Preliminary performance testing of the APAS® Independence with APAS® PharmaQC artificial intelligence system for environmental monitoring in sterile manufacturing.

Steven Giglio, Julie Winson and Rhys Hill
Clever Culture Systems, Adelaide, Australia
September 2023

ABSTRACT

Microbiological methods remain deeply rooted in traditional practices and new technologies are being considered to improve the agility, accuracy, consistency, and data integrity within pharmaceutical laboratories. One such technology, APAS PharmaQC, is an autonomous plate reader that uses artificial intelligence to detect microbial growth on environmental monitoring plates without any significant change to ways of working. Pilot primary validation with a suite of model organisms has demonstrated levels of performance for Linearity, Accuracy, Ruggedness, Precision, and Limit of Detection that are in line with compendial guidelines. These data demonstrate that APAS PharmaQC can provide benefits to pharmaceutical laboratories conducting environmental monitoring.

INTRODUCTION

Microbiology stands as a deeply traditional scientific field with little variation over the last 100 years in composition of agars and design of laboratory techniques. However, recent developments in microbiology and automation have introduced advanced tools to aid laboratory professionals in plate preparation and skilled reading tasks. In clinical microbiology, the integration of artificial intelligence (AI) and image analysis has demonstrated high performance (see https://cleverculturesystems.com/scientific-library/) which can alleviate workload pressures and improve overall quality of results. Similarly, the introduction of AI image analysis is now being implemented for utility in pharmaceutical microbiology.

APAS Independence with APAS PharmaQC interpretive software (hereafter called APAS PharmaQC) is an instrument that utilises AI to read and interpret culture plates used in routine environmental monitoring (EM) procedures of sterile manufacturing. Simply put, the algorithms within APAS PharmaQC have been trained to read the culture plate just like a human would, automatically reporting plates with no growth and providing an estimated colony count for plates showing microbial growth, delineating bacterial and mould counts, and flagging plates with spreading organisms. Plates demonstrating growth are also physically separated for human verification either manually with plate in hand, or through review of images on APAS PharmaQC web user interface. The system captures a permanent image of agar plates, delivers an objective evaluation of growth, provides consistency in interpretation of plates, and ensures data integrity by managing records in

line with the requirements of CFR 21 Part 11. Importantly, implementation of APAS PharmaQC does not necessarily require changes to how plates are prepared – that is, which manufacturer plates are sourced from, the number of plates used, method of sample collection and incubation, and how the results are reported and retained, minimising the extent of validation required.

The challenge when deciding how to determine performance in environmental monitoring has been to develop performance tests for a new technology to address the compendial definitions for alternative microbiological tests. Limit of Detection (LOD) is a good example. While clearly this is traditionally defined as a single colony that a human can see (and therefore detect) testing this explicitly within systems that are able to detect micro-colonies at early time points or use other mechanisms (such as AI or luminescence for example) for CFU detection presents a challenge. Determination of non-inferiority in this case means that the alternative method is 'no worse than' a microbiologist, but early detection will always be 'better' and time points for CFU determination will not be identical. There may also be changes to counting interpretation over time, and thus how would people react to an initial count, which may change over time? In this respect risk-based decision matrices determining fit-for-purpose applications would lend themselves for determining utility, but also comes with significant validation effort to define operational parameters for action and alert level responses at defined times.

In some cases, contriving samples with enough 1 CFU across multiple organism morphologies is also technically challenging, leading to a low number of plates that may not provide statistical veracity.

At the time of writing, there were no guidance documents specifically for the adoption of an instrument that uses AI to determine the presence or absence of growth on a plate, or to generate colony counts, in the pharmaceutical sector. PDA technical report 33 (1) suggests that "only the automated aspect of the method requires validation" which may be interpreted as a head-to-head comparison for instrument count versus a reference count (human read). A recent publication by Deutschmann et al., (2) nicely articulates broader considerations for evaluating and adopting new technologies for colony counting systems, however (and rightly so) warns that the views and recommendations in their position paper may not necessarily satisfy local authorities.

Even the compendial methods guiding validation such as USP <1223> (3), USP <61> (4), Ph Eur 5.1.6 (5), and Ph Eur 2.6.12 (6), are somewhat aloof, not being overly prescriptive in parts, almost dismissive in others, and may not necessarily be fit-for-purpose for emerging technologies. This is not an uncommon problem across all microbiology disciplines, and it therefore becomes germane for laboratory professionals to demonstrate non-inferiority and equivalency using a conglomerate of approaches and justifying applications

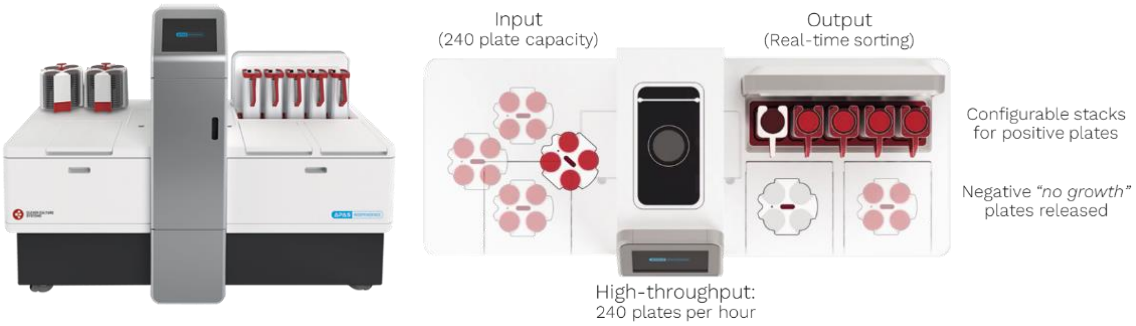


Figure 1. APAS PharmaQC instrument and topographical view

of tests in accordance with compendial definitions or other standards.

This paper serves as a discussion document which outlines an approach for a pilot primary validation for the APAS PharmaQC system and presents preliminary performance data of the system and addressing traditional compendial testing using AI and model organisms.

METHODOLOGY

APAS PharmaQC

APAS PharmaQC is a stand-alone instrument that fully automates culture plate imaging and interpretation of TSA plates with and without neutralisers, supporting most of the major media manufacturers seen in industry today (see Figure 1). The use of APAS PharmaQC is to substitute the primary read of an incubated EM plate without impacting the existing methods of plate preparation and incubation. After the appropriate incubation time, plates are loaded into the input area which has a holding capacity of 240 plates. The instrument is programmed to run plates though the central imaging station where both top- and bottom-lit images are acquired. The interpretive algorithms for growth detection and counting are then applied and if there is no growth on the plate it is triaged to the bottom output carriers for immediate release without user intervention. If growth is detected, plates are triaged to the upper output stacks where growth and counts are presented to qualified microbiologists for verification. In all cases, a contemporaneous report is issued to the Laboratory Information Management System (LIMS), and the image is available for digital review and storage, with full audit trails available. APAS PharmaQC allows a percentage of no growth plates to be reviewed (set by the user) to monitor performance.

Organisms and media used for testing.

Several model organisms were used for different primary validation activities. The reason for final test organism selection and source of information is provided in Table 1. These organisms broadly represent those groups likely to be encountered in environmental monitoring for sterile manufacturing areas (7) and include common skin and airborne organisms. Challenge organisms used in quality control testing of the agar as well as those organisms listed in the compendial methods have also been considered for testing. Organisms used in a pivotal validation paper (2) for colony counting

Table 1. Organisms used for primary validation

Organism	Reason for selection	Source	Used in test
Escherichia coli (NCTC 12923/ATCC 8739)	Fast grower Organism for QC testing of media.	PDA Journal Validation Paper (S. Deutschmann, B. Carpenter , C. Duignan, A Systematic Approach for the Evaluation, Validation and Implementation of Automated Colony Counting Systems, PDA Journal of Pharmaceutical Science and Technology 2022)	LOD, Ruggedness, Precision
Bacillus spizizenii (NCTC 10400/ATCC 6633)	Variable feathered edge colonies. Gram positive bacillus associated with people and the environment	PDA Journal Validation Paper Organism for QC testing of media USP 61, EU 2.6.12	LOD, Ruggedness, Precision, Linearity, Accuracy
Aspergillus brasiliensis (NCPF 2275/ATCC 16404)	Common filamentous fungus found in the environment	PDA Journal Validation Paper Organism for QC testing of media USP 61, EU 2.6.12	LOD, Linearity, Accuracy
Staphylococcus aureus (NCTC 10788/ATCC 6538)	Large yellow colony Gram positive coccus associated with people	Organism for QC testing of media. USP 61, EU 2.6.12	LOD, Ruggedness, Precision, Linearity, Accuracy
Pseudomonas aeruginosa (NCTC 12924/ATCC 9027)	Variable sized flat colony Non fermentative Gram neg bacteria common in water	Organism for QC testing of media USP 61, EU 2.6.12	LOD, Ruggedness, Precision
Moraxella osloensis (ATCC 19976)	Common environmental organism	Personal communication from industry representative	Linearity, Accuracy
Candida albicans (NCPF 3179/ ATCC 10231)	Common yeast found on people	Organism for QC testing of media	LOD, Linearity, Accuracy
Micrococcus luteus (ATCC 4698)	Common human-borne organism/skin contaminant	VALSource presentation - Marc Glogovsky Senior Consultant - Microbiology March 2019.	Linearity, Accuracy
Staphylococcus epidermidis (ATCC 14990)	Common human-borne organism/skin contaminant	Environmental Monitoring Capturing, Cultivating & Collating Viable Data	Linearity, Accuracy

technologies have also been considered. Combined, these organisms serve to demonstrate broad specificity. Media used for these pilot studies was Thermo Fisher TSA agar.

Limit of Detection (LOD)

In the traditional sense of LOD, this is typically the lowest level of detectable organism by an alternative method when compared to a compendial method. In colony counting terms, this is a single colony, and to some extent is already tested in Linearity and Accuracy. Practically, deposition of a single CFU of organism is challenging and variation in colony size is hard to test with model organisms.

In the context of APAS PharmaQC, the algorithm is always designed to detect a single colony of anything, as it will

differentiate agar from non-agar/organism, and then apply strict rules to define a colony of any bacteria, yeast, or mould. Therefore, the purpose of LOD in the context of APAS PharmaQC is to confirm that the software itself has been designed to detect colonies at the design limit of the optical system of the instrument which is 0.5 mm. This does not necessarily mean that all organism colony LOD must be 0.5 mm, as often microcolonies are hard to detect depending on the organism. Instead, a single organism type should return a LOD close to 0.5 mm and a spread of LOD measurements is expected based on organism used.

Five replicate plates were prepared from the panel organisms so that approximately 20-50 CFU were present on each plate and incubated for sufficient time so that colonies were tested when they just became visible until when they were clearly visible and typical in size. Plates were imaged at each time point. The aim was to have approximately 100 colonies to test for LOD across various time points. A microbiologist then used a specialised application to digitally annotate selected colonies within a series of images of the same plate taken over time, and the diameter in mm at which APAS PharmaQC detected these annotated colonies was determined when 95% of colonies were detectable.

Precision and Ruggedness of colony counts

The basis of this test is to measure the consistency of the APAS PharmaQC result for a single plate when assessed multiple times in a single instrument and across multiple instruments. It measures the ability of APAS PharmaQC to consistently provide within-run, and inter-instrument performance of colony count. Additionally, and within each instrument, the angle of plate presentation to the imaging station relative to the first position (being zero degrees) is varied to 120° and 270° to measure any impact of angle of plate presentation on the determination of colony counts. Determining rotational bias of image acquisition is an important part of any system that uses images to determine counts.

Each organism was prepared so that 5 replicate plates and a target CFU of 10-100 per plate were available for testing. *A. brasiliensis* was not included as the colonies are large and growth becomes confluent, decreasing the significance of this test. After incubation at 30°C and at day 3 and day 5, each plate was imaged 5 times in the same position in the instrument. The plate was then spun 120° relative to the first position and imaged 5 times, and then spun 270° relative to the first position and imaged 5 more times. The variance (measured as coefficient of variation (CV)) between the resulting 15 APAS colony counts was measured per plate and per instrument (Precision). The test for the same plate was performed on 3 different instruments, and the variance measured across instruments (Ruggedness).

Accuracy and Linearity of colony counts

Test organisms were prepared from fresh overnight cultures or bioballs such that target counts of 0.5, 5-10, 50, 100, 200, 250 CFU per plate were cultured. Each target count had 3 replicates prepared. Plates were incubated at 30°C for all organisms.

Each plate was read by a microbiologist to provide the reference result which was compared to the APAS PharmaQC result. Accuracy at each target count was determined by comparing the reference count with the APAS PharmaQC count. Linearity R^2 was determined by plotting the reference count against the APAS PharmaQC count and assessing correlation. The y-intercept and slope were also determined. Bland Altman plots were generated to represent the difference between the reference and APAS PharmaQC count per sample, relative to the mean of the counts.

RESULTS

Limit of Detection

Several organisms were used for LOD testing to include sufficient diversity. The lowest LOD was 0.6 mm for *C. albicans* (Table 2). This was the size when 95% of colonies could be detected, with 80% of isolates detected at 0.5 mm (data not shown). The variation in LOD across organisms is not unexpected as internal time-lapse videos show that the characteristic of growth is indeed different across organism and is not necessarily concentric, i.e., organisms do not grow uniformly from a single CFU in a radial manner. Although testing for *A. brasiliensis* was performed, the intent of this LOD test application does not provide useful information because typically mould investigations rely on extended incubations and sometimes specialised media, but also is typically an end point detection test where any colonies are usually large and obvious.

Table 2. Limit-of-Detection summary table

Organism	Limit-of-Detection (mm)
C. albicans	0.6
S. aureus	0.7
P. aeruginosa	2.5
A. brasiliensis	Not determined
B. spizizenii	1.3
E. coli	2.0

Ruggedness and Precision

Table 3 and Table 4 show the mean CFU and CV for measurement of Ruggudness and Precision. APAS 1 to 3 columns represent the results from 3 individual instruments, where the CV is measured for each replicate plate within each instrument (Precision). The All APAS column is a measure of agreement across APAS 1 to 3 where the CV is determined for each replicate plate (Ruggedness).

When looking at Precision on Day 3, the tightest precision was achieved with *S. aureus* with a 1.1% CV, while *B. spizizenii* CV ranged from 7-24.5%. Similarly for Ruggedness on Day 3, the CV for *S. aureus* ranged from 1.3 – 2.2% and for *B. spizizenii* the CV ranged from 8.3-20.4%. A similar trend is evident on Day 5 data.

Ph Eur 5.1.6 (5) states that for an alternative method a CV of 10-15% is acceptable for precision, but it must not be larger than that of the compendial method. In this case, the CV of the compendial method in not known. USP<1223> (3) provides guidance that 15-35% relative standard deviation/CV may be acceptable, although results outside this range both in the high and low side are possible. All APAS PharmaQC values (bar one Ruggedness CV for *B. spizizenii*) are within these combined

Table 3. Ruggedness and Precision for Day 3

Org Day 3	Growth level CFU per plate	Replicate	APAS1		APAS2		APAS3		All APAS	
			Mean	%CV	Mean	%CV	Mean	%CV	Mean	%CV
E. coli	10-100 CFU	1	48.7	4.3	51.8	5.7	48.3	6.2	49.6	6.2
		2	43.9	6.5	46.1	8.1	45.3	6.5	45.1	7.2
		3	66.7	3.6	63.5	5.5	74.7	8.4	68.3	9.4
		4	64.1	4.8	67.2	3.4	69.9	3.7	67	5.3
		5	68.9	5.2	68.3	6.3	73.2	6.1	70.2	6.5
S. aureus	10-100 CFU	1	81.7	2.3	81.2	2.4	80.5	1.7	81.1	2.2
		2	80.7	1.3	81.1	1.4	80.7	1.4	80.8	1.3
		3	69.2	2.1	69.6	1.3	69.8	2.1	69.5	1.9
		4	81.8	1.2	82.2	1.1	81.9	1.7	82	1.4
		5	97.1	1.5	96.1	1.7	95.7	1.5	96.3	1.6
P. aeruginosa	10-100 CFU	1	80.7	4.1	71.1	5.1	71.8	5.3	74.5	7.6
		2	65	6	58	13.2	64	7.9	62.3	10.3
		3	86.7	6.8	81.6	5.5	84.4	10.6	84.2	8.2
		4	81.9	9.3	78.9	8.5	74.5	8.4	78.4	9.4
		5	89	5.4	82.9	5.7	80	6.2	84	7.2
B. spizizenii	10-100 CFU	1	27.5	14.7	24.3	12.8	30.6	19.4	27.5	18.6
		2	19.3	14.7	16.5	8.8	17.1	9.3	17.7	13.3
		3	52.3	12	60.3	11.4	45.1	17.1	52.6	17.7
		4	34.9	7	35.1	8.6	37.6	7.7	35.9	8.3
		5	45.7	13.2	36.5	24.5	44.3	17.9	42.2	20.4

Table 4. Ruggedness and Precision for Day 5

Org Day 5	Growth level CFU per plate	Replicate	APAS1		APAS2		APAS3		All APAS	
			Mean	%CV	Mean	%CV	Mean	%CV	Mean	%CV
E. coli	10-100 CFU	1	54.9	8.3	107.5	6.8	73.5	15.3	78.7	29.8
		2	50.3	7.1	85.6	15.4	59.7	10.1	65.2	26.5
		3	74.6	8.6	121.5	7.3	93	8.2	96.4	21.7
		4	70.3	6.9	105.1	8.9	88.7	10.2	88	18.6
		5	70.9	4.3	123.1	7.2	94.8	7.4	96.3	23.4
S. aureus	10-100 CFU	1	79.7	2.3	79.3	2.3	80.5	1.7	79.8	2.2
		2	79.1	1.9	79.2	2.2	78.7	1.2	79	1.8
		3	66.5	3.4	65.5	2.5	67.1	2.8	66.4	3
		4	80.5	1.9	78.1	2.7	78.7	3.3	79.1	2.9
		5	94.9	2.1	95.3	1.6	96.5	2.1	95.6	2.1
P. aeruginosa	10-100 CFU	1	73.3	5.7	82.3	8.9	73.9	5.4	76.5	8.7
		2	63.9	13.7	75.5	8.3	66.1	11.8	68.5	13.2
		3	85.5	9	81.7	5.6	84.4	8.6	83.9	8
		4	73.7	9	77.1	11.5	77.8	13	76.2	11.4
		5	75.5	6.4	79	6.4	76.3	15	76.9	10
B. spizizenii	10-100 CFU	1	44.1	12.6	82.3	16.3	38.9	15.6	55.1	39
		2	23.7	11.5	51.6	17.4	23.9	7.7	33.1	43.3
		3	68.2	6.4	95.1	24.8	59.1	6.7	74.1	27.9
		4	41.8	16.9	75.2	9.5	38.3	9.6	51.8	34.5
		5	67	14.3	95.3	10.7	66.8	11.3	76.4	21.2

Table 5. Summary data for Linearity and Accuracy

Organism	1-10				1-50				1 - 100				Overall			
	Cases	r^2	Slope	Intercept	Cases	r^2	Slope	Intercept	Cases	r^2	Slope	Intercept	Cases	r^2	Slope	Intercept
A. brasiliensis	9	0.3971	1.19	1.44	17	0.7891	0.41	3.11	26	0.5116	0.22	5.85	36	0.4626	0.16	6.59
B. spizizenii	6	0.3405	0.67	5.59	10	0.9291	1.4	2.25	17	0.7601	0.79	11.83	36	0.7639	0.64	17.78
C. albicans	8	1	1	0	24	0.9898	1.03	-0.43	32	0.984	0.99	0.3	36	0.9861	0.99	0.23
M. luteus	3	1	1.17	-0.17	12	0.9879	0.9	1.42	20	0.8997	1.04	-1.69	36	0.9416	0.91	4.42
S. aureus	6	0.8276	1	0.17	8	0.9932	0.98	0.2	18	0.9922	0.94	0.53	36	0.9877	0.83	4.44
S. epidermidis	3	1	1	0	18	0.9816	0.9	1.42	24	0.9698	0.84	2.67	36	0.9644	0.85	1.99
M. osloensis	8	0.2196	1.15	2.09	16	0.9522	0.98	2.59	22	0.9834	0.97	2.66	36	0.9962	0.96	2.62

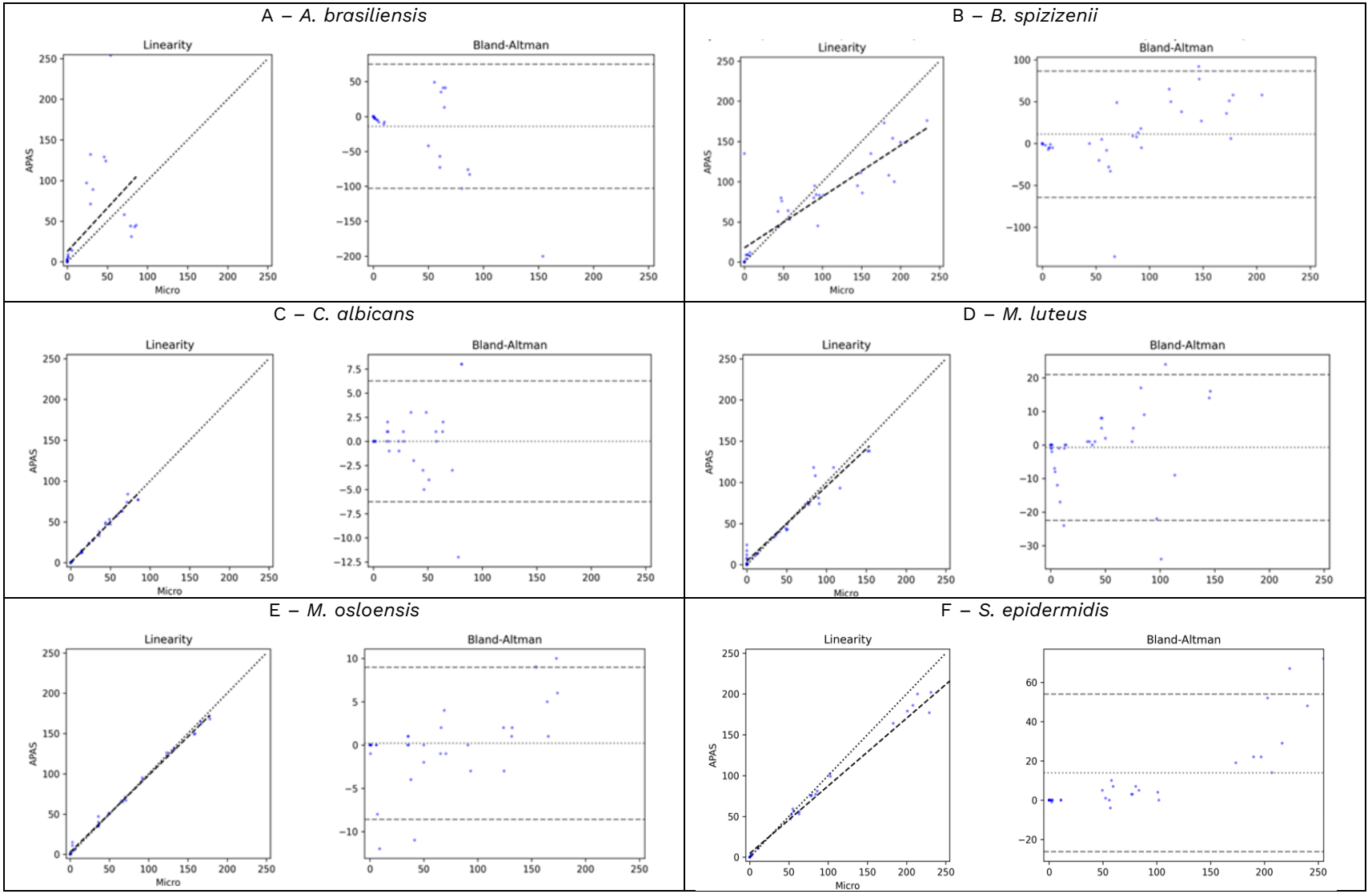


Figure 2. Linearity and Bland-Altman plots

compendial limits.

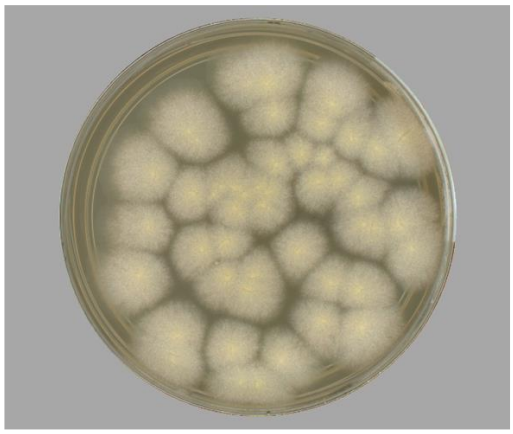
Accuracy and Linearity

Accuracy of the test is a measure of the closeness of the test result (APAS PharmaQC) to the reference result (manual counting). Accuracy is demonstrated if the tests show that the slope of the regression line does not differ significantly from 1 and if the y-intercept is not significantly different from 0 (5). Linearity is a measure of correlation, where the R^2 value should be close to 1 to indicate results that are proportional to the concentration of microorganisms present in the sample within a given range. USP<1223> (3) states that the correlation is acceptable if R^2 is at least 0.9025. Ph Eur 5.1.6 (5) has similar requirements to USP<1223> where R^2 must be at least 0.9 and the slope is between 0.8 and 1.2. Table 5 provides the summary data for measurements of accuracy and linearity. The data is stratified by CFU bands to provide insights into ranges of linearity at counts less than 10, less than 50, and less than 100. An overall measure across all CFU is also provided. Figure 2 provides graphical representation of the linearity study, and Bland-Altman plots containing long-dashed lines which represent the ±2 standard deviation limits.

For *S. aureus*, *C. albicans*, *M. luteus*, *S. epidermidis*, and *M. osloensis*, overall R^2 of 0.9416 to 0.9962 with slopes of 0.83 to 0.99 and y-intercepts of 0.23 to 4.44 were achieved. The Bland-Altman plots support the high levels of agreement with most points within ±2 standard deviation limits. *B. spizizenii* R^2 was as high as 0.9291 when looking at counts from 1-50 CFU, but overall R^2 is 0.7639. *B. spizizenii* is a challenge organism and

Figure 3. Example of B. spizizenii growth demonstrating variable morphology, size, and confluence

Day 3



Day 5

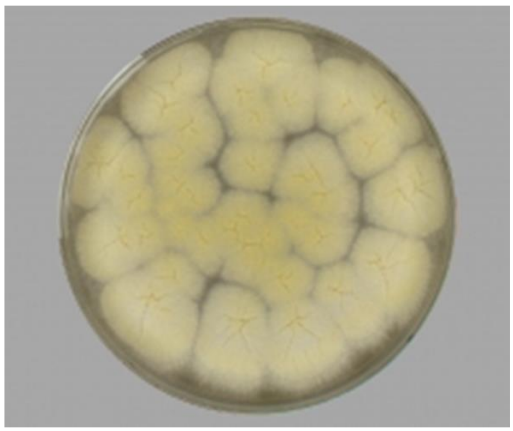


Figure 4. Example of A. brasiliensis growth changes over time demonstrating counting challenges

Figure 3 highlights the counting difficulties for humans and technology alike. The organism presents with variable morphology (smooth to feathered edges), variable size depending on the position and grouping of colonies, while merging colonies developing confluence are always contentious for determining true accuracy. Similar challenges exist for *A. brasiliensis* as depicted in Figure 4.

DISCUSSION

A pilot secondary validation study at AstraZeneca's Macclesfield site has provided encouraging results that the APAS PharmaQC system can effectively screen no growth plates from the workflow with high efficiency and 100% accuracy for detecting growth (8). Importantly, the system detected an initially false negative screen from the routine workflow which was subsequently rectified according to internal quality procedures. APAS PharmaQC integrated within the normal ways of working during this pilot study, with no additional modifications to media type used, incubation conditions or time, or any other procedures.

Underpinning this performance are the primary validation activities which interrogate specific performance characteristics of the technology. Preliminary APAS PharmaQC pilot primary validation data demonstrates performance in line with compendial guidelines for most tests and organisms. The expected challenges with counting *B. spizizenii* and *A. brasiliensis* were encountered, however the primary purpose APAS PharmaQC is to remove no growth plates from the workflow and provide an estimate of counts so that microbiologists can verify. Implementation of an alternative microbiology method requires a deep understanding of the technology and how to apply traditional compendial definitions to emerging technologies. The decision to choose a technology for implementation is multi-faceted (2), and certainly an important part is understanding the performance data to determine results equivalency and non-inferiority. Understandably, no single test provides the silver bullet to determine this, instead, laboratory professionals must use the data within the realms of understanding any technology limitations to identify an approach that will support their contamination control strategy and processes.

While work with APAS PharmaQC is ongoing (expanding pilot secondary validation data, additional tests for colony edge detection, Robustness/Interference testing, Operational Range, and executing formal primary and secondary validation studies), preliminary results demonstrate a high level of performance that supports the use of APAS PharmaQC for routine EM testing in sterile manufacturing. Given that greater than 98% of plates in EM do not have any growth present, using APAS PharmaQC to automatically report and remove these alone provides significant value add.

CLEVER CULTURE SYSTEMS

REFERENCES

1. PDA Technical Report #33, Evaluation, Validation, and Implementation of New Microbiological Testing Methods. PDA Technical Report No. 33 (revised) Evaluation, Validation, and Implementation of. Alternative and Rapid Microbiological Methods, Parenteral Drug Association, Bethesda, Maryland, 2013.
2. Deutschmann, S., Carpenter, B., Duignan, C., Knutsen, C., Salvas, J., Wysocki, L., Plourde, L., Johnson, L. and Eder, W., 2022. A Systematic Approach for the Evaluation, Validation, and Implementation of Automated Colony Counting Systems. PDA Journal of Pharmaceutical Science and Technology, 76(6), pp.509-526.
3. USP<1223> Validation of Alternative Microbiological Methods
4. USP<61>Microbiological Examination of Nonsterile Products: Microbial Enumeration Tests
5. Ph Eur 5.1.6 Alternative Methods for Control of Microbiological Quality
6. Ph Eur 2.6.12 Total Viable Aerobic Count
7. VALSource presentation - Marc Glogovsky Senior Consultant - Microbiology March 2019. Environmental Monitoring Capturing, Cultivating & Collating Viable Data
8. Australian Stock Exchange Announcement – LBT Innovations August 2023, First APAS® PharmaQC Performance Data from AstraZeneca Installation

DISCLOSURES

SG, JW, and RH are employees of Clever Culture Systems