

NGS Laboratory Bacterial Validation Plan

Eric L Vaughn, DrPH, MPH

DC Department of Forensic Sciences – Public Health Laboratory

APHL ID Con, Pasadena, CA

March 25, 2025

Disclaimer

- The views and positions expressed in this presentation do not constitute an endorsement for any product mentioned.

Purpose

- Validate 4 independent bacterial analysis platforms and methods
 - BioNumerics v7.6
 - In-house pipeline
 - CDC's Phoenix
 - Commercial software

Process

- Determine the methods
- Determine the total number of samples
- Determine the comparator method and results
- Perform Accuracy - sensitivity and specificity (S/S) for all 4 groups
 - Test all isolates
- Perform repeatability testing
 - Re-test all isolates
 - Compare to sensitivity and specificity
- Perform reproducibility testing
 - Re-test a subset of isolates
 - Compare to repeatability

Sample Selection and Breakdown - BioNumerics

- Isolates broken up into 3 groups:
 - Group 1 – True Positives (TP):
 - Verify for Organism (Genus and species) and Serotype.
 - Comparator Method: CDC EDLB QA Panel results
 - Group 2 – TP:
 - BioNumerics only: Verify for Organism (Genus and species), serotype, quality Q-30 score, contamination, coverage, sequence length, core percent and relatedness.
 - Comparator Method: PulseNet proficiency and certification test results
 - Group 3 – True Negatives (TN):
 - 10 Isolates that failed PulseNet QC metrics

Sample Selection and Breakdown – All other methods

- Isolates broken up into 3 groups:
 - Group 1 – TP:
 - Verify for Organism (Genus and species), serotype, Q-30 score, contamination, coverage, sequence length, core percent and relatedness.
 - Comparator Method: PulseNet proficiency and certification test results.*
 - Group 2 – TP:
 - Verify for Organism (Genus and species), serotype, and antimicrobial resistance markers (AMR).
 - Comparator Method: Isolates pulled from NCBI, ATCC, and other repositories.
 - Group 3 – TN:
 - 10 Isolates that failed PulseNet QC metrics*

Total testing numbers by group

	BioNumerics	In-House	Phoenix	Commercial
Group 1 – TP	38	20	20	8
Group 2 – TP	37	50	50	8
Group 3 – TN	10	10	10	3
Total	85	80	80	19

Total testing throughout the validation

	BioNumerics	In-House	Phoenix	Commercial
Accuracy (S/S) – TP	75	70	70	16
Accuracy (S/S) – TN	10	10	10	3
Repeatability – TP	75	70	70	16
Repeatability – TN	10	10	10	3
Reproducibility – TP	20	20	20	16
Reproducibility – TN	5	5	5	3
Total	195	185	185	57

Acceptable Criteria

- All methods must have a value 90% or more of the results in agreement based on the following formulas:
 - Sensitivity: $(TP / (TP + FN)) \times 100$
 - Specificity: $(TN / (FP + TN)) \times 100$
 - Repeatability: $((TP + TN) / (TP + TN + FP + FN)) \times 100$
 - Reproducibility: $(\# \text{ of results in agreement} / \text{total } \# \text{ of results}) \times 100$

Results

BioNumerics Calculations			
Test	Number	Test	Number
Results in Agreement		Not in agreement	
S & S	85	S & S	0
A/R	85	A/R	0
P/R	25	P/R	0
Total	195	Total	0
Total # of results			195

Phoenix Calculations			
Test	Number	Test	Number
Results in Agreement		Not in agreement	
S & S	79	S & S	1
A/R	79	A/R	1
P/R	23	P/R	2
Total	181	Total	4
Total # of results			185

BioNumerics Calculations				
Results Values				
# of results in agreement			195	
Total # of results			195	
+/- Values				
TP	75	FP	0	
TN	10	FN	0	
Total Calculations				Passing (Y/N)
Accuracy - Sensitivity	100.0	%		Yes
Accuracy - Specificity	100.0	%		Yes
Accuracy and Repeatability	100.0	%		Yes
Precision and Reproducibility	100.0	%		Yes

Phoenix Calculations				
Results Values				
# of results in agreement			181	
Total # of results			185	
+/- Values				
TP	70	FP	0	
TN	10	FN	4	
Total Calculations				Passing (Y/N)
Accuracy - Sensitivity	94.6	%		Yes
Accuracy - Specificity	100.0	%		Yes
Accuracy and Repeatability	95.2	%		Yes
Precision and Reproducibility	97.8	%		Yes

Results

Commercial Calculations			
Test	Number	Test	Number
Results in Agreement		Not in agreement	
S & S	19	S & S	0
A/R	19	A/R	0
P/R	19	P/R	0
Total	57	Total	0
Total # of results			57

In-House Calculations			
Test	Number	Test	Number
Results in Agreement		Not in agreement	
S & S	48	S & S	32
A/R	48	A/R	32
P/R	18	P/R	7
Total	114	Total	71
Total # of results			185

Commercial Calculations				
	Results Values			
	# of results in agreement			57
	Total # of results			57
	+/- Values			
	TP	16	FP	0
	TN	3	FN	0
Total Calculations				Passing (Y/N)
Accuracy - Sensitivity		100.0	%	Yes
Accuracy - Specificity		100.0	%	Yes
Accuracy and Repeatability		100.0	%	Yes
Precision and Reproducibility		100.0	%	Yes

In-House Calculations				
	Results Values			
	# of results in agreement			114
	Total # of results			185
	+/- Values			
	TP	70	FP	24
	TN	10	FN	28
Total Calculations				Passing (Y/N)
Accuracy - Sensitivity		71.4	%	No
Accuracy - Specificity		29.4	%	No
Accuracy and Repeatability		60.6	%	No
Precision and Reproducibility		61.6	%	No

Validation Plan Summary

- BioNumerics, Phoenix, and the commercial software performed as expected.
 - All will be used for official analysis of bacterial isolates
- In-house did not perform as expected
 - More reference genomes needed
 - May be using outdated code
 - Will only be used for quick rule-out when needed.

Moving Forward

- Validation gave a set of samples that can be used for future validations and verifications
- PulseNet 2.0 plan pending internal approvals using BioNumerics sample groups.
- Data for future training and employee competency testing.

Acknowledgements

- DC DFS Public Health Laboratory Leadership
 - Jocelyn Hauser, PhD
 - Pushker Raj, PhD
- Next Generation Sequencing Core
 - Wadih Bchara
 - Taylor Boyd
 - Janis Doss, PhD
 - Joel Steyer, PhD
 - Ikenna Uzomah

**For questions, please see my contact
information here:**

Email: eric.vaughn@dc.gov

Thank You!