

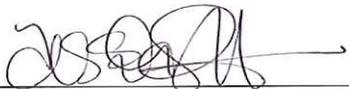
Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

This study has been technically reviewed and approved for use by:



Jennifer Zeffer, FBU Technical Leader | 020315
Date

This study was conducted and written by:



Jessica Skillman, Forensic Scientist III | 020315
Date

Introduction

Throughout the country, laboratories that conduct forensic DNA analysis have acknowledged a need for more consistent and detailed guidelines on the analysis, interpretation, reporting and statistical requirements for low-level and mixture DNA profiles. In addition, the DC DFS Scientific Advisory Board made a series of recommendations on November 5, 2014 to enhance the current practices and procedures used by the DC DFS Forensic Biology Unit¹. These recommendations specifically address the parameters established in the internal validation of the AmpFISTR® Identifiler® Plus amplification kit². In response to both the community and the DFS Scientific Advisory Board, the following studies were conducted:

- A reevaluation of the data supporting the analytical threshold, peak height ratio, and stochastic threshold,
- A stutter study which evaluated forward and reverse stutter,
- An evaluation of different approaches for data analysis, and
- A verification of the selected approach for data analysis.

While the AmpFISTR® Identifiler® Plus Internal Validation included an evaluation of both 28 and 29 cycle parameters, the laboratory currently only uses the 28 cycle procedure. This reevaluation will only include and apply to data amplified for 28 cycles.

I. Threshold Re-assessment

In order to address the recommendations made by the DC DFS Scientific Advisory Board, a reevaluation of the laboratory analytical threshold was needed. The internal validation set the analytical threshold at 70 rfu for all dye channels based on data quality and quantity, however, more recent literature has suggested a different set of techniques for establishing analytical threshold. Over time, a more robust set of data has also been collected from samples across kit lot numbers and instrumentation.

Using a set of random samples selected from the AmpFISTR® Identifiler® Plus Internal Validation and recent performance checks (Phase I Move, Phase II Move, 3130xl 2013 and 3130xl 2014³) different methods were used to calculate a set of proposed analytical thresholds. In addition, using the previously stated data, stochastic threshold and peak height ratios were evaluated to determine if modifications would be needed with different analytical thresholds.

a. Analytical Threshold

Multiple definitions and expressions exist for the lower limit of detection for analyzing data⁴. For this laboratory, the term analytical threshold will be defined as the minimum signal at which a peak can reliably be distinguished from noise. At this time, there are two generally accepted ways to calculate analytical threshold. Equation 1 (shown below) is suggested by the Scientific Working Group on DNA Analysis Methods (SWGDM) in section 1.1. of the Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories⁵.

Equation 1: Analytical Threshold = 2 (Maximum peak height - Minimum peak height)

Equations 2 and 3 were developed by the International Union of Pure & Applied Chemists (IUPAC). Limit of detection is defined as the smallest measure that can be detected with reasonable certainty. This equation is believed to result in an analytical threshold with 89-99.86% confidence that noise will be below this value⁶.

Equation 2: Limit of Detection = Average peak height + (3 × Standard Deviation peak height)

Another important calculation is the Limit of Quantitation (LOQ). The LOQ is the estimated limit in which the signal is not only reliably detected but also the peak height is reliably measured⁷.

Equation 3: Limit of Quantitation = Average peak height + (10 × Standard Deviation peak height)

Based on definition, all three equations were used to determine possible analytical thresholds for this laboratory's forensic DNA analysis. While the SWGDAM equation has

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

been used across the forensic community for the analysis of data from capillary electrophoresis instruments, the IUPAC equations are a mathematically supported approach for any type of analytical procedure.

All three equations were applied to the results from a set of 29 negative samples and a set of 40 samples which were amplified with ideal target quantities of DNA (0.5-1.0ng). These samples were randomly selected from the AmpFISTR® Identifiler® Plus Internal Validation and subsequent performance checks in order to include data from both of the laboratory's 3130xl instruments, different capillary lot numbers and reagent lot numbers. All sample results were analyzed with an analytical threshold of 1 rfu (relative fluorescence unit) in GeneMapper ID-X v1.4. For negative samples, all pull up and spikes were removed from calculations. For the samples which contained DNA, all peaks within four base pairs of a concordant peak were removed along with any artifacts such as pull up or background. The data was exported to a Microsoft Excel spreadsheet. Maximums, minimums, averages and standard deviations were calculated overall and separately for the different dye channels. The following results were obtained:

Negative Samples

SWGDM

Analytical Threshold = 2(Maximum peak height - minimum peak height)

	max	min	Analytical Threshold
blue	13	1	24
green	31	1	60
yellow	19	1	36
red	21	1	40
overall	31	1	60

IUPAC

Limit of Detection = Average peak height + (3 x Standard Deviation peak height)

Limit of Quantitation = Average peak height + (10 x Standard Deviation peak height)

	average	standard deviation	Limit of Detection	Limit of Quantitation
blue	2.42	1.17	5.94	14.16
green	3.09	1.45	7.43	17.58
yellow	6.19	2.29	13.08	29.14
red	6.69	2.58	14.44	32.53
overall	4.70	2.71	12.82	31.79

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Samples with Ideal Targets

SWGAM

Analytical Threshold = 2(Maximum peak height - minimum peak height)

	max	min	Analytical Threshold
blue	25	1	48
green	20	1	38
yellow	37	1	72
red	29	2	54
overall	37	1	72

IUPAC

Limit of Detection = Average peak height + (3 x Standard Deviation peak height)

Limit of Quantitation = Average peak height + (10 x Standard Deviation peak height)

	average	standard deviation	Limit of Detection	Limit of Quantitation
blue	2.68	1.85	8.23	21.20
green	3.27	1.94	9.09	22.69
yellow	6.63	3.20	16.23	38.62
red	7.32	3.41	17.55	41.42
overall	5.71	3.52	16.26	40.89

For Section III of this re-evaluation, all values were rounded and the following table used for different methods of data analysis:

	Equation 1 (Negatives)	Equation 1 (Samples)	Equation 2 (Negatives)	Equation 2 (Samples)	Equation 3 (Negatives)	Equation 3 (Samples)
blue	20	60	10	10	10	20
green	60	50	10	10	20	20
yellow	40	70	10	20	30	40
red	40	60	10	20	30	40
overall	60	70	10	20	30	40

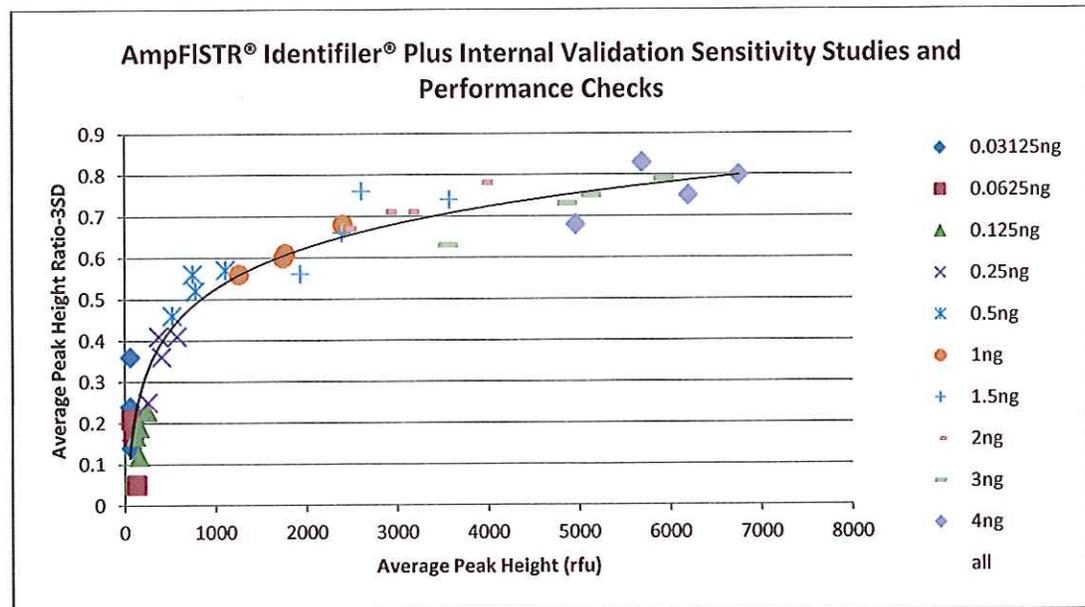
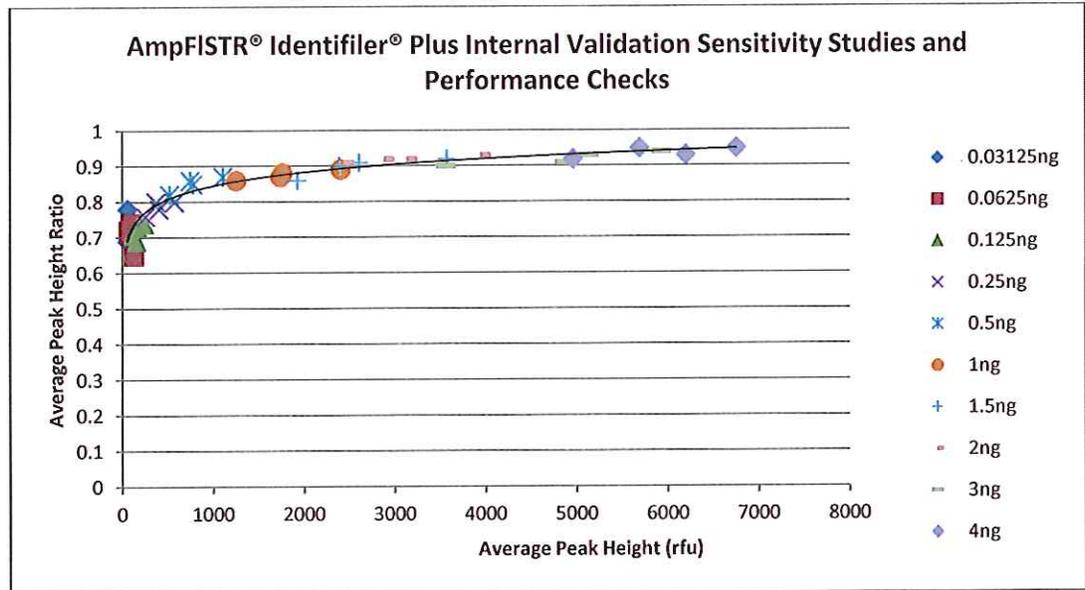
*A higher value was obtained for the negative samples than the samples which contained DNA. In order to determine the most appropriate value to use for the green dye channel, samples will be analyzed using the lower of the two values. Data will then be evaluated to determine which threshold is most appropriate for use.

b. Peak Height Ratio

Currently the laboratory uses a 55% peak height ratio for all samples in order to determine possible indications of a mixture and/or to calculate deduced profiles from intimate samples. This peak height ratio was determined using the original 70 rfu analytical threshold recommended in the AmpFISTR® Identifiler® Plus Internal Validation. Based on the data evaluated in the previous section, a change in peak height ratio may be necessary if the analytical threshold is changed.

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

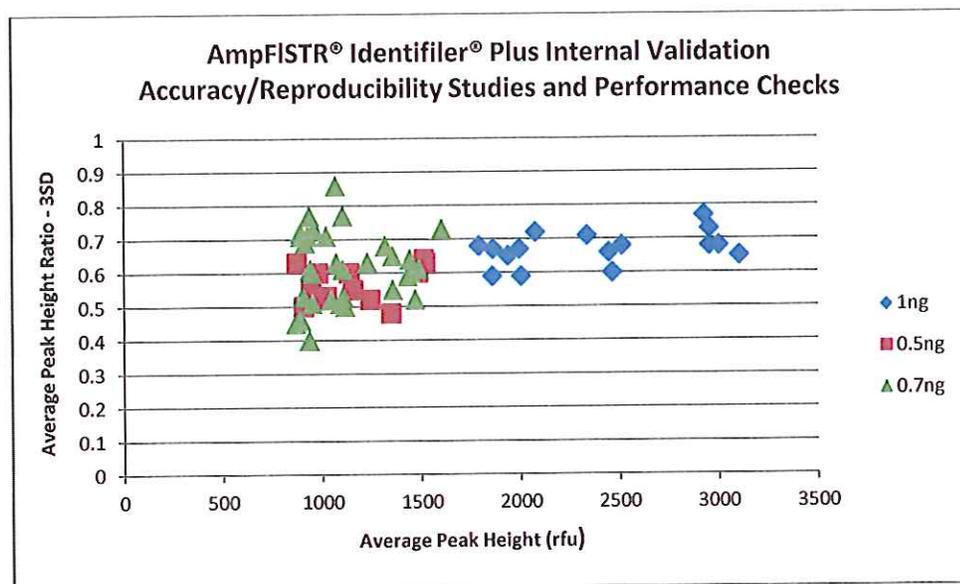
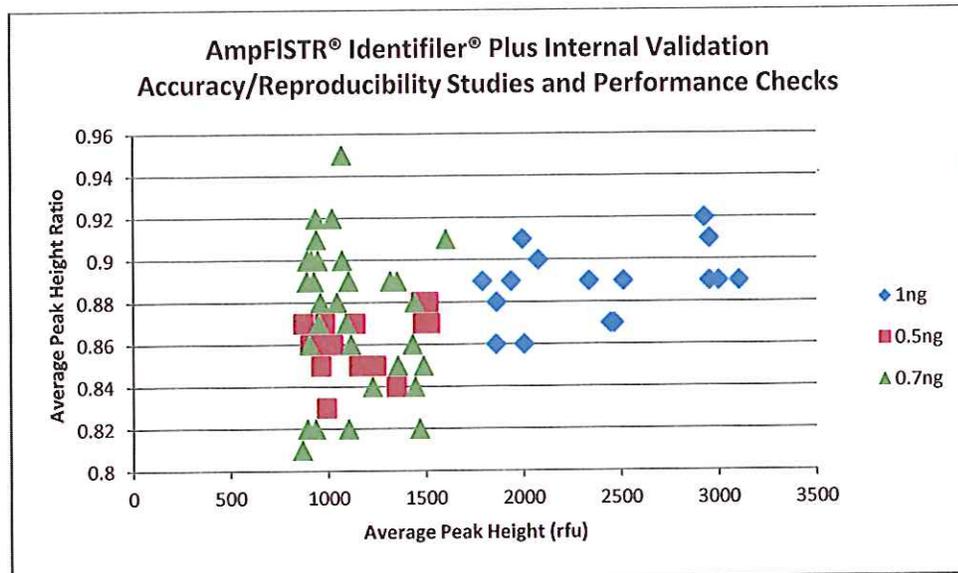
In order to reevaluate the peak height ratio recommended by the validation, additional samples from instrument performance checks were added to create a data set which not only covered different kit lot numbers but also different instruments.



In general, average peak height ratios were above 60% for all sensitivity samples; however the inclusion of standard deviations showed a significant decrease in peak height ratios at approximately 1000rfu. This demonstrates the stochastic effects which can occur in samples amplified at targets less than 250pg. At greater than 1000 rfu, the

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

average peak height ratio with standard deviations was 56% or greater. At peak heights below 1000 rfu, a single source sample may show peak imbalance at one or multiple loci.



For samples amplified in the ideal target range (0.5ng-1ng), all average peak height ratios were greater than 80%; however, with the inclusion of standard deviations, peak height ratios could be as low as 40% for samples with peak heights at or below 1000rfu. This is consistent with the data obtained from the sensitivity samples.

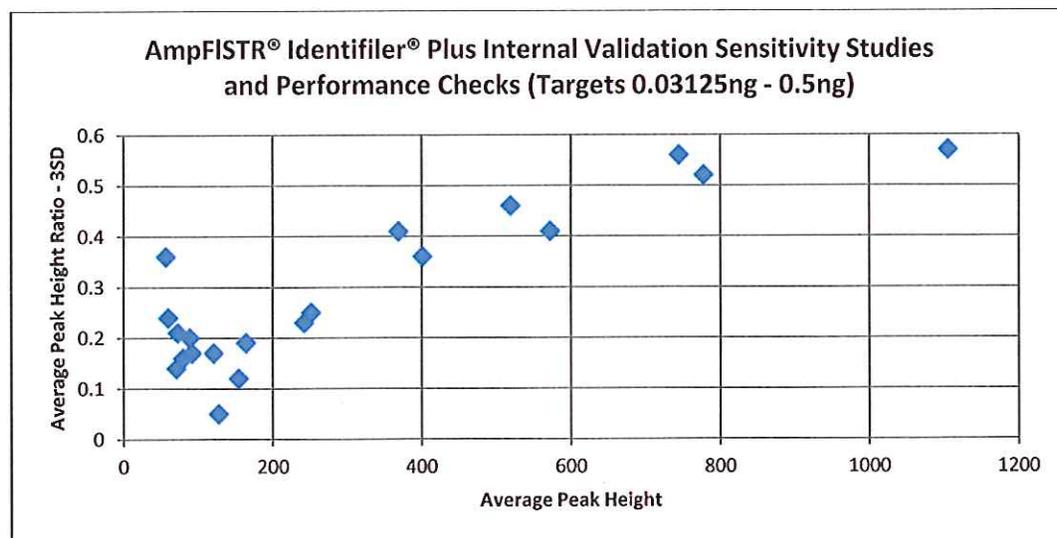
Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Based on this information, no adjustment to the current peak height ratio (55%) is recommended even if the analytical threshold is lowered. The additional peaks detected with a lower analytical threshold would only further support the variability observed at samples less than 1000 rfu.

c. Stochastic Threshold

The current stochastic threshold was determined by evaluating all the data and finding the alleles with the highest peak heights that were missing their heterozygous partner. A value of 200 rfu was recommended based on peaks that were observed at 196 rfu, 129 rfu, 215 rfu, and 229 rfu. This was done using the 70 rfu threshold determined by the validation.

Another method to determine stochastic threshold is to graphically evaluate the peak height ratios of various peak heights obtained in the sensitivity study of the internal validation. Using that information, the peak height at which peak height ratios may cause the heterozygous partner peak to fall below analytical threshold can be determined. The graph below shows data from the 0.03125 ng through 0.5 ng targets in the AmpFISTR® Identifiler® Plus Internal Validation and performance checks. While peak height ratios at 600 rfu fall below the 55% threshold, a heterozygous partner peak would still be expected to be above threshold ($600 \text{ rfu} \times 0.50 = 300 \text{ rfu}$). At approximately 200 rfu, the peak height ratio and average peak height are low enough that a heterozygous partner peak may not be above threshold ($200 \text{ rfu} \times 0.25 = 50 \text{ rfu}$). This indicates that the stochastic threshold with a 70 rfu analytical threshold may need to be adjusted. However, further evaluation with the newly proposed analytical thresholds demonstrated that a 200 rfu stochastic threshold is still appropriate (See Section IVc).

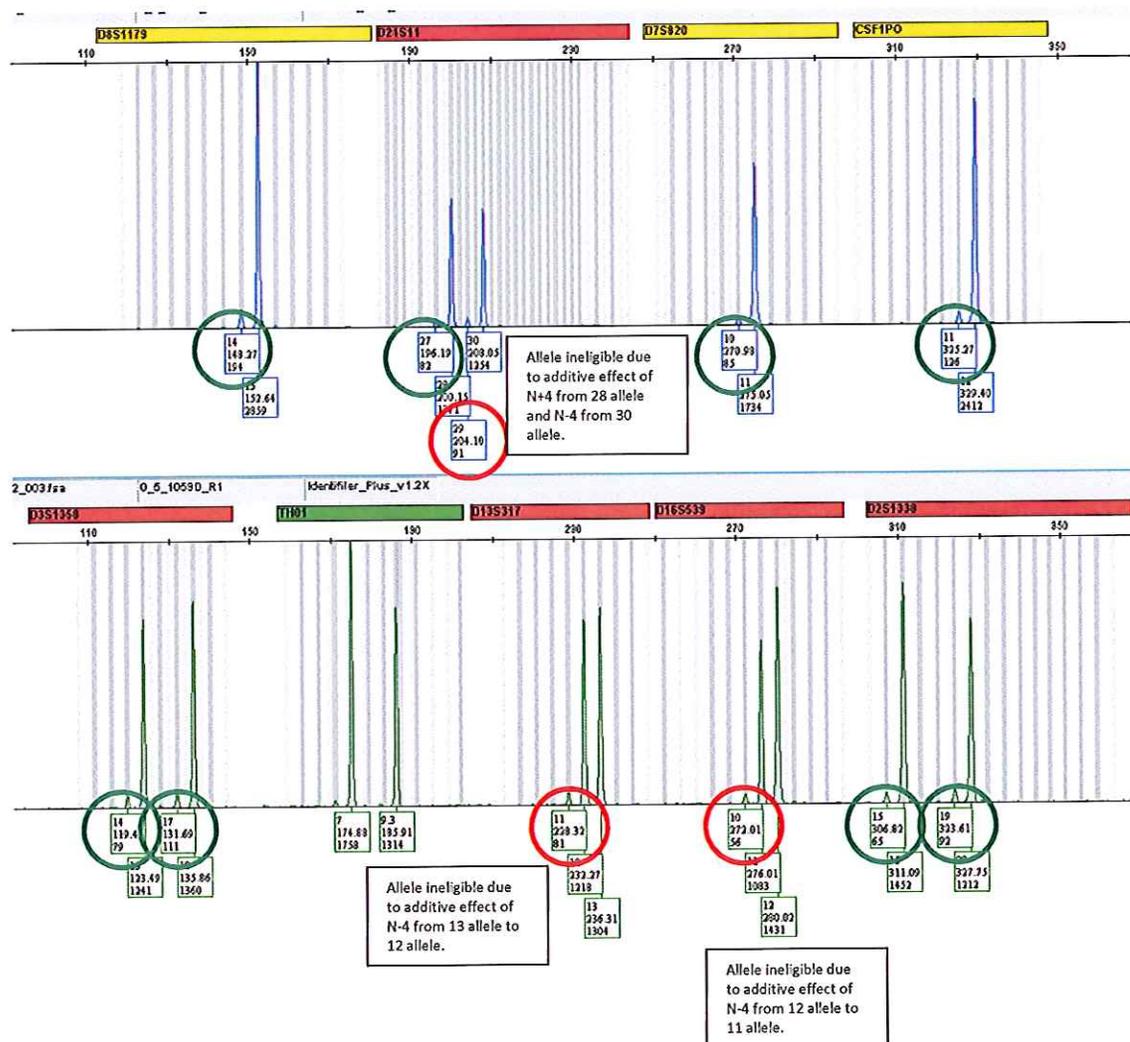


Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

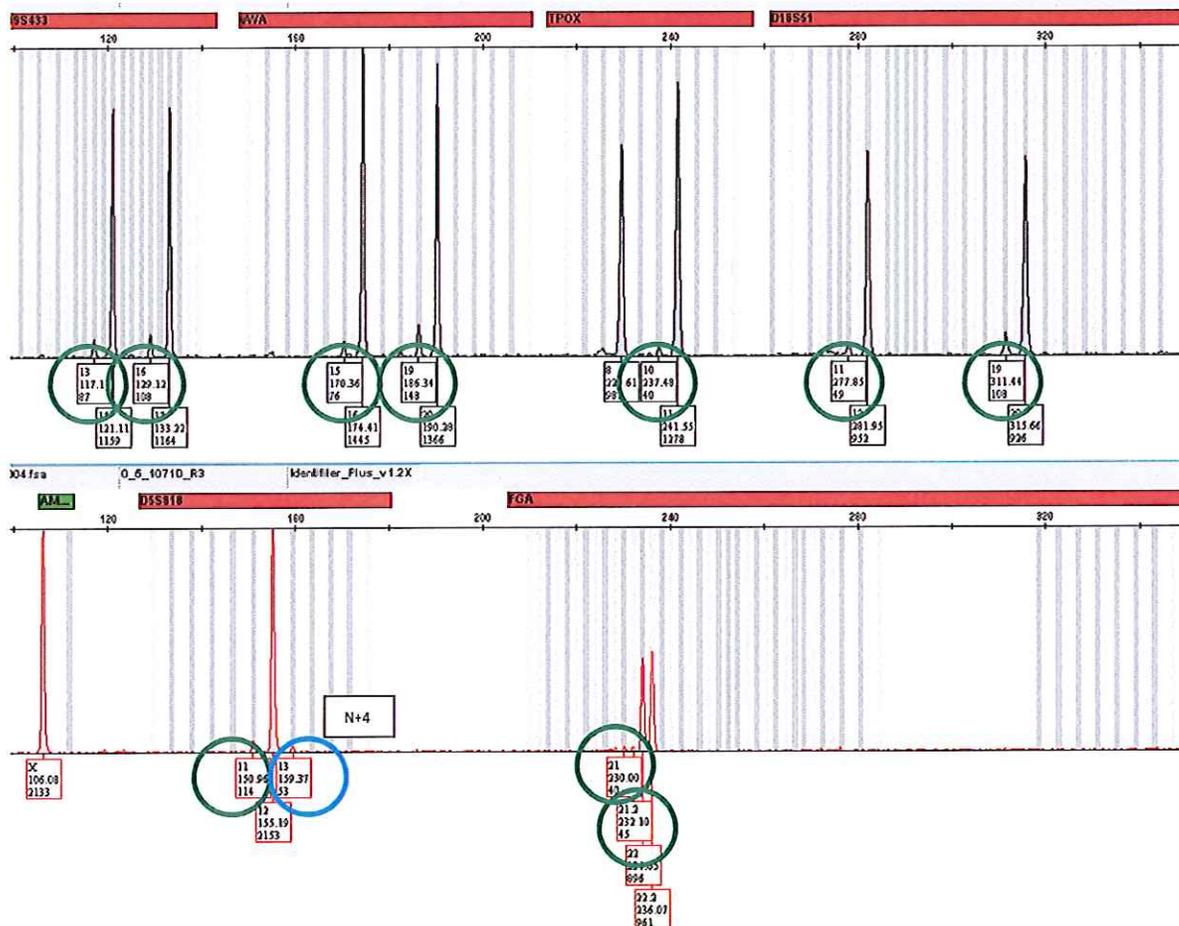
II. Stutter Study

During the AmpFISTR® Identifiler® Plus Internal Validation, a stutter study was not conducted. Discordant peaks were evaluated and it was determined that the manufacturer validated stutter percentages were sufficient for eliminating stutter while ensuring the detection of possible low level minor contributors. In an effort to further enhance the laboratory's analysis and interpretation of mixture and low level DNA profiles, a stutter study was completed and compared to the manufacturer's validation.

All single source samples used in the AmpFISTR® Identifiler® Plus Internal Validation were reanalyzed at 40 rfu with the stutter filter disabled. Forward and reverse stutter was calculated for each eligible allele and recorded in a Microsoft Excel spreadsheet. Eligible alleles (green circles) and ineligible alleles (red circles) were determined using the parameters in the following images:



Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation



Averages and standard deviations were then calculated for each locus using all the recorded percentages. An overall forward stutter average and standard deviation was also calculated. The results are listed in the following charts:

Locus	Average N-4 Stutter Percentage	Standard Deviation	Average + 3 Standard Deviations	Manufacturer Recommendations
D8S1179	5.92	1.52	10.48	10.32
D21S11	6.80	1.02	9.87	10.67
D7S820	4.12	1.27	7.93	9.69
CSF1PO	4.93	1.19	8.50	9.20
D3S1358	7.73	1.16	11.22	12.27
TH01	2.12	0.64	4.04	4.08
D13S317	5.40	1.22	9.07	9.93
D16S539	5.09	1.55	9.76	10.39
D2S1338	7.13	1.59	11.90	12.44
D19S433	7.22	1.42	11.47	11.21

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

vWA	6.47	1.39	10.63	12.45
TPOX	3.12	1.26	6.92	6.38
D18S51	7.63	1.99	13.60	13.68
D5S818	4.44	1.34	8.46	10.06
FGA	6.70	1.72	11.86	13.03

Overall, the reverse stutter (N-4) values are lower than the values published by the manufacturer; however the majority are within one percent. This may be due to the use of different data sets and/or analysis parameters used to conduct the study. The samples used from the AmpFISTR® Identifiler® Plus Internal Validation were not chosen anonymously and may not have contained as much allelic variability as the stutter study conducted by the manufacturer. In the subsequent studies of this reevaluation, the new stutter percentages will be applied and evaluated.

Locus	Average N+4 Stutter	Standard Deviation	Average + 3SD	Manufacturer Average + 3SD
D8S1179	1.10	0.36	2.18	1.64
D21S11	3.70	**	3.70	1.77
D7S820	*	*	*	1.94
CSF1PO	1.23	0.66	3.19	1.39
D3S1358	2.03	1.10	5.32	1.74
TH01	*	*	*	0.95
D13S317	2.37	1.78	7.72	1.38
D16S539	1.10	0.28	1.95	1.56
D2S1338	6.40	**	6.40	5.88
D19S433	*	*	*	5.09
vWA	4.81	1.67	9.83	3.42
TPOX	1.48	0.37	2.59	2.48
D18S51	6.33	2.96	15.20	2.15
D5S818	1.80	0.65	3.74	1.97
FGA	2.55	0.64	4.46	2.67
Overall	2.38	1.93	8.16	

* no instances of N+4 were observed at this locus

**only one instance of N+4 was observed at this locus

At the time of the AmpFISTR® Identifiler® Plus Internal Validation, no data had been published by the manufacturer regarding expected forward stutter (N+4) percentages. Recently, a technical note containing the above listed data was released⁸. The average forward stutter from this internal study is higher than the values published by the manufacturer, however, as with the reverse stutter assessment, different data sets and analysis parameters were used. The manufacturer study included significantly more instances of forward stutter due to a lower analytical threshold (10 rfu). Additional data may be required in order to establish a more reliable forward stutter filter.

Based on the manufacturer data, most loci would not require a forward stutter filter because the peaks would be present well below analytical threshold (1000-2000rfu x 2% - 20-40rfu). Using the values obtained from this study, occasionally a forward stutter peak may reach a level of up to 15%, especially at homozygous alleles, alleles with longer repeats, and/or alleles at the following loci: D13S317, vWA and D18S51. However, if the analytical threshold is adjusted, a forward stutter filter may be necessary for analysis and interpretation.

III. Different approaches for data analysis and profile interpretation

Using the data obtained in the above studies, six different methods of data analysis were developed for evaluation. Each was tested using the same data set to determine a best fit for the laboratory. This data set included samples from the AmpFISTR® Identifiler® Plus Internal Validation, instrument performance checks and samples from the NIST Interlab Mixture Studies from 2005 and 2013⁹. The electropherograms and exported Excel data from each method were evaluated to determine the number of concordant peaks obtained, number of discordant peaks obtained, causes of discordant peaks, appropriate peak height ratios and stochastic thresholds. Also considered for each method were the effects on data review, profile interpretation and analyst consistency.

All data was analyzed using GeneMapper ID-X v1.4 with the analysis thresholds determined in Section I above and the reverse stutter percentages calculated in Section II above. Forward stutter percentages were not applied unless noted.

- a. Method A: SWGDAM analytical thresholds derived from samples which contain DNA were rounded to the nearest ten and used to analyze the sample set. Forward stutter percentages calculated in Section IV of this study were also applied to the data.

Analytical Threshold (rfu)

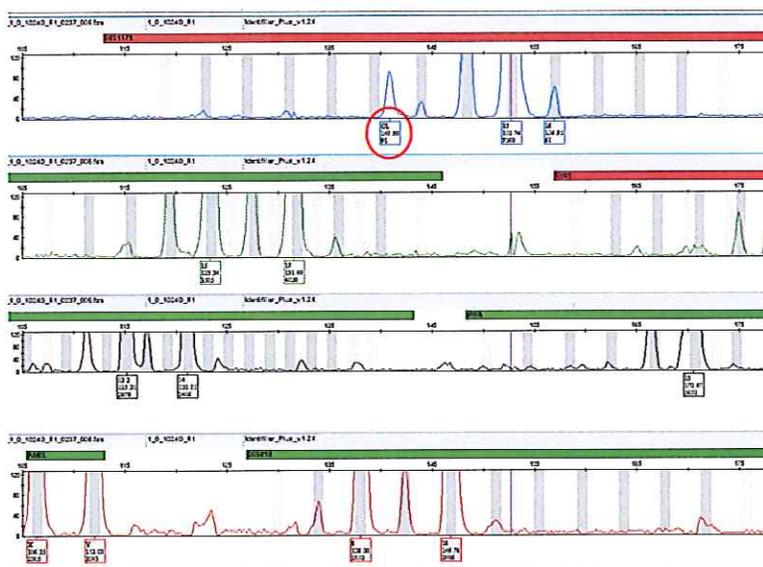
Blue:	60
Green:	50
Yellow:	70
Red:	60
Orange:	70

Results

Total Number of alleles	4936
Total Number of alleles < 70 rfu	61
Total Number of OL/non-matching	126
Total Number < 70 rfu	54
Total Number of pull up	99
Total Number of background	15

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Total Number of Minus A	0
Total Number of elevated stutter	7
Total Number of N+4	4
Total Number of Off-scale	0
Total Number of shoulder	0
Total Number of unknown	1



This method used the highest of all the suggested analytical thresholds and was the most similar to the current method used by the laboratory. However, the application of a dye-specific analytical threshold resulted in the detection of 61 additional concordant alleles while adding 54 additional discordant alleles. The only discordant allele whose origin could not be determined was at 91 rfu and was present at a locus already labeled as off-scale. This allele would have also been detected using the laboratory’s current analytical threshold. The locus, under both sets of parameters, would have been interpreted with caution and the sample possibly re-amplified. All other discordant alleles were easily identified as pull up, background or stutter. NIST Mixture 2005 and 2013 samples were not included in these calculations because they were not amplified or typed using this laboratory’s instrumentation. They will be evaluated and assessed in a separate section of this study.

Due to the similarity in analytical threshold to the laboratory’s current method of data analysis, there is no recommendation to change the peak height ratio. However, based on the data obtained in the stochastic threshold study above, an adjustment to 250 rfu may be a more appropriate stochastic threshold.

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

This method may allow a slight increase in sensitivity without adding a significant amount of time determining whether a low level peak is an artifact or true allele. The slight increase in sensitivity may allow the detection of additional low level peaks in single source and mixture samples, especially at loci in the green dye channel. The slight increase in stochastic threshold, if determined to be appropriate, may also allow a more conservative approach to conclusions and statistical calculations for samples with low level minor contributors.

- b. Method B: Limit of Quantitation (LOQ) analytical thresholds derived from samples which contain DNA were rounded to the nearest ten and used to analyze the sample set.

Analytical Threshold (rfu)

Blue:	20
Green:	20
Yellow:	40
Red:	40
Orange:	70

Results

Total Number of alleles	5075
Total Number of alleles < 70 rfu	208
Total Number of OL/non-matching	937
Total Number < 70 rfu	858
Total Number of pull up	620
Total Number of background	171
Total Number of Minus A	3
Total Number of elevated stutter	23
Total Number of N+4	100
Total Number of Off-scale	0
Total Number of shoulder	10
Total Number of unknown	10

This method dropped the analytical threshold significantly from the laboratory's current method of analysis. While an additional 208 concordant alleles were detected, 858 additional discordant alleles were also detected. Most discordant alleles were from pull up and were not an indication of a poor matrix. The samples analyzed for these studies were all run within a short time period of a new spectral calibration and were within the accepted 4% value suggested by the manufacturer. A significant number of forward

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

stutter peaks were identified indicating that this method may require the use of a forward stutter filter. 10 discordant peaks were unable to be identified indicating the analytical threshold may not be sufficient for eliminating unreliable data.

A change in peak height ratio and stochastic threshold may be necessary due to the increased number of peaks detected at low levels. If this method is chosen, a further study using the data from the amplification cut-off study conducted in 2014 could be used to determine these new values at low level DNA concentrations.

This method increases sensitivity significantly, however introduces a more complicated set of interpretation guidelines. Analyst judgment would often be required in order to determine the validity of a called peak and multiple samples would need re-injected, re-run or re-amplified for confirmation. An increased amount of forward stutter may also cause difficulty in assessing mixtures with low level contributors.

- c. Method C: SWGDAM analytical thresholds derived from samples which contain DNA were used for the samples which contain at least one allele at or above 1000rfu. For samples in which all alleles are below 1000 rfu, samples were reanalyzed using the SWGDAM analytical thresholds derived from negative samples. Forward stutter percentages calculated in Section IV of this study were also applied to the data.

Samples with at least one allele >1000rfu

Analytical Threshold (rfu)

Blue:	60
Green:	50
Yellow:	70
Red:	60
Orange:	70

Samples with all alleles <1000rfu

Analytical Threshold (rfu)

Blue:	20
Green:	50*(see Study I above)
Yellow:	40
Red:	40
Orange:	70

Results

Total Number of alleles	4972
Total Number of alleles < 70 rfu	97
Total Number of OL/non-matching	127

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Total Number < 70 rfu	55
Total Number of pull up	99
Total Number of background	15
Total Number of Minus A	0
Total Number of elevated stutter	8
Total Number of N+4	4
Total Number of Off-scale	0
Total Number of shoulder	0
Total Number of unknown	1

Based on the data obtained from the peak height ratio and stochastic threshold studies above, a method was suggested in which two separate sets of analysis parameters were applied depending on peak heights. According to the study, stochastic effects caused peak height ratios to begin to vary at approximately 1000 rfu. In addition, it was observed in the analytical threshold study that negative samples produced lower overall baseline noise than samples which contained ideal quantities of DNA. Therefore, sample results at or above 1000 rfu were considered ideal and a higher threshold applied to eliminate common artifacts and unreliable data. Samples with all results below 1000 rfu were analyzed with the lower threshold calculated from the negative samples to ensure low level data was still reliably detected.

This method added 36 concordant peaks that were previously not detected with Method A. Only one additional discordant peak was added which was attributed to elevated stutter. Like Method A, no change in peak height ratio is needed for data with peaks greater than 1000 rfu. Additionally, no change in stochastic threshold would be needed since samples with peak heights below 1000 rfu would be analyzed with the lower analytical threshold. However, an additional study may be necessary using the data obtained in the amplification cutoff study conducted in 2014 to determine if different peak height ratio and stochastic thresholds should be used for data under 1000 rfu.

This method provides better sensitivity without the addition of possible unreliable data. The slight increase in sensitivity may allow the detection of additional low level peaks in single source and mixture samples.

- d. Method D: LOQ analytical thresholds derived from samples which contain DNA were used for all samples which contain at least one allele at or above 1000 rfu. For samples in which all alleles are below 1000 rfu, samples were reanalyzed using the Limit of Detection (LOD) derived from samples which contain DNA.

Samples with at least one allele >1000rfu

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Analytical Threshold (rfu)

Blue:	20
Green:	20
Yellow:	40
Red:	40
Orange:	70

Samples with all alleles >1000rfu

Analytical Threshold (rfu)

Blue:	10
Green:	10
Yellow:	20
Red:	20
Orange:	70

Results

Total Number of alleles	5088
Total Number of alleles < 70 rfu	221
Total Number of OL/non-matching	944
Total Number < 70 rfu	865
Total Number of pull up	621
Total Number of background	172
Total Number of Minus A	3
Total Number of elevated stutter	26
Total Number of N+4	101
Total Number of Off-scale	0
Total Number of shoulder	10
Total Number of unknown	11

For this method an additional 13 concordant alleles were detected and seven additional discordant alleles were detected from Method B. As with Method B, most discordant alleles were due to pull up and an additional peak of unknown origin was detected, indicating the analytical thresholds in this method may not be sufficient for eliminating unreliable data.

A change in peak height ratio and stochastic threshold may be necessary for both the above and below 1000 rfu analysis parameters due to the increased number of peaks detected at low levels. If this method is chosen, a further study using the data from the

amplification cut-off study conducted in 2014 could be used to determine these new values at low level DNA concentrations.

This method increases sensitivity significantly, however introduces a more complicated set of interpretation guidelines. Analyst judgment would often be required in order to determine the validity of a called peak and multiple samples would need re-injected, re-run or re-amplified for confirmation. An increased amount of forward stutter may also cause difficulty in assessing mixtures with low level contributors. The samples in this study were difficult to assess and required significantly more time to analyze. For interpretation, many single source samples appeared as mixtures and two-person mixtures appeared as three or more individuals.

- e. Method E: SWGDAM analytical thresholds derived from samples which contain DNA were used for all samples. If a sample appeared to contain signals below analytical threshold, LOQ derived from samples which contain DNA was used to reanalyze. The data obtained from the lower threshold was only used to determine the number of contributors and/or exclude an individual.

All Samples

Analytical Threshold (rfu)

Blue:	60
Green:	50
Yellow:	70
Red:	60
Orange:	70

Designated samples with possible alleles below threshold

Contributor Informative Threshold (rfu)

Blue:	20
Green:	20
Yellow:	40
Red:	40
Orange:	70

Using this method would require a determination of the reliability of the additional information obtained when analyzed at the Contributor Informative Threshold (CIT). If the data is very reliable, then the information may not only be used to determine the number of contributors in a sample, but also whether an individual may be excluded as a possible contributor. If the data is less reliable, then the information may only be used to determine possible number of contributors.

Additionally, this method would require that the analysts examining the data have consistent opinions regarding when the use of the CIT should be employed. To assess the usability of Method E, three analysts were handed the results from Method A and asked to indicate which samples should be reanalyzed at the CIT. Two analysts performed their review using the printed electropherograms to mimic technical review and one analyst reviewed the data electronically to mimic a second read.

Each analyst returned with a different number and set of samples (9, 51 and 16). In addition, only nine samples were consistent between two individuals. This indicates that Method E may be subjective to the method used for assessment and the individual's training and experience. This could lead to inconsistency within the unit. This method will not be further assessed as an option for data analysis.

- f. Method F: LOQ analytical thresholds derived from samples which contain DNA were used for all samples. If a sample appeared to contain signals below analytical threshold, LOD derived from samples which contain DNA was used to reanalyze. The data obtained from the lower threshold was only used to determine the number or contributors and/or exclude an individual.

All Samples

Analytical Threshold (rfu)

Blue:	20
Green:	20
Yellow:	40
Red:	40
Orange:	70

Designated samples with possible alleles below threshold

Contributor Informative Threshold (rfu)

Blue:	10
Green:	10
Yellow:	20
Red:	20
Orange:	70

Using this method, like Method E, would require a determination of the reliability of the additional information obtained when analyzed at the Contributor Informative Threshold (CIT). If the data is very reliable, then the information may not only be used to determine the number of contributors in a sample, but also whether an individual may be excluded as a possible contributor. If the data is less reliable, then the information may only be used to determine possible number of contributors.

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Additionally, this method would require that the analysts examining the data have consistent opinions regarding when the use of the CIT should be employed. To assess the usability of Method F, three analysts were handed the results from Method B and asked to indicate which samples should be reanalyzed at the CIT. Two analysts performed their review using the printed electropherograms to mimic technical review and one analyst reviewed the data electronically to mimic a second read.

As with Method E, each analyst returned with a different number and set of samples (42, 61 and 1). In addition, only 26 samples were consistent between two individuals. This indicates that Method F may be subjective to the method used for assessment and the individual’s training and experience. This could lead to inconsistency within the unit. This method will not be further assessed as an option for data analysis.

Method Summary Table

		Method A	Method B	Method C	Method D
Initial Analytical threshold	B	60	20	60	20
	G	50	20	50	20
	Y	70	40	70	40
	R	60	40	60	40
<1000rfu Analytical threshold	B			20	10
	G			50	10
	Y			40	20
	R			40	20
Number of Concordant Peaks		4936	5075	4972	5088
Number of Discordant Peaks		126	937	127	944
Number of N+4		4	100	4	101
Number of Unknown		1*	10	1*	11

*locus was off-scale

IV. Selection and Verification of Analysis Method

Once data was compiled for the above listed studies, an initial review was conducted by the author of this project, the Forensic Biology Unit Technical Leader and FSL Quality Manager. The results of each method were discussed and while Method C was identified as the most appropriate for analysis and interpretation, the additional steps below were recommended prior to a final decision and implementation.

- a. Method G: SWGDAM analytical thresholds derived from negative samples were rounded to the nearest ten and used to analyze the sample set.

Analytical Threshold (rfu)
 Blue: 20
 Green: 50
 Yellow: 40
 Red: 40
 Orange: 70

Results

Total Number of alleles	5041
Total Number of alleles < 70 rfu	173
Total Number of OL/non-matching	604
Total Number < 70 rfu	525
Total Number of pull up	430
Total Number of background	85
Total Number of Minus A	2
Total Number of elevated stutter	13
Total Number of N+4	63
Total Number of Off-scale	0
Total Number of shoulder	7
Total Number of unknown	4

This method dropped the analytical threshold significantly from Method A in the blue, yellow and red dye channels. While an additional 105 concordant alleles were detected, 480 additional discordant alleles were also detected. Most discordant alleles were from pull up and were not an indication of a poor matrix. The samples analyzed for these studies were all run within a short time period of a new spectral calibration and were within the accepted 4% value suggested by the manufacturer. Less forward stutter peaks were identified than Method B however the use of a forward stutter filter may still be necessary. Four discordant peaks were unable to be identified indicating the analytical threshold may not be sufficient for eliminating unreliable data.

A change in peak height ratio and stochastic threshold may be necessary due to the increased number of peaks detected at low levels. If this method is chosen, a further study using the data from the amplification cut-off study conducted in 2014 could be used to determine these new values at low level DNA concentrations.

This method increased sensitivity significantly, however may introduce a more complicated set of interpretation guidelines. Analyst judgment would often be required in order to determine the validity of a called peak and multiple samples would need re-injected, re-run or re-amplified for confirmation. An increased amount of forward stutter may also cause difficulty in assessing mixtures with low level contributors.

Once data was compiled for this Method, an additional review was conducted by the author of this project, the Forensic Biology Unit Technical Leader and the FSL Quality

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

Manager. It was determined that Method C is the most appropriate method for data analysis.

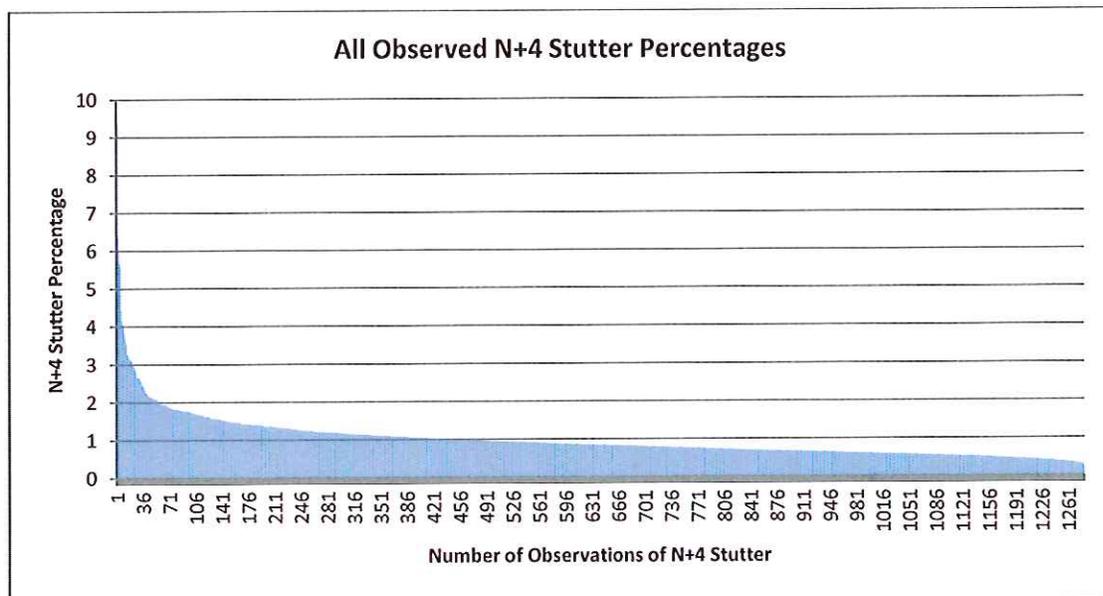
b. Additional N+4 Evaluation

In Section II of this reevaluation, it was determined that there was not enough data to establish an appropriate forward stutter filter. All the samples used for the stutter study were reanalyzed with a lower threshold of 10 rfu. Each peak in the forward stutter position was evaluated and eliminated from calculations if possible background or pull up may be influencing peak height. The results of the re-analysis are below.

Locus	Average N+4 Stutter Percentage	Standard Deviation	Average + 3 SD	Manufacturer Average + 3 SD
D8S1179	0.80	0.34	1.81	1.64
D21S11	1.02	0.61	2.84	1.77
D7S820	0.99	0.51	2.52	1.94
CSF1PO	0.93	0.58	2.67	1.39
D3S1358	0.86	0.42	2.13	1.74
TH01	0.75	0.47	2.17	0.95
D13S317	0.90	0.70	3.01	1.38
D16S539	0.76	0.35	1.82	1.56
D2S1338	1.42	1.11	4.74	5.88
D19S433	1.26	0.55	2.90	5.09
vWA	0.98	0.62	2.85	3.42
TPOX	1.04	0.41	2.27	2.48
D18S51	1.25	1.13	4.64	2.15
D5S818	1.16	0.50	2.66	1.97
FGA	1.29	1.08	4.54	2.67
OVERALL	0.99	0.67	3.02	

At the time of the AmpFISTR® Identifiler® Plus Internal Validation, no data had been released by the manufacturer regarding expected forward stutter (N+4) percentages. Recently a technical note containing the above data was released. For most loci, the average forward stutter from this internal study was higher than the values listed by the manufacturer, however, as with the reverse stutter assessment, different data sets and analysis parameters were used. Despite the differences in average forward stutter per locus, both this study and the manufacturer study demonstrated that the majority of forward stutter is less than 2% across all loci.

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation



It is important to note that the current analytical threshold would not require a forward stutter filter for samples amplified in an appropriate range of 1000-2000rfu. Forward stutter would be present at levels under analytical threshold (20-40 rfu). However, recent mixture interpretation studies have suggested the importance of factoring in stutter when determining the peak height of a possible contributor(s).

In addition, there were also instances of forward stutter which significantly exceeded their expectation (greater than 9%). This may be sample specific, however, and should always be considered when examining profiles with possible low level contributors.

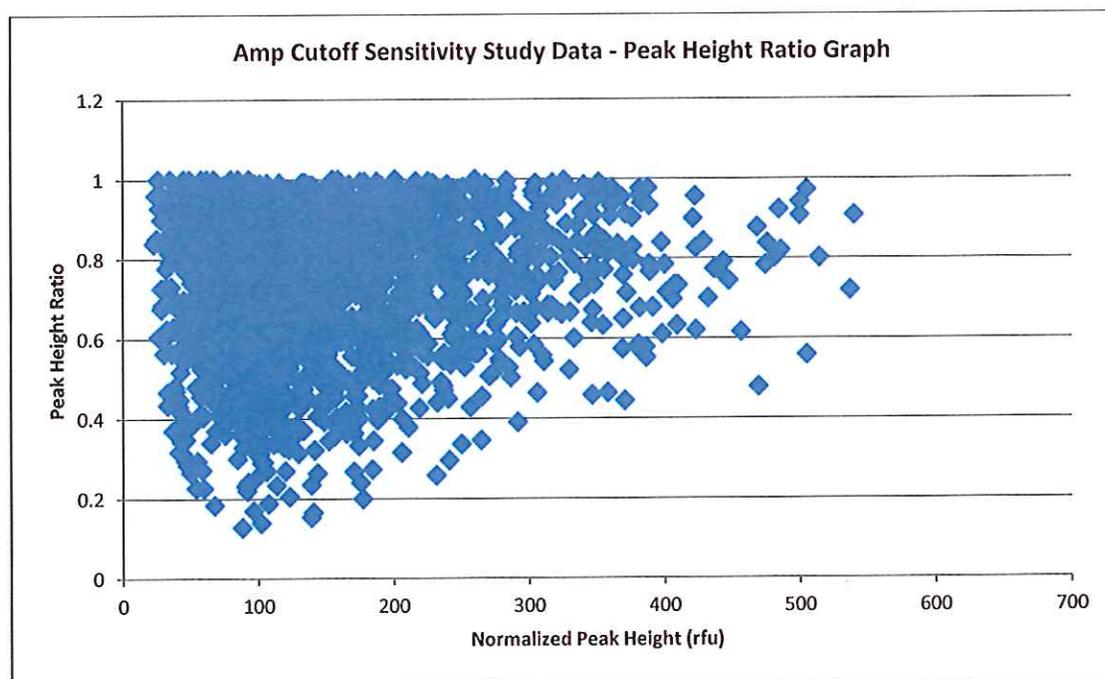
c. Amplification Cutoff Samples

In previous studies of this project, recommendations were made to verify stochastic threshold and peak height ratio if the analytical threshold is adjusted. For Method C, a sufficient number of samples greater than 1000 rfu were evaluated in the validation samples in the first section of this report, however more samples less than 1000 rfu were needed to properly assess the values.

In 2014, a study was conducted to assess an appropriate DNA quantity at which amplification would not result in a useful profile¹⁰. The sensitivity section of this study contained 236 single source samples, all but one under 1000 rfu. These samples were analyzed using the parameters listed in Method C and then evaluated to determine the highest allele missing its heterozygous partner. Two instances of sister allele dropout were observed at 173 rfu. This indicated the stochastic threshold of 200rfu was appropriate for both samples above and below 1000 rfu.

Reevaluation of AmpFISTR® Identifiler® Plus Internal Validation

For peak height ratio, the graph below indicates the peak height ratios that were observed in the amplification cutoff sensitivity samples. 82.5% of loci produced peak height ratios greater than or equal to 55%. Of the loci with peak height ratios below 55%, the majority were also below the 200 rfu stochastic threshold. This indicates that the overall peak height ratios of samples under 1000 rfu will be 55% or greater, however more variability will be observed. This is consistent with the previous peak height ratio study and validation.



It should be noted that when concordance was assessed for the samples in this study, a discordant allele was observed at 22rfu at D8S1179. The sample was amplified at a level below the laboratory's current amplification cutoff value and produced only one other allele which was determined to be concordant. In casework, this profile would be reported as inconclusive and not be suitable for comparison.

d. NIST Mixture 2005

Method A produced 303 concordant alleles, three of which would not have been previously detected using the 70rfu analytical threshold. The total number of discordant alleles was unable to be properly evaluated due to off scale data.

Method B produced 306 concordant alleles, six of which would not have been previously detected using the 70rfu analytical threshold. One of these alleles was in a sample which was a mixture with a locus that included a heterozygous contributor and a tri-allele contributor. The tri-allele contributor was not completely detected in Method

A. The total number of discordant alleles was again unable to be properly evaluated due to off scale data.

Methods C and D do not apply to the NIST Mixture 2005 study because none of the mixtures contained all peaks below 1000rfu.

Methods E and F were not further evaluated due to the information obtained in the above study.

For this mixture set, it is difficult to determine the best method for analysis due to a significant amount of off scale data.

e. NIST Mixture 2013

Method A

Total Number of alleles	244
Total Number of alleles < 70 rfu no discordant alleles observed	6

Method B

Total Number of alleles	250
Total Number of alleles < 70 rfu	12
Total Number of OL/non-matching (1 pull up, 1 background)	2
Total Number < 70 rfu	2

Method C

One sample was eligible to be reanalyzed. This sample produced six additional concordant alleles and no discordant alleles.

Method D

One sample was eligible to be reanalyzed. No additional concordant or discordant alleles were detected for this sample.

Methods E and F were not further evaluated due to the information obtained in the above study.

For this mixture set, Method C resulted in the most number of detected alleles without the addition of any artifacts.

Conclusions

In the first section of this reevaluation, the thresholds recommended in the AmpFISTR® Identifiler® Plus Internal Validation were re-calculated using both historically used methods and recently published methods for analytical threshold, peak height ratio and stochastic threshold. In addition, Section II

provided calculated stutter percentages based on data generated by this laboratory's procedures and instrumentation. Each of these values was then assessed and a set of methods for data analysis were tested. Each method was evaluated based on quantity and quality of results.

All the methods added concordant alleles which were previously not detected by the originally recommended 70 rfu threshold. Methods A and C were both able to do this without a major increase in the number of identifiable artifacts. However, Methods B and D introduced a significant number of discordant alleles, some of which were unable to be identified as an artifact. Both of these methods were not further evaluated due to the need to sort through the data and determine whether a called peak is an unknown artifact or true allele. This would lead to more samples which would need re-injected, re-run and/or re-amplified for confirmation and the possibility of inconsistency between analysts.

Method C was developed and tested to determine whether low level samples could reliably be analyzed with a different set of parameters. At approximately 1000 rfu, peak height ratios began to vary significantly and the lower set of analytical thresholds was able to include more concordant alleles than Method A without introducing a significant number of discordant alleles. Additional studies were conducted to verify the forward stutter, peak height ratio and analytical thresholds for Method C.

Methods E and F suggest an additional threshold, the contributor informative threshold (CIT), be used when samples may contain alleles below the analytical threshold. While this method may add to the amount of information an analyst uses for conclusions and statistical calculations, variability between analysts may lead to inconsistency in use. Determining whether an allele is possibly present below threshold can vary not only with the method being used for evaluation, electronic or paper, but also the display of the plot settings (size, scale, labeling, etc.). Methods E and F were determined to be unsuitable for use.

In conclusion, Method C is recommended for data analysis and interpretation using the following analysis parameters and interpretation thresholds:

Samples with at least one allele >1000rfu
Analytical Threshold (rfu)
Blue: 60
Green: 50
Yellow: 70
Red: 60
Orange: 70

Samples with all alleles <1000rfu
Analytical Threshold (rfu)
Blue: 20
Green: 50*(see Study I above)
Yellow: 40

Red: 40
Orange: 70

Reverse Stutter Filter = values stated in Section II
Forward Stutter Filter = values stated in Section IVb
Peak Height Ratio \geq 55%
Stochastic Threshold \geq 200 rfu

Additional notes:

All analysis for this reevaluation was conducted using GMID-X Version 1.4. Please refer to the appropriate performance check to verify results expected from GMID-X Version 1.3¹¹.

During review, all forward and reverse stutter calculations were verified. The charts listed in this document and the supporting excel files reflect any slight changes in final values. No data was reanalyzed or updated for Sections III and IV.

All fsa files, GMID-X projects, analysis methods, Excel spreadsheets, Notepad documents and references will be maintained as electronic copies on CD after completion of technical review.

Appendix A – electropherograms from Section II – Stutter Study

Appendix B – electropherograms from Section III – Methods A-D

Appendix C – electropherograms from Section IVb – N+4 Study

Appendix D – electropherograms from Section IVc – Amp Cutoff Samples

¹ Letter to Dr. Max M. Houck regarding DNA Mixture Interpretation Recommendations from Irvin B. Litofsky, November 5, 2014.

² Eschinger, Lesley, Kolowski, Jason, Skillman, Jessica. Internal Validation of AmpFISTR®Identifiler® Plus PCR Amplification Kit. District of Columbia, Metropolitan Police Department, Forensic Biology Unit, September 13, 2012.

³ Various Performance Check binders from 2012, 2013, 2014. District of Columbia, Department of Forensic Science, Forensic Science Laboratory, Forensic Biology Unit. Various dates, reviewers and authors.

⁴ Analytical Threshold and Sensitivity: Establishing RFU Threshold for Forensic DNA Analysis. Bregu, J, Conklin D, Coronado E, Terrill M, Cotton R, Grgicak C. Journal of Forensic Science, January 2013, Vol. 58, No. 1.

⁵ Scientific Working Group on DNA Analysis Methods Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. http://swgdam.org/Interpretation_Guidelines_January_2010.pdf, approved 01/14/2014.

⁶ Grgicak, C. Analytical Thresholds: Determination of Minimum Distinguishable Signals presentation. 21st International Symposium on Human Identification, Mixture Interpretation Workshop: Principles, Protocols, and Practice. October 11, 2010. San Antonio, TX. <http://www.cstl.nist.gov/strbase/mixture/3%20-%20Analytical%20Threshold.pdf>

⁷ McCord, B. DNA Typing and Threshold Setting. <http://dna.fiu.edu/Advanced%20DNA%20Typing%20lectures/DNA%20typing%20and%20threshold%20setting%20of%20genetics%20class3%20.pdf>

⁸ Technical Note: Considerations for the Evaluation of Plus Stutter for AmpFISTR® PCR Amplification Kits in Human Identification Laboratories (Version2, March 2014). Thermo Fisher Scientific/Life Technologies.

⁹ NIST Interlab Mixture Studies, 2005 and 2013. <http://www.cstl.nist.gov/strbase/>

¹⁰ Skillman, Jessica, Zeffer, Jennifer. Amplification Cutoff Validation. District of Columbia, Department of Forensic Science, Forensic Science Laboratory, Forensic Biology Unit. To be released after technical review.

¹¹ Skillman, Jessica, Zeffer, Jennifer. Performance Check GeneMapper ID-X Version 1.4. District of Columbia, Department of Forensic Science, Forensic Science Laboratory, Forensic Biology Unit. 02/01/2015.