a high risk of false negatives. Cutoffs are increased in this case.
2. Reduce false positive rate: Newborn screening programs have to balance the risk of false negatives (undetected cases, primarily milder forms and late onset forms) with risk of false positives (additional cost, undue anxiety for parents). If the rate of false positive is

disproportionally high for the incidence of the disorder and no affected newborns are found, cutoffs are decreased.

3. Seasonal changes: High heat and humidity is known to reduce the activity of lysosomal and other enzymes. This results in a higher rate of false positives in certain seasons. This has previously been observed in testing for biotinidase, galactosemia, and G6PD[8]. Details of this phenomena can be found in Appendix B.

7.3. Age Specific Cutoffs

At the outset of the study, there was no expectation of age related changes in LSD activity as this was the first newborn screening study for these LSDs. During the study, MSPHL observed that the activity values for samples from older newborns (14+ days of age) were significantly lower than the values for younger newborns for the GAA, GBA, and GLA assays. Age specific cutoffs were implemented for newborns greater than or equal to 14 days of life at sample collection in May 2013. These cutoffs were set at approximately the same percentile rank as the 1-6 day cutoffs to reduce the false positive rate.

Similarly, another set of cutoffs for samples collected between 7-13 days of life were applied in 2014 after similar analysis revealed low activity values for that population relative to the 1-6 day population. More information about the population statistics at each age range can be found in Appendix A.

7.4. Cutoff Summary

Tables 5-8 list the cutoff values that were used during the pilot and pivotal phases of the clinical study. Cutoffs were changed as data was collected supporting the validity of the cutoffs.

Table 5 : IDUA Cutoff Values During Study Period (Cutoffs in bold type indicate the pivotal phase)

	IDUA Cutoffs	**High Risk**			**Borderline**		
Study Period	**Effective Date**	**0-6 Days**	**7-13 Days**	**14+ Days**	**0-6 Days**	**7-13 Days**	**14+ Days**
Pilot	1/15/2013	4.0	4.0	4.0	5.0	5.0	5.0
	5/15/2013	3.0	3.0	3.0	5.0	5.0	5.0
Pivotal	7/3/2013	2.0	2.0	2.0	5.0	5.0	5.0
	11/18/2014	1.5	1.5	1.5	5.0	5.0	5.0

Table 6 : GAA Cutoff Values During Study Period

	GAA Cutoffs	High Risk			Borderline		
	Effective Date	**0-6 Days**	**7-13 Days**	**14+ Days**	**0-6 Days**	**7-13 Days**	**14+ Days**
Pilot	1/15/2013	8.0	8.0	8.0	10.0	10.0	10.0
	3/9/2013	7.0	7.0	7.0	10.0	10.0	10.0
Pivotal	5/16/2013	7.0	7.0	4.5	10.0	10.0	10.0
	11/12/2013	7.2	7.2	4.5	10.0	10.0	10.0
	6/23/2014	7.2	4.5	4.5	10.0	10.0	10.0

Table 7 : GBA Cutoff Values During Study Period

	GBA Cutoffs	High Risk			Borderline		
	Effective Date	**0-6 Days**	**7-13 Days**	**14+ Days**	**0-6 Days**	**7-13 Days**	**14+ Days**
Pilot	1/15/2013	4.5	4.5	4.5	7.0	7.0	7.0
	3/9/2013	7.0	7.0	7.0	8.0	8.0	8.0
	5/16/2013	7.0	7.0	5.0	8.0	8.0	8.0
	6/5/2013	5.5	5.5	5.0	8.0	8.0	8.0
Pivotal	8/28/2013	5.5	5.5	5.0	7.0	7.0	7.0
	6/23/2014	5.5	4.0	4.0	7.0	7.0	7.0

Table 8 : GLA Cutoff Values During Study Period

	GLA Cutoffs	High Risk			Borderline		
	Effective Date	**0-6 Days**	**7-13 Days**	**14+ Days**	**0-6 Days**	**7-13 Days**	**14+ Days**
Pilot	1/15/2013	5.5	5.5	5.5	7.0	7.0	7.0
	3/9/2013	6.2	6.2	6.2	7.5	7.5	7.5
	5/16/2013	6.2	6.2	3.7	7.5	7.5	7.5
Pivotal	6/5/2013	7.0	7.0	3.7	8.0	8.0	8.0
	12/17/2013	8.0	8.0	3.7	10.0	10	10
	4/23/2014	8.0	5.0	3.7	10.0	5.0	5.0
	6/4/2014	7.0	5.0	3.7	9.0	5.0	5.0
	6/23/2014	7.0	5.0	3.0	9.0	5.0	5.0

8. STUDY LIMITATIONS

8.1. Acquiring New Samples

Newborn screening labs may request an additional screen from a newborn for a number of reasons; it may be done automatically in the case of a premature birth or upon receipt of a sample designated as poor quality. Another screen may also be collected based on an indeterminate result of any of the screening tests performed using the sample.

During the clinical study, MSPHL was not able to obtain a repeat specimen based on the LSD results because the LSD program was in the statewide validation phase. As a result, when the output of the Risk Assessment was to designate the sample as low risk, another sample could not be collected, and the sample was presumed normal unless another sample was subsequently collected for reasons other than the LSD test results.

The ability to request another sample based on the LSD test results would have removed subjectivity from the screening process during the clinical study. It is recommended that laboratories request an additional sample from the newborn for samples that test below the high risk cutoff but are not determined to be at high risk for disease.

Because another sample could not be requested based on the LSD results, Risk Assessment was a very important part of the testing algorithm since it ultimately determined whether or not a newborn would be referred for confirmatory testing..

8.2. Cutoffs below LoD

During the clinical study, two of the cutoffs for the IDUA assay were below LoD (2.77). Additionally, one of the age specific cutoffs (14+) for the GLA assay was below the LoD (3.18). MSPHL follow up staff contacted the diagnostic referral centers to confirm that there were no additional newborns diagnosed with MPS I in the clinical study period. The two screens collected from the only prospectively identified MPS I affected newborn identified during the study support this, as the average activity value for both samples was less than the LoD.

8.3. Age Specific Cutoffs

Since this was the first prospective clinical study performed for these assays, there was no solid information about the effect of demographics (age, gender, gestational age, etc.) on lysosomal enzyme activity values. During the course of the study, MSPHL determined that three of the four enzymes exhibit a significant decrease in activity in older newborns, and that age specific cutoffs would be required to accurately assess samples from older newborns. There were still a very large number of presumed normal screens collected in the 7-13 and 14+ day ranges (10,620 and 12,880, respectively). The lognormalized population distributions for each age range closely mirror the distributions for the 1-6 day group in shape and variability, although the mean decreases for each age range.

Very few affected samples were collected in the 7-13 and 14+ age ranges to perform a receiving operator characteristic (ROC) analysis, but the cutoffs were set near the same percentile rank as the 1-6 day populations.

8.4. DBS Variability

Collection of sample by blood spotting introduces some intrinsic variability into the sample testing process (between punches within the same spot, between spots on the same card, and between cards from the same newborn). Although guidance on sample collection states that the circles marked on the filter paper should be filled and evenly saturated with a single drop of blood, samples that do not meet this standard are regularly received by NBS laboratories. Variables in preparation of a blood spot include multispotted samples (multiple small blood spots used to make 1 larger blood spot), blood applied to both sides of the card, layered blood spots, compressed blood spots, blood spot volume, hematocrit, chromatographic effects of absorption, humidity, blood sample source, and filter paper performance characteristics. Prior studies have been conducted to quantify the variability of specific analytes on DBS[9]; enzyme based assays may be even more sensitive to the causes of variability due to the additional factor of denaturation.

The intrinsic variability within and between blood spots leads to a high coefficient of variability near the cutoffs. The risk of this variability was reduced in several ways:

- Borderline and high risk cutoffs: It is recommended that the borderline cutoff be set above the high risk cutoff by a minimum of two times the assay reproducibility. This accounts for the potential of punch to punch variability causing a false negative result.
- Repeat of samples below borderline cutoffs: All samples that test below the borderline cutoff should be repeated in duplicate. By including additional measurements, the measurement error, including spot variability, is reduced proportionally to the square root of the number of tests.

9. CLINICAL DATA ANALYSIS

The clinical study was conducted using specimens received at MSPHL between January 11, 2013 and January 14, 2015. The testing of these specimens occurred from January 15, 2013 and January 14, 2015, inclusive. The study period was retrospectively separated into "pilot" and "pivotal" phases. The pilot phase included newborns born on or before 8/26/2013. The pivotal phase included newborns born on or after 8/27/2013.

Detailed analysis is provided for the pivotal phase of the study period. This analysis uses the cutoffs that were appropriate at the time of testing and for the specific age range referenced in Tables 5-8. An overall summary of the entire study including both pilot and pivotal phases is also provided. A total number of 154,412 newborns were screened. Of those screened, some newborns had samples that did not meet the sample acceptance criteria listed below.

9.1. Specimens (Samples) included in analysis

The total prospective newborn specimens received by MSPHL during the clinical study period was 182,917. When analyzing system performance during both the pilot and pivotal phases, the following samples were excluded from the 182,917 samples received:

- **Samples collected at < 24 h of life**: Samples collected at less than 24 hours of life trigger a mandatory retest in Missouri. These samples were excluded from the analysis. (n=3,713)
- **Samples with no recorded age at collection**: If no age at collection was recorded, the sample was excluded from the analysis because age-specific cutoff values could not be applied. (n=580)
- **Samples with no valid data point**: Some samples did not have a valid test result. These samples were excluded from the analysis. (n=21)
- **Samples designated as poor quality on receipt by MSPHL**; Reasons for this designation include incomplete saturation, supersaturation, and contamination– in addition to other criteria. These samples were excluded from the analysis. (n=3,055)

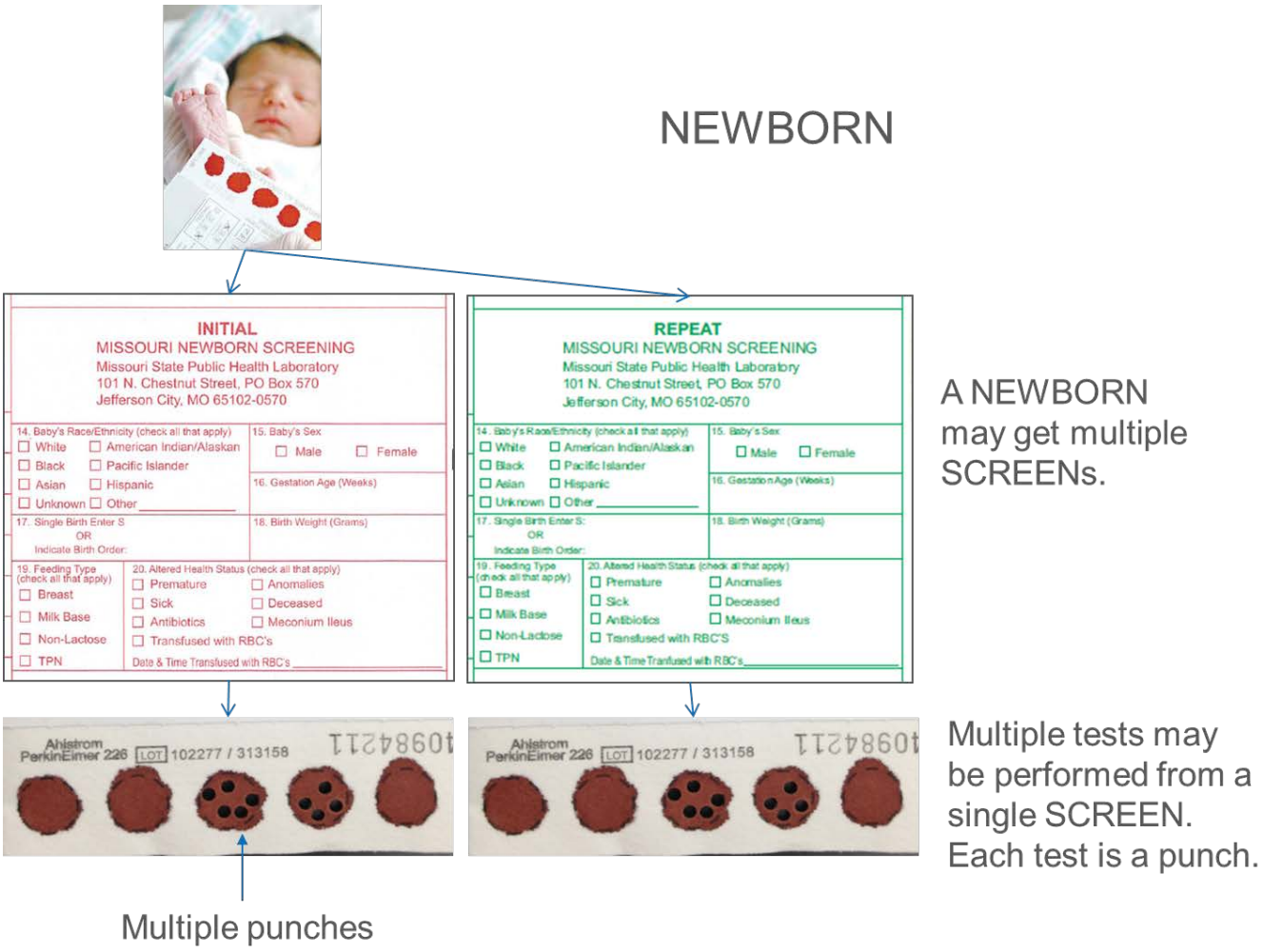
The total number of valid samples after these exclusions was 175,548.

9.2. Newborns included in analysis

Of the 175,548 valid specimens, the resulting number of newborns with at least one valid screen was 153,697. The number of specimens is higher than the number of newborns since multiple specimens were collected from the same newborn for a number of reasons, including – but not limited to - premature birth, low birth weight, poor quality initial sample, and sample collection before 24 hours of life (all of these cases require a mandated repeat screen per Missouri state guidelines). Figure 2 illustrates the relationship between newborns, screens, and tests. MSPHL tested each collected screen independently of any previous results from the newborn.

Of the 153,697 newborns with at least one valid screen included in the analysis, 48,608 were born during the pilot phase and 105,089 were born during the pivotal phase.

Figure 2 - Illustration of Relationship Between Newborns, Screens, and Tests (DBS Screening Card)



9.3. Pivotal Phase Results

The pivotal phase included newborns born on or after 8/27/2013. A total of 105,089 newborns were screened during the pivotal phase. Device performance and clinical results for each assay during the pivotal phase are summarized in **Sections 9.3.1 to 9.3.3**. This analysis includes all valid samples from the newborns, including repeat screens.

Table 9 provides descriptions for each row in the pivotal phase result tables.

Table 9 : Description of Rows in Assay Performance Tables

Grouping / Row Title		Description of Row
	Newborns	The total number of newborns included in the pivotal phase analysis
1st Test	**All screens with first result above borderline**	The number of newborns where the first valid test result for all screens for the newborn were above the borderline cutoff and therefore *presumed normal*.
	At least one screen with first result below borderline	The number of newborns where the first valid test result for at least one screen was below the borderline cutoff and therefore was *repeated*. This is also presented in parentheses as a percentage of newborns screened.
Average of all Tests	**All screens w/ avg. above high risk**	After retesting, all screens from the newborn had an average activity value above the high risk cutoff and were therefore *presumed normal.*
	At least one screen w/ avg. below high risk	After retesting, at least one screen from the newborn had an average activity value below the high risk cutoff and *Risk Assessment was performed prior to referral decision.* This is also presented in parentheses as a percentage of newborns screened.
	Referred / Not Referred:	After Risk Assessment, the number of newborns that were *presumed affected and referred* or *presumed normal and not referred*. This is also presented in parentheses as a percentage of newborns screened.
Referred Sample Summary		The number of newborns in each category of final clinical diagnosis (i.e. true positive, false positive, etc.)
Samples Not Referred Summary		The number of newborns with average activity values below the high risk cutoff that were not referred for each of the reasons provided related to the Risk Assessment.
Performance Summary	**Total Presumed Normal**	The number of newborns *presumed normal* after the first test, repeat testing, or Risk Assessment
	Total Presumed Affected	The number of newborns *presumed affected and referred* after Risk Assessment
	True Positives	The number of newborns *confirmed to have the disorder* after diagnostic and molecular follow up testing
	False Positives	The number of newborns *confirmed to be absent of disease* after follow up testing; this includes carriers and pseudodeficiencies
	Refused/Moved	The number of newborns who *did not receive confirmatory testing* either because it was refused or because the family moved out of Missouri
	Below HR / Not Referred	The number of newborns with activity values below the high risk cutoff that were *not referred* after Risk Assessment
	False Positive Rate (FPR)	The number of newborns with false positive results divided by the total number of newborns – minus true positives and newborns who did not receive follow up testing
	False Positive Rate (FPR) incl. below HR/not referred	The number of newborns with activity values below the high risk cutoff (including those referred and not referred) divided by the total number of newborns – minus true positives and newborns who did not receive follow up testing

9.3.1. Presumed Affected

The vast majority of the newborns were classified as presumed normal after a single test. Of the newborns that required additional testing to arrive at a screen result, most resolved to a presumed normal result after retesting; only a very small percentage were subjected to risk assessment for each disorder.

		IDUA	GAA	GBA	GLA
	Newborns	105,089	105,089	105,089	105,089
1st Test	All screens with first result above borderline	104,245	103,691	104,336	103,670
	At least one screen with first result below borderline	844 (0.80%)	1,398 (1.33%)	753 (0.72%)	1,419 (1.35%
Average of all tests	All screens with average above high risk	793	1,288	690	1,219
	At least one screen with average below high risk	51 (0.03%	110 (0.10%)	63 (0.06%)	200 (0.19%)
	Referred	33 (0.03%)	45 (0.04%	8 (0.01%)	60 (0.06%)
	Not Referred	18 (0.02%	65 (0.06%)	55 (0.05%)	140 (0.13%)
Referred Sample Summary	True Positive	0	7	2	30
	Normal, False Positive	9	23	3	26
	Pseudodeficiency, False Positive	20	8	0	0
	Carrier, False Positive	2	7	2	0
	Refused	1	0	1	3
	Moved	1	0	0	1

9.3.2. Presumed Normal after risk analysis

		IDUA	GAA	GBA	GLA
Samples Not Referred Summary	Other normal sample from newborn	13	42	34	84
	Suspected poor quality sample	2	12	12	29
	Altered health status	0	2	3	6
	Other	3	9	6	21

The Risk Assessment categories in the above table include:

- Other normal sample from newborn: prior sample from newborn (n=131), later sample from newborn (n=42).

- Suspected poor quality sample: other assay below borderline (n=49), spot variability (n=3), contaminated sample (n=1), other assay low normal (n=2).
- Altered health status: transfused (n=11).
- Other: Outliers excluded (n=31), different cutoff applied (n=3), retrospectively referred (n=2), age related enzyme decrease (n=1), and multiple reasons (n=2).

9.3.3. Clinical Performance Summary

		IDUA	GAA	GBA	GLA
Performance Summary	Total Presumed Normal	105,056	105,044	105,081	105,029
	Total Presumed Affected	33	45	8	60
	True Positives	0	7	2	30
	False Positives	31	38	5	26
	Refused/Moved	2	--	1	4
	Below HR/Not Referred	18	65	55	140
	False Positive Rate (FPR)	0.029%	0.036%	0.005%	0.025%
	False Positive Rate (FPR) – included below HR/not referred	0.047%	0.098%	0.057%	0.158%

The last line of the above table, False Positive Rate (FPR) – included below HR/not referred, references the number of newborns with activity values below the high risk cutoff (including those referred and not referred) divided by the total number of newborns – minus true positives and newborns who did not receive follow up testing. This would be the worst-case false positive rate if risk Assessment were not performed.

9.3.4. Retest Rate

A total of 120,118 specimens from 105,089 newborns were tested during the pivotal phase. The number of specimens is higher than the number of newborns since many newborns had multiple specimens collected, either due to their health status (low birth weight, premature, transfused) or due to other routine newborn screening results.

A total of 137,153 individual tests were performed on the 120,118 specimens, resulting in an overall retest rate of 1.142 or 14.2% per specimen. Of these 137,153 individual tests 7,785 tests were due to invalid data points generated by the device representing an invalid rate of 5.7% (per specimen). The remaining 8.5% were due to retesting triggered by activity values below borderline cutoff for one of the assays (as required by the screening protocol).

Due to the multiplexed nature of the test a retest trigged by one assay will result in all 4 assays being repeated. The average retest rate on a per assay basis is 14.2 / 4 = 3.55%. The average invalid rate on a per assay basis is 5.7 / 4 = 1.43%.

9.4. Two Year Clinical Study Summary

During the entire study period (both pilot and pivotal), 275 newborns were presumed affected and referred for confirmatory diagnosis. The results of the confirmatory testing are listed in Table 10.

Table 10: Presumed affected individuals – entire study period

	IDUA	GAA	GBA	GLA
Presumed Affected	73	79	19	104
- Refused	1	0	1	4
- Moved	1	1		1
- Normal, False Positive	31	33	13	47
- Carrier, False Positive	4	14	2	0
- Pseudodeficient, False Positive	35	14	0	0
- True Positive	**1**	**17**	**3**	**52**
- Unclassified Onset	1	0	1	45
- Unknown Onset / Genotype of Unknown Significance	0	3	2	4
- Late Onset	0	9	0	3
- Infantile Onset	0	5	0	n/a
- Classical Infantile	0	3	0	n/a
- Non-Classical Infantile	0	2	0	n/a
False positive rate	0.045%	0.039%	0.010%	0.030%

9.4.1. Disease Incidence

Based on the number of newborns analyzed for this two-year study period (n=153,697), the incidence of each of the disorders was calculated and is summarized below.

Table 11: Incidence Rate During Clinical Study

	Incidence from MSPHL Study
MPS I (IDUA)	1 : 153,697
Pompe (GAA)	1 : 9,041
Gaucher (GBA)	1 : 51,232
Fabry (GLA)	1 : 2,956

10. CONCLUSIONS

10.1. True Positives

10.1.1. Confirmed Positive Newborns

MPS I: The only confirmed MPS I positive newborn was screened during the pilot phase of the study. This was reported as a severe case of MPS I (Hurler disorder).

Pompe: Seventeen newborns were confirmed affected with Pompe disorder during the clinical study. Five of these newborns were confirmed to have infantile onset Pompe disorder, which results in death at a median age of approximately 12 months. Nine newborns were confirmed with late onset Pompe disorder, which can begin to cause symptoms as early as childhood. Three of the newborns were confirmed to have genetic mutations of unknown significance; they will require long term follow up to test for disease onset.

Gaucher: Three newborns were confirmed affected with Gaucher disorder. One newborn was confirmed to have Type 1 – non-neuronopathic Gaucher disorder of unknown onset. The other two newborns were found to have mutations of unknown significance; they will also require long term follow up.

Fabry: 52 newborns were confirmed affected with Fabry disorder during the study. The status of 45 of these newborns was reported by the referral center as "Fabry disorder" without a classification of onset. Three newborns were confirmed with mutations consistent with late onset Fabry disorder and four newborns were confirmed to have mutations of unknown significance.

10.1.2. Disease Incidence

Based on the number of newborns analyzed for this two-year study period (n=153,697), the incidence of each of the disorders was calculated and is summarized below, along with previously published incidence.

Table 12: Incidence Rate During Clinical Study

	Incidence from MSPHL Study	Published Incidence
MPS I (IDUA)	1 : 153,697	1:54,000 – 1:185,000
Pompe (GAA)	1 : 9,041	1:28,000
Gaucher (GBA)	1 : 51,232	1:57,000
Fabry (GLA)	1 : 2,956	1:1,500 – 1:13,000

For Pompe disorder the MSPHL incidence was about 3 times higher than published rates of incidence. For MPS I, Gaucher, and Fabry the MSPHL incidence is comparable to the published rates of incidence.

10.2. False Negatives

10.2.1. Surveillance

The Missouri Department of Health and Senior Services has an active surveillance program to track any reports of false negative results to the contracted metabolic centers they use for confirming diagnosis of any of these LSDs. Based on information from this surveillance program there were no known false negative results reported during the 2 year study or in 15 months following the conclusion of the study. Newborns screened during the clinical study with early onset disorders would have been reported to one of these metabolic centers.

10.2.2. Incidence

Based on the number of newborns analyzed during the two-year study period, the incidence rates in the study were 1: 153,697 for MPS I, 1: 9,041 for Pompe, 1: 51,232 for Gaucher, and 1: 2,956 for Fabry, resulting in a combined incidence rate of 1:2,105 (Table 12). The fact that the individual incidence rates for each disorder in the study agree with, or are higher than, published literature suggests that risk related to false negatives is minimal.

10.2.3. Limitations

- False negative reporting is based on limited data on late onset forms of the disorders since it would take several years to identify an undetected late onset case.
- Certain late onset forms for Pompe disorder may have GAA enzymatic activity in the normal range and result in a false negative[10].
- For female Fabry patients GLA enzyme activity is highly variable and it typically falls in the normal range. Therefore, GLA enzyme activity for females should be interpreted with caution as most female carriers can have enzyme activity in the normal range which results in a normal screen.

10.3. False Positives

During the clinical study, several patients with reduced enzymatic activity were confirmed by follow-up testing to be either carriers or pseudodeficient for the referred condition. Carriers were identified for MPS I, Pompe and Gaucher disorders, while pseudodeficiencies were identified for MPS I and Pompe disorders. Carriers and pseudodeficient newborns may exhibit enzymatic activity below the high risk cutoff while remaining asymptomatic for the disorder and are included in the false positive calculations.

Table 13 indicates the false positive rate during the pivotal phase and entire study period for each assay.

Table 13 : False Positive Rates During the Pivotal Phase and Entire Study Period

	IDUA	GAA	GBA	GLA
Pivotal False Positive Rate (FPR)	0.029%	0.036%	0.005%	0.025%
Overall False Positive Rate (FPR)	0.045%	0.039%	0.010%	0.030%

The false positive rates during the pivotal phase and the entire study period for each assay are well below the typical goal of a 0.1% FPR for newborn screening assays.

A number of newborns had average test results below the high risk cutoff during the study but were not referred based on the results of the Risk Assessment. If the Risk Assessment was not performed and these newborns would presumably be false positive results given that the surveillance program has not identified any false negatives. If these were added as false positives, the adjusted worst-case false positive rate would increase as shown in Table 14.

Table 14 : False Positive Rates Including Samples Below High Risk But Not Referred as False Positives

	IDUA	GAA	GBA	GLA
Pivotal False Positive Rate (FPR) including newborns below the high risk cutoff and not referred as false positives *** Note: This does not represent how the study was conducted, and is retrospective in nature**	0.047%	0.098%	0.057%	0.158%

Even without the Risk Assessment, the worst-case false positive rates for the IDUA, GAA, and GBA assays were still below the target of 0.1% FPR[11]. For the GLA assay, the worst-case FPR is above the 0.1% target. However, this worst-case FPR is acceptable given the incidence rate of Fabry disorder (1 in 2,956). Tests for disorders with similar incidence rates, like congenital hypothyroidism (CH), typically accept much higher false positive rates.[11]

10.4. Reference Ranges

All presumed normal initial screen samples during the entire study period were also analyzed to determine the reference interval of normal samples. Invalid data points were excluded and values for samples that were tested more than once were averaged.

The reference interval was calculated using Analyse-it® version 4.0 using quantiles to estimate the reference interval. The selected reference limit was 0.1%. The final high risk cutoff for each assay for each age at collection group is also listed for comparison.

Table 15 : Calculated normal reference intervals

IDUA	N	0.10% Quantile	Final High Risk Cutoff
1-6 day	151,960	3.17	1.50
7-13 day	10,620	2.34	1.50
14+ day	12,880	1.82	1.50
GAA	**N**	**0.10% Quantile**	**Final High Risk Cutoff**
1-6 day	151,960	7.13	7.20
7-13 day	10,620	5.23	4.50
14+ day	12,880	4.27	4.50
GBA	**N**	**0.10% Quantile**	**Final High Risk Cutoff**
1-6 day	151,960	6.10	5.50
7-13 day	10,620	4.61	4.00
14+ day	12,880	4.07	4.00
GLA	**N**	**0.10% Quantile**	**Final High Risk Cutoff**
1-6 day	151,960	6.70	7.00
7-13 day	10,620	4.45	5.00
14+ day	12,880	3.41	3.00

10.5. Retest Rate

A total of 120,118 specimens from 105,089 newborns were tested during the pivotal phase. 137,153 individual tests were performed on the 120,118 specimens, resulting in an overall retest rate of 1.142 or 14.2% per specimen. Of these 137,153 individual tests, 7,785 tests were due to invalid data points generated by the device representing an invalid rate of 5.7% (per specimen). The remaining 8.5% were due to retesting triggered by activity values below borderline cutoff for one of the assays (as required by the screening protocol).

Due to the multiplexed nature of the test a retest trigged by one assay will result in all 4 assays being repeated. The average retest rate on a per assay basis is 14.2 / 4 = 3.55%. The average invalid rate on a per assay basis is 5.7 / 4 = 1.43%.

APPENDIX A: AGE AT COLLECTION RELATED CHANGES IN ACTIVITY

During the study, MSPHL observed that the activity values for samples from older newborns (14+ days of age) were significantly lower than the values for younger newborns for the GAA, GBA, and GLA assays. They implemented age specific cutoffs for newborns greater than or equal to 14 days of life at sample collection in May 2013. They determined these cutoffs by evaluating both the median activity value of the population in that age range for each assay and setting the cutoff to the same percentile as the 1-6 day cutoff.

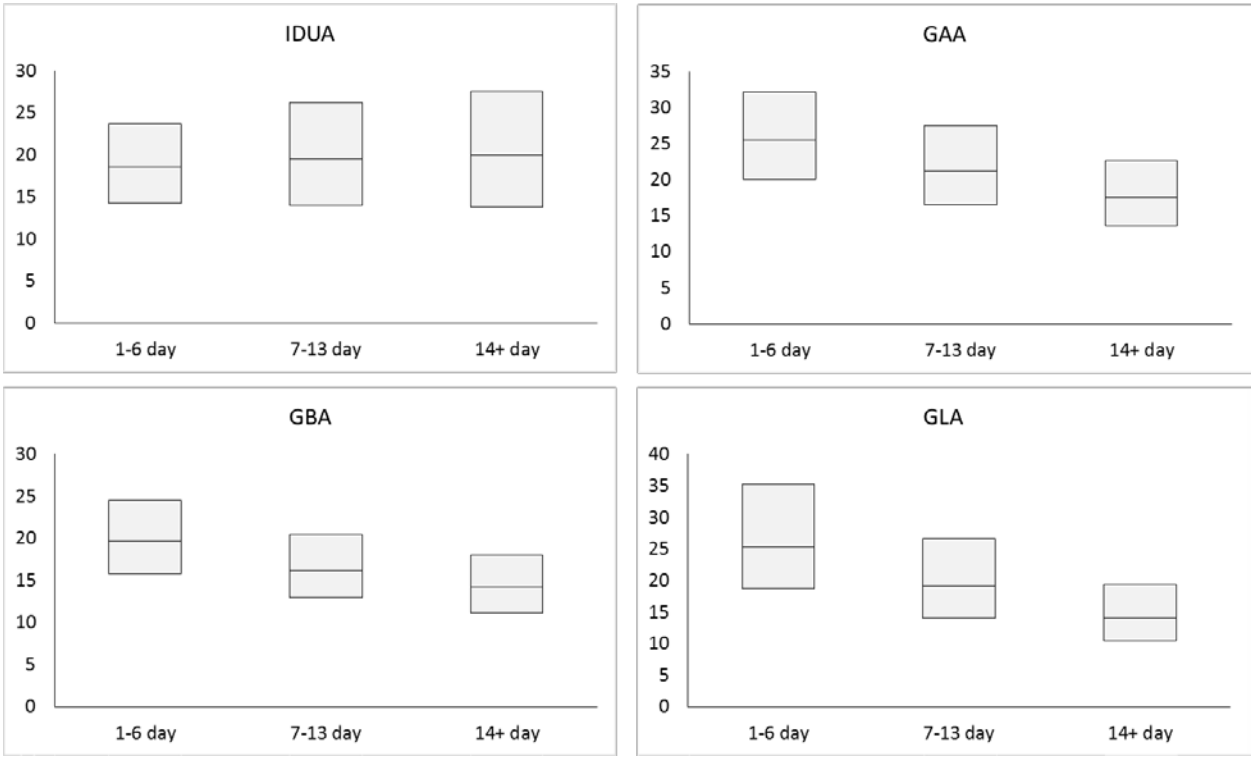
Similarly, another set of cutoffs for samples collected between 7-13 days of life were applied in 2014 after similar analysis revealed low activity values for that population relative to the 1-6 day population.

These population groups are generally lognormally distributed, as commonly produced by biological processes, so only parametric statistics can be applied until the activity values for each population are log-normalized. Once the population has been log-normalized, the standard deviation can be calculated to compare variability between assays and age groups. The standard and log-normalized population statistics are listed in Table 16.

Table 16: Mean and median activities – and log-normalized mean – of each assay, grouped by age at collection

	Raw Population Values					Lognormalized Values		
IDUA	**N**	**Mean**	**Q1**	**Median**	**Q3**	**N**	**Mean**	**StDev**
1-6 day	151,960	19.68	14.30	18.56	23.72	151,950	2.90	0.40
7-13 day	10,620	21.10	13.95	19.48	26.24	10,619	2.94	0.49
14+ day	12,880	21.90	13.82	19.99	27.55	12,879	2.95	0.55
GAA	**N**	**Mean**	**Q1**	**Median**	**Q3**	**N**	**Mean**	**StDev**
1-6 day	151,960	27.00	20.05	25.42	32.15	151,960	3.23	0.36
7-13 day	10,620	23.52	16.55	21.19	27.48	10,620	3.07	0.41
14+ day	12,880	19.29	13.56	17.48	22.67	12,880	2.87	0.41
GBA	**N**	**Mean**	**Q1**	**Median**	**Q3**	**N**	**Mean**	**StDev**
1-6 day	151,960	20.92	15.81	19.70	24.54	151,960	2.98	0.34
7-13 day	10,620	17.69	12.97	16.15	20.39	10,620	2.80	0.37
14+ day	12,880	15.65	11.14	14.20	18.00	12,880	2.67	0.40
GLA	**N**	**Mean**	**Q1**	**Median**	**Q3**	**N**	**Mean**	**StDev**
1-6 day	151,960	29.86	18.69	25.32	35.31	151,960	3.26	0.49
7-13 day	10,620	22.66	14.01	19.03	26.65	10,620	2.98	0.51
14+ day	12,880	16.82	10.45	14.01	19.41	12,880	2.68	0.51

Figure 3 : Boxplots for the Reference Ranges for the IDUA, GAA, GBA, and GLA Assays. Boxes represent Quartiles and Medians.



For the GAA, GBA, and GLA assays, this analysis confirms a significant decrease in activity values between the 1-6, 7-13, and 14+ day groups (the IDUA activity marginally increases over the same period). There is generally a slight increase in the standard deviation of the population with increasing age with the exception of IDUA, where there is a more significant increase. This is reflected in the reference ranges calculated for each assay, which can be found in Table 15 and Table 16.

APPENDIX B: WEATHER RELATED CHANGES IN ACTIVITY

The data analysis on prospective newborn specimens collected during the study (1/15/2013 – 1/14/2015 – inclusive) is reported below. Median enzymatic activity was calculated for every week of testing and plotted over time. Figures below show the plots of weekly median newborn enzymatic activity over time for GBA and GLA.

From the plots, it can be seen there is a clear reduction in enzymatic activity during the summer months at higher temperature. Since the DBS samples are shipped at room temperature, they are subjected to different temperature and humidity conditions during transit.

Based on the plots of weekly median shown below, it is recommended that the cutoff values can be adjusted with seasonal changes because of the variation in enzymatic activity. It is also recommended that the users monitor the patient medians over time.

The cutoff values for GLA were increased on 12/17/2013; at the time, this was recorded as a change due to performance of a new reagent lot. This was during the first year of testing and the cyclical nature of the assay performance throughout the year was not clearly understood. The cutoff value for GLA was lowered on 6/4/2014 as the temperature warmed and the median activity value decreased. Figures 4-7 illustrate the fluctuation in patient median for each of the four assays over the entire study period.

Figure 4 : Seasonal fluctuation in weekly patient median - IDUA

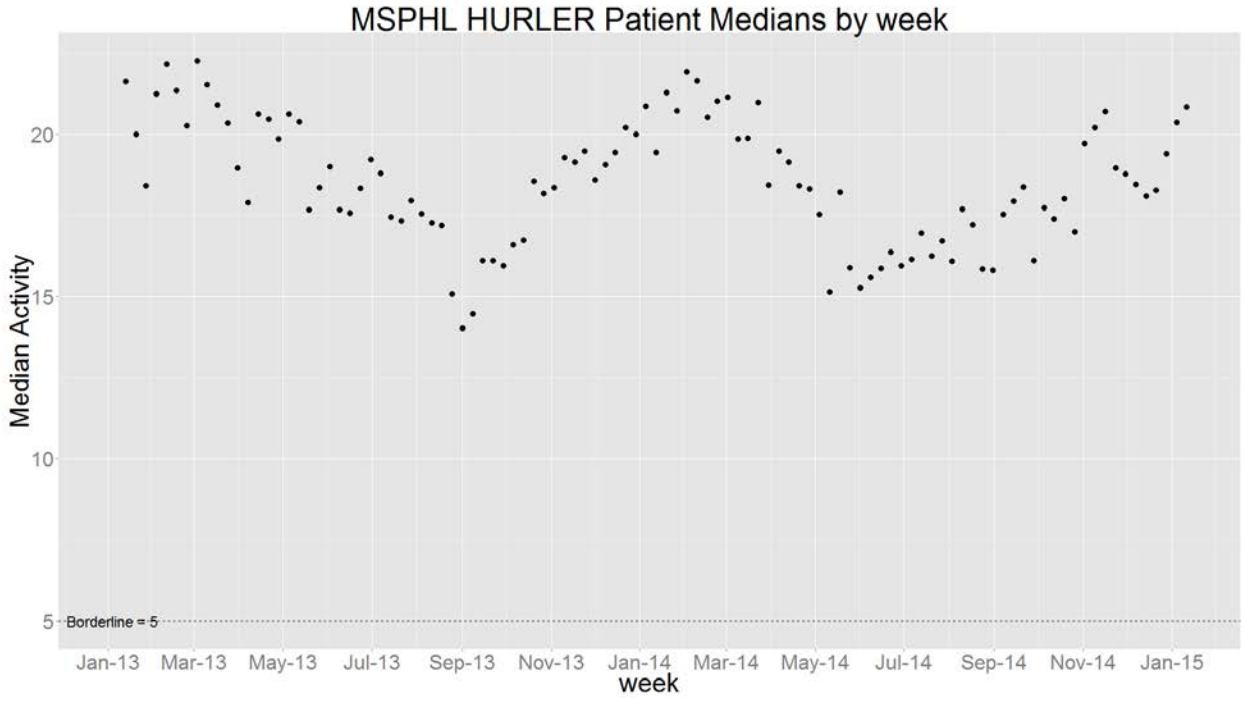


Figure 5 : Seasonal fluctuation in weekly patient median - GAA

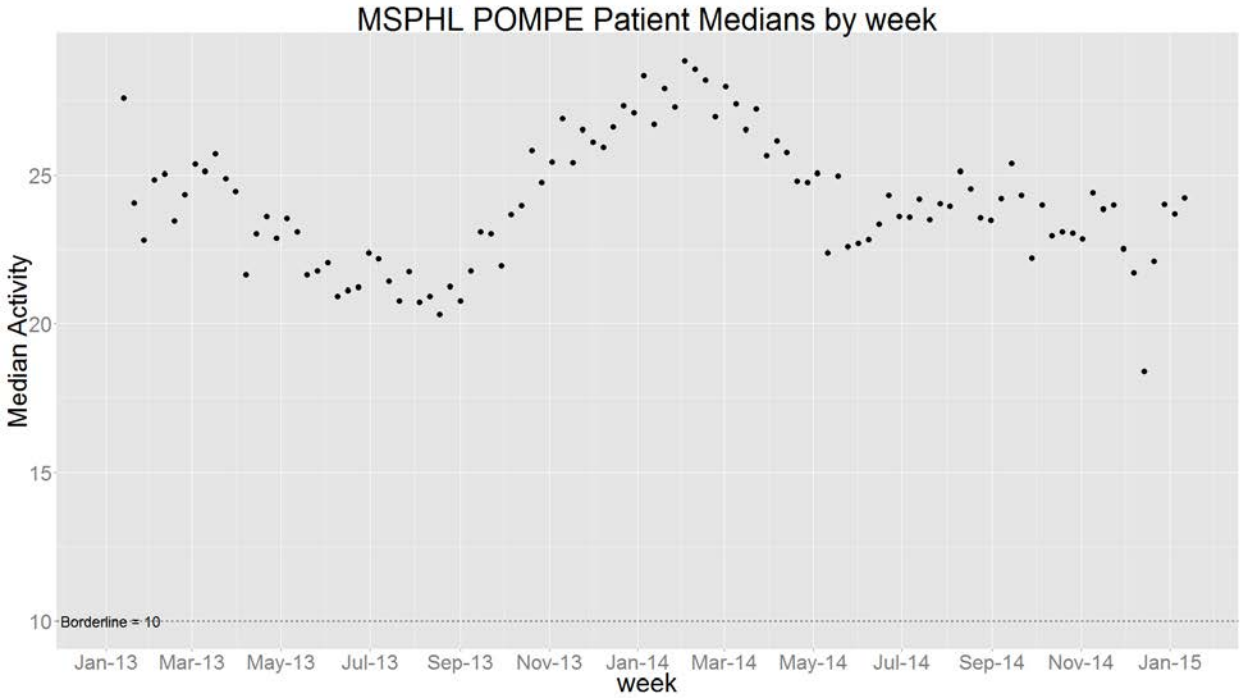


Figure 6 : Seasonal fluctuation in weekly patient median - GBA

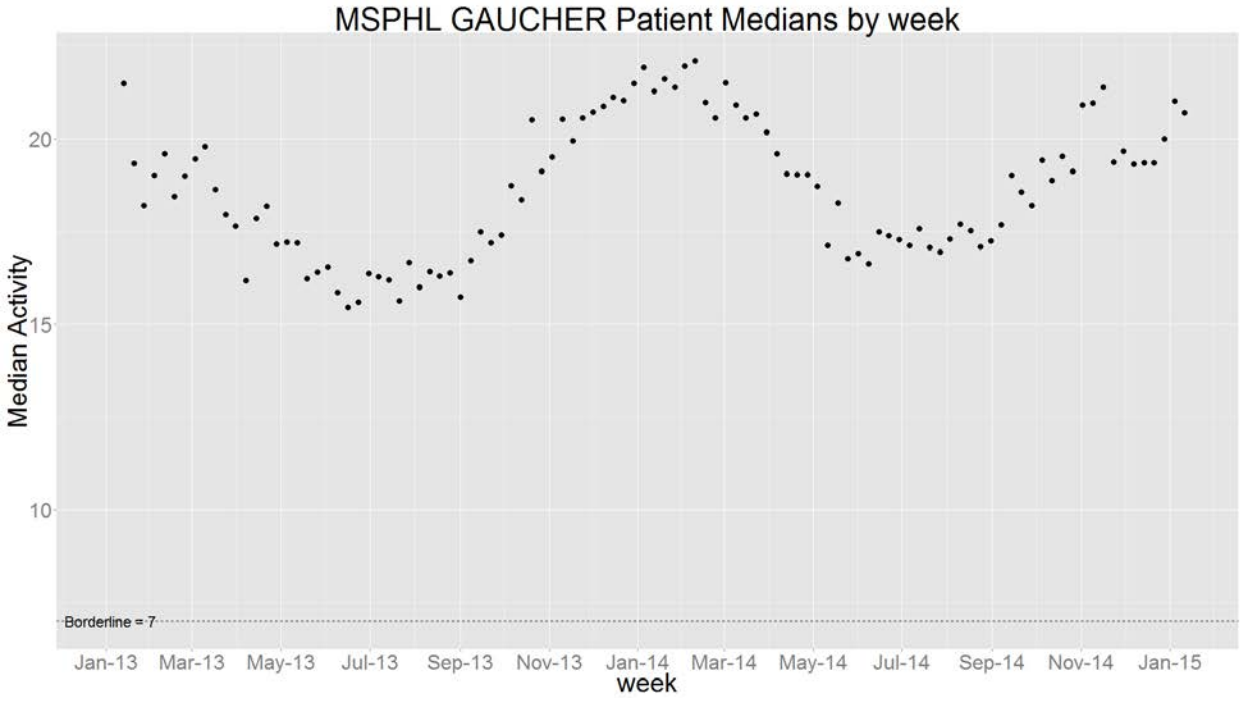
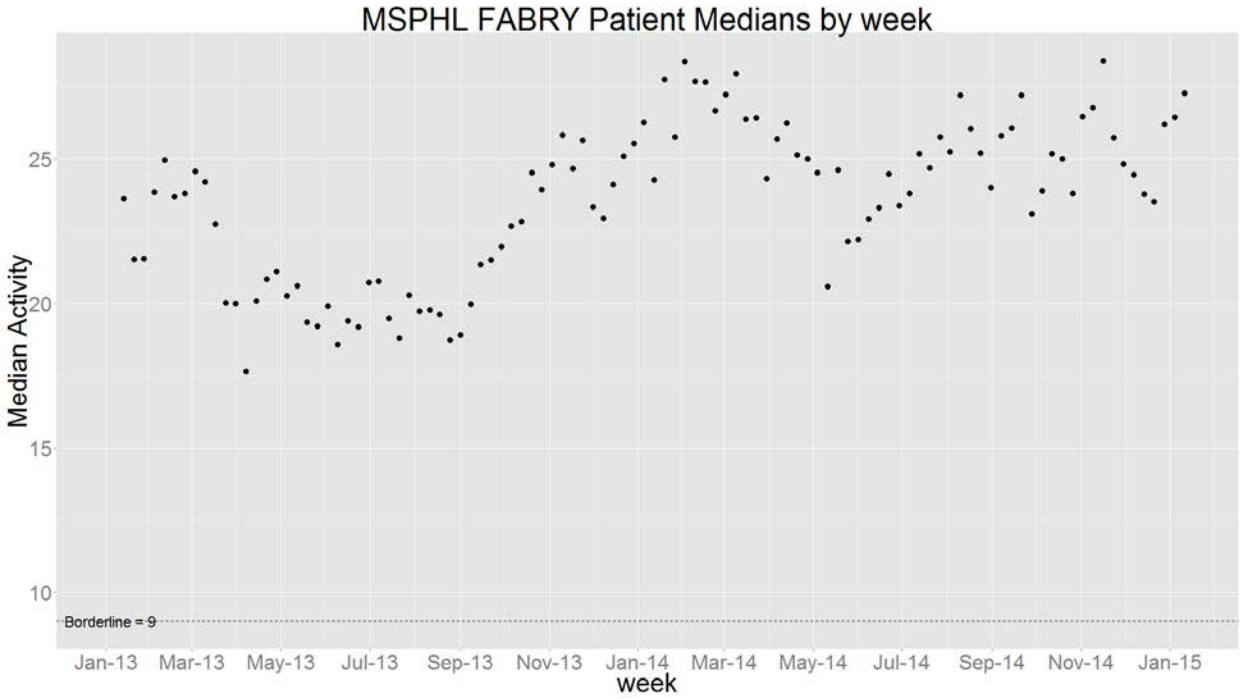


Figure 7 : Seasonal fluctuation in weekly patient median - GLA



APPENDIX C: DEVICE CHANGES

Pilot Phase

During the pilot phase there were 5 minor changes (addressing latent defects in software) and one major change (Seeker Cartridge) to improve the robustness of the System. All changes were validated and found to have no impact on the clinical performance of the device.

Date effective	Change summary	Rationale
Feb. 6, 2013	Assay Protocol (update to 48x4v5)	Modified assay protocol to reduce Stop Buffer (STB) dispense failures by reducing load on electrowetting effector during dispense
Feb. 28, 2013	Spot Logic (build 16109) and Assay Protocol (48x4v7)	Modified assay protocol to unify Stop Buffer (STB) dispense sequence through entire assay
Mar. 26, 2013	Spot Logic (build 1.2.1) and Assay Protocol (48x4v8)	Modified sequence used to merge Stop Buffer (STB) with reaction droplet to reduce droplet operation failure rate
May 1, 2013	Spot Logic (build 1.3.3)	Updated non-enzymatic hydrolysis (NEH) thresholds for new reagent lot to reduce rate of "n/a" results
May 21, 2013	Assay Protocol (update to 48x4v8)	Extended reservoir pulsing through IDUA detection sequence to reduce GLA reagent dispense failures
Aug. 16, 2013	Cartridge change (1093 to 1305) Assay Protocol (48x4v10) Spot Logic (build 1.4.1	The printed circuit board sub-component of the Seeker Cartridge was changed to make the assay protocol more robust. Specifically the change involved rewiring some of the electrodes of PCB in reagent and stop buffer reservoirs to prevent the formation of stray droplets during dispensing operations. Assay protocol and Spot Logic were revised to support this cartridge change.

Pivotal Phase

During the pivotal phase there was 1 minor change to the stop buffer formulation to improve robustness of the system.

Date effective	Change summary	Rationale
Nov 17, 2014	Stop Buffer (STB) formulation change.	Concentration of Tween in in Stop Buffer (STB) was increased from 0.01% to 0.04% w/v to improve robustness of droplet movement towards the end of the

		assay protocol.

REFERENCES

[1] Staretz-Chacham O, Lang TC, LaMarca ME, Krasnewich D, Sidransky E. Lysosomal storage disorders in the newborn. *Pediatrics*. 2009; 123;1191-1207.
[2] Linthorst GE, Poorthuis BJHM, Hollak CEM. Enzyme activity for determination of presence of Fabry disease in women results in 40% false-negative results. *Journal of the American College of Cardiology*. 2008; 51(21):2082.
[3] Biegstraaten M, Arngrímsson R, Barbey F, Boks L, Cecchi F, Deegan PB, *et al*. Recommendations for initiation and cessation of enzyme replacement therapy in patients with Fabry disease: the European Fabry Working Group consensus document. *Orphanet Journal of Rare Diseases*. 2015; 10(1).
[4] https://www.sos.mo.gov/adrules/csr/current/19csr/19c25-36.pdf
[5] CLSI. Blood Collection on Filter Paper for Newborn Screening Programs; CLSI document NBS01-A6.
[6] CLSI. Newborn Screening for Preterm, Low Birth Weight, and Sick Newborns; CLSI document NBS03-A.
[7] CLSI. Newborn Screening by Tandem Mass Spectrometry; Approved Guideline; CLSI document NBS04-A.
[8] Freer, D. E. (2005). Observations on heat/humidity denaturation of enzymes in filter-paper blood spots from newborns. *Clinical chemistry*, *51*(6), 1060-1062.
[9] Mei, J. V., Alexander, J. R., Adam, B. W., & Hannon, W. H. (2001). Use of filter paper for the collection and analysis of human whole blood specimens. *The Journal of nutrition*, *131*(5), 1631S-1636S.
[10] Chiang S-C, Hwu W-L, Lee N-C, Hsu L-W, Chien Y-H. Algorithm for Pompe disease newborn screening: Results from the Taiwan screening program. Mol Genet Metab. 2012;106 (3):281-286.
[11] Kwon C, Farrell PM. The Magnitude and Challenge of False-Positive Newborn Screening Test Results. Arch Pediatr Adolesc Med. 2000;154(7):714-718. doi:10.1001/archpedi.154.7.714.