A Comparative Study of Commercial ATP Hygiene Monitoring Systems

Based on data generated by Silliker Group Inc

Abstract

The performance of 5 leading commercially available of ATP hygiene monitoring systems were compared in the largest independent study of its kind. Linearity, Sensitivity, Repeatability, Precision and Accuracy are the major requirements of an effective ATP system. These performance criteria were determined using many replicates and dilutions of ATP, foodstuffs and micro-organisms.

The findings indicate that some of the systems show large variations and background interference which significantly affects their ability to meet the performance criteria.

The best system was revealed to be Hygiena SystemSURE Plus and the poorest systems were Charm Novalum and Neogen Accupoint.

Overview

1. Scope of the study
2. Components of an ATP system
3. Performance results based on studies criteria
   - Section 1: ATP performance results
   - Section 2: Foodstuffs performance results
   - Section 3: Micro-organisms performance results
4. Universal ATP device performance – Snapshot
5. Summary
6. Information links

Disclaimer:
This presentation was prepared by Hygiena and is solely based on data generated from the comparative study of ATP hygiene monitoring systems by Silliker Inc. Food Science Center Report RPN 13922 (2010).This presentation contains all the conclusions drawn by Silliker and further detailed analysis of the raw data. Silliker Inc supports all claims made in its report RPN 13922 however the additional data analysis has not been verified by Silliker Inc.

Scope of the Study:
Determine the performance of leading systems for the detection of ATP, foodstuffs and microorganisms

Purpose & Intent for ATP Hygiene Monitoring

The purpose of ATP bioluminescence for hygiene monitoring is to provide a simple, rapid, direct, objective test for cleaning verification. It is a sophisticated, sensitive indicator test to instantly determine the hygienic status and potential risk of the object being sampled.

Unlike microbiological tests that take days to yield results, ATP testing provides valuable information in seconds. The results from ATP surface hygiene monitoring are different to those of microbial enumeration methods and give additional information that the microbial test cannot provide.

ATP tests are not intended to replace microbial tests. However there is concurrent direct correlation between the results of the two methods because cleaning simultaneously removes both organic residues and microbes.

Performance Criteria

The key performance criteria evaluated were:
1. Linearity
2. Sensitivity
3. Repeatability
4. Accuracy

These performance criteria were determined by experimentation in controlled laboratory conditions using 3 sample types;

1. ATP dilutions pipetted directly to the swab bud

2. Dilutions of foodstuff: pipetted directly to the swab bud and tested on wet & dried-on stainless steel surfaces.

3. Dilutions of microbial cultures typical of those of concern to food & beverage processors including: *Escherichia coli, Lactobacillus plantarum, Pseudomonas aeruginosa, Salmonella typhimurium* and *Staphylococcus aureus* and one yeast culture, *Saccharomyces cerevisiea*.

ATP Detection Systems:

Three key components make up a system

An ATP system consist of 3 components.
Each component is a critical aspect of overall performance.

1. **Instrument – Luminometer (2 types were evaluated)**
   - Photodiode: sensitive, robust, requires low voltage, does not drift with time. It is low cost and has low background noise†.
   - Photomultiplier Tube: Sensitive, fragile, requires high voltage, drifts with time . It is expensive and has high background noise.

2. **Bioluminescent Chemistry (2 variations were evaluated)**
   - Liquid stable chemistry – new technology allows for immediate reaction with sample, gives greater precision, accuracy, and more consistency. Less manufacturing processes and therefore has lower cost.
   - Lyophilized chemistry – old technology (>30 years) requires complex expensive manufacturing, dry storage and rehydration at point of use that has larger variability.

   All chemistry uses luciferase/luciferin enzymes to generate light. The quality and quantity of the enzymes and the other components determine the performance of the chemistry.

† Background noise is light or electrical interference from the instrument or reagent swab device that causes a system to give a RLU reading in the absence of ATP. High background noise is commonly seen with photomultiplier tube machines. Some machines like the Charm Novalum deal with this by building in a background deduction algorithm. This reduces sensitivity. Other systems convert RLU to log RLU units (or zones) to disguise the high background and variation. Signal – Background Noise = True and Meaningful result

3. **Reagent Swab Device – Design & Wetting Agent**
   - ATP test device design and components play important roles in performance and cost per test.
   - ATP test devices are pre-wetted with an extractant to break up biofilms, collect and release ATP from a sample.
   - Four of the test devices tested used woven swab tips and one used a sponge like tip (Neogen).

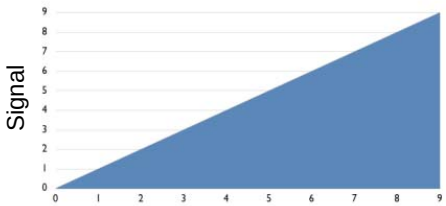
Five Commercial ATP Systems Used In Study

- **BioControl Lightning MVP & Surface sample devices**
  Photomultiplier tube based system / lyophilized chemistry / woven swab
- **Charm Science Novalum & Pocketswab Plus**
  Photomultiplier tube based system / lyophilized chemistry / woven swab
- **Hygiena SystemSURE Plus with Ultrasnap & Supersnap devices**
  Photodiode based system / liquid stable chemistry / woven swab
- **Neogen AccuPoint instrument and surface sampler**
  Photodiode based system /lyophilized chemistry / sponge swab
- **3M UniLite NG CleanTrace and CleanTrace swabs**
  Photomultiplier based system / liquid stable chemistry / woven swab

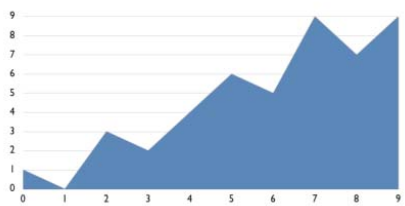
Performance Criteria:
Linearity, Sensitivity, Repeatability, and Accuracy

Linearity: Expression of predictability and reliability of the result.

Good Linearity r > 0.90

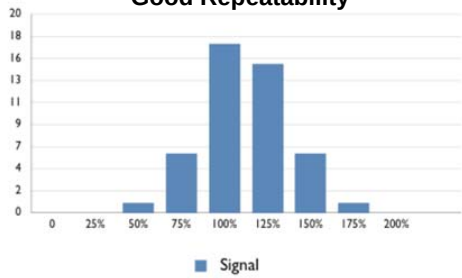


Poor Linearity r < 0.80

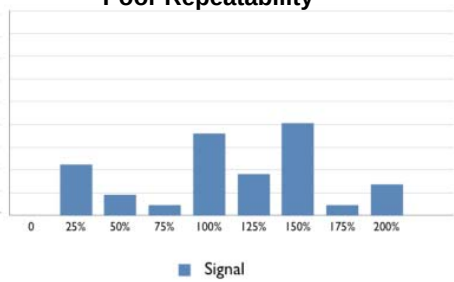


Repeatability: Variation between measurements by the same operator using the same test sample. Expression of consistency and reliability of results.

Good Repeatability

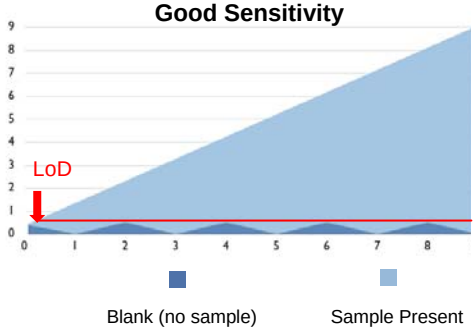


Poor Repeatability

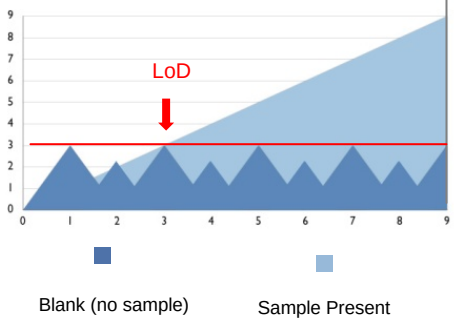


Sensitivity: Smallest amount that can be detected above the background of the system i.e. Limit of Detection (LoD)

Low Background
Good Sensitivity

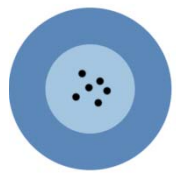


High Background
Poor Sensitivity

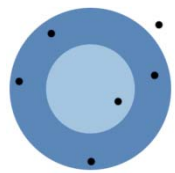


Accuracy: Recovery and detection of all available ATP in the sample.

Accurate

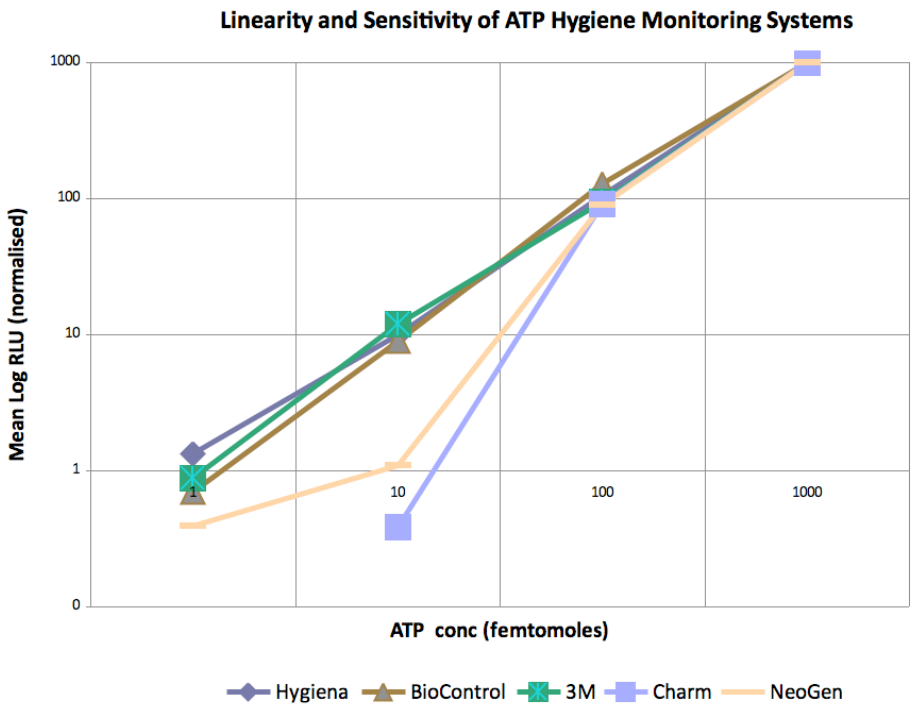


Inaccurate



A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Linearity:
Expression of predictability and reliability of the result



Graph

The linearity graph shows a direct, proportional, straight-line relationship between RLU and ATP.

$$y = mx + c$$

where 'y' = RLU and 'x' = ATP, and both increase in a constant predictable way. This means precise, reliable detection of low ATP levels at low RLU values.
Linearity is described by the term Correlation Coefficient (r) which shows how well the data approaches the perfect fit i.e. r = 1.000

Best Performance

BioControl, Hygiena & 3M all showed good linearity for the detection of ATP r = 0.94 – 0.98

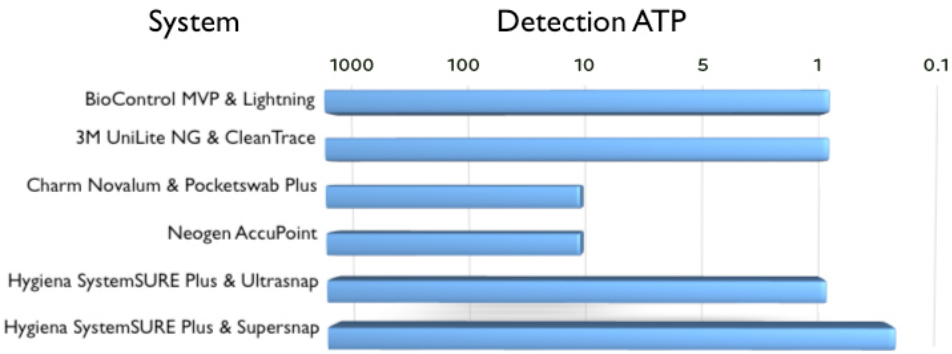
Poor Performance

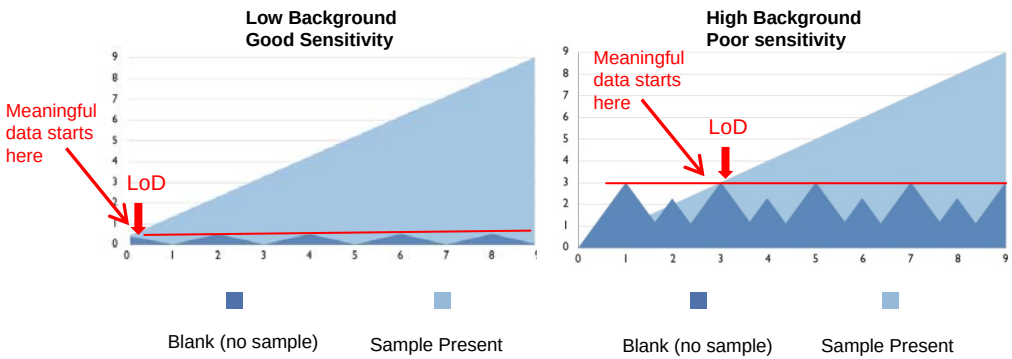
Neogen & Charm systems were not linear at low ATP levels; both systems displayed 0 RLU in the presence of ATP detected by other systems

Data located in Table 7 and 13 of Silliker report and Appendixes A, B, C

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Sensitivity:
Smallest detectable amount of sample





Background noise is the signal detected by the system when a blank device is tested i.e. no sample present. High background causes poor sensitivity. Some systems like the Charm Science Novalum build in a background deduction algorithm to compensate for the high background noise.

Graph

Sensitivity is defined as the Limit of Detection. It is the smallest amount detectable above the background of the system.

Background noise is the signal detected by the systems in the absence of ATP that can come from both the instruments (as electrical interference) and the reagent swab devices (as chemical interference from impurities).

Signal – Background Noise = True meaningful result
A low background noise means a clear signal with little interference that enables the detection of the lowest amount of sample i.e. maximum sensitivity.
The graph shows the limit of detection (LoD) for each ATP test system

Best Performance

Hygiena SystemSURE Plus & Supersnap – LoD = .017 fmols

Average Performance

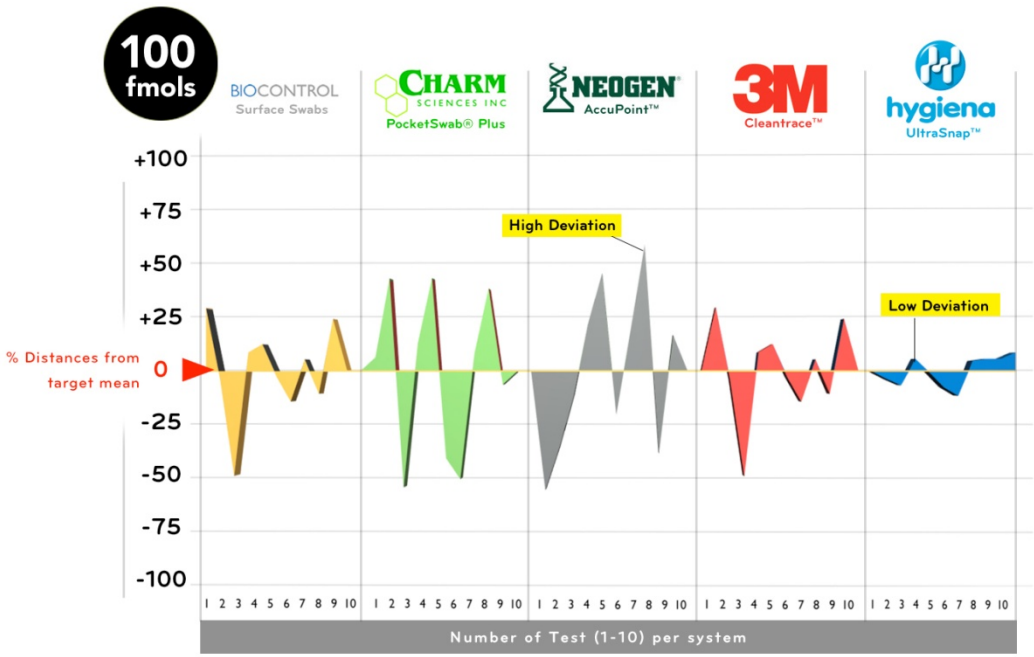
BioControl Lightning MVP & MVP swabs - LoD ~ 1.0
Hygiena SystemSURE Plus & Ultrasnap - LoD ~ 1.0
3M NG & CleanTrace - LoD ~ 1.0

Poor Performance

Charm Novalum & Pocketswab Plus - LoD = 10.0
Neogen Accupoint & Accupoint swabs - LoD = 10.0

Data located in Table 11, 12 and 13 of Silliker report and Appendixes A, B, C

Repeatability:
Variation between measurement by the same operator using the same test sample



Variation is described by the term Coefficient of Variation (CV%). The higher the CV% then the greater the variability which means the results is less consistent and more unreliable.

The overall variation across the whole range of ATP measurements show;

Best Performance
Hygiena and Supersnap – CV = 9

Average Performance
3M = 26% CV
Hygiena and Ultrasnap = 28%CV
BioControl = 39% CV

Poor Performance
Charm Science = 86% CV
Neogen = 123% CV

Greater differences between systems were highlighted when further analysis was performed. See graphs opposite.

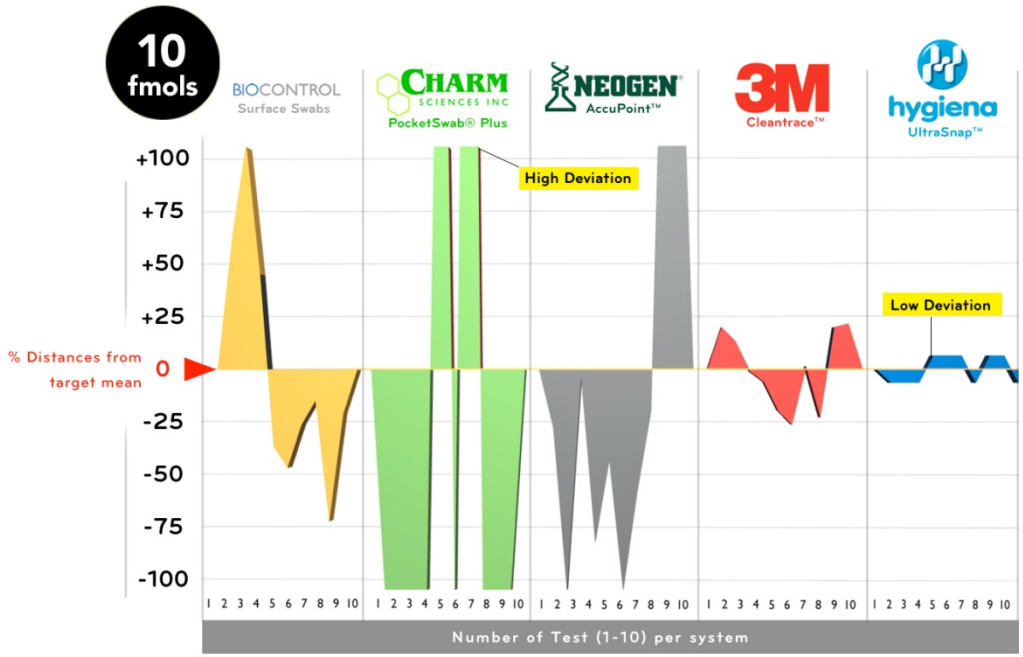
Recommended Pass / Fail limits are usually set between 10 and 100 fmols ATP so good repeatability is essential at these critical values.
The graphs opposite show that Hygiena is the only system with low deviation which means it delivers the most reliable, consistent and dependable results.

This is particularly important where the highest standards of quality and safety requires the detection of very low levels of ATP (see next page).

Data located in table 9 and 10 of Silliker report and raw data in Appendix C.

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Repeatability:
At critical limits



High standards of cleanliness are required for high care production processes where safety and quality are of paramount importance.
Under these circumstances, the reliable and precise detection of low levels of ATP is essential.

The graphs opposite reveals large variations and differences between systems at low ATP values (10fmols) that are essential for high specific cleaning duties.

Best Performance
Hygiena and Supersnap – CV = 7%CV

Average Performance
Hygiena and Ultrasnap = 10%CV
3M = 17% CV

Poor Performance
BioControl = 53% CV
Neogen = 116% CV
Charm Science = 214% CV

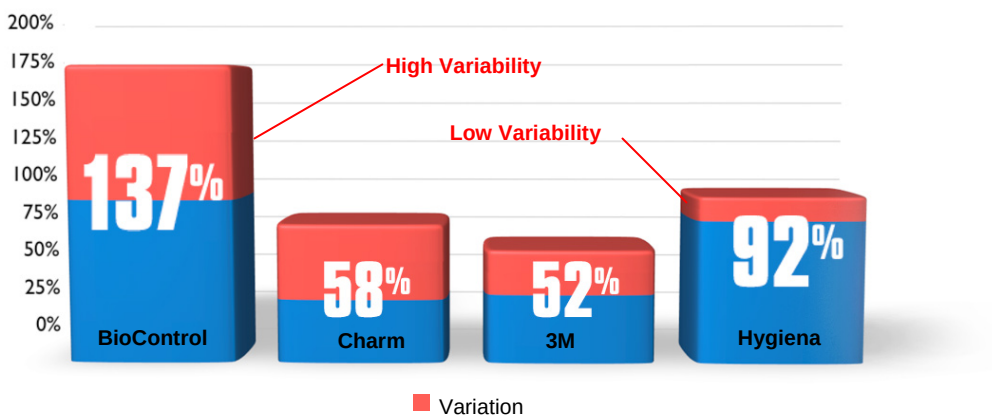
The graphs show that only Hygiena SystemSURE Plus can consistently deliver dependable results at low ATP levels.

Data located in table 9 and 10 of Silliker report and raw data in Appendix C.

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Accuracy:
Recovery and detection of all available ATP to reflect the true value of the sample

Recovery of ATP on Swab (%)



Note: Neogen Accupoint was not a part of this section of the study because of the design of the test device. Performance in the other sections of the study indicate that recovery of sample would be poor.

ATP (at 100 fmols) was added to each test device and measurements were made (using 10 replicates) to determine how much of the available sample was actually detected. If 100% of the ATP was detected then the system is accurate and gives a true meaningful results.

The best system will be closest to 100%

Less than 100% means that only part of the sample was detected due to some interference within the system. This means that the system does not give a true result and is not accurate.

Best Performance
Hygiena – 92% recovery

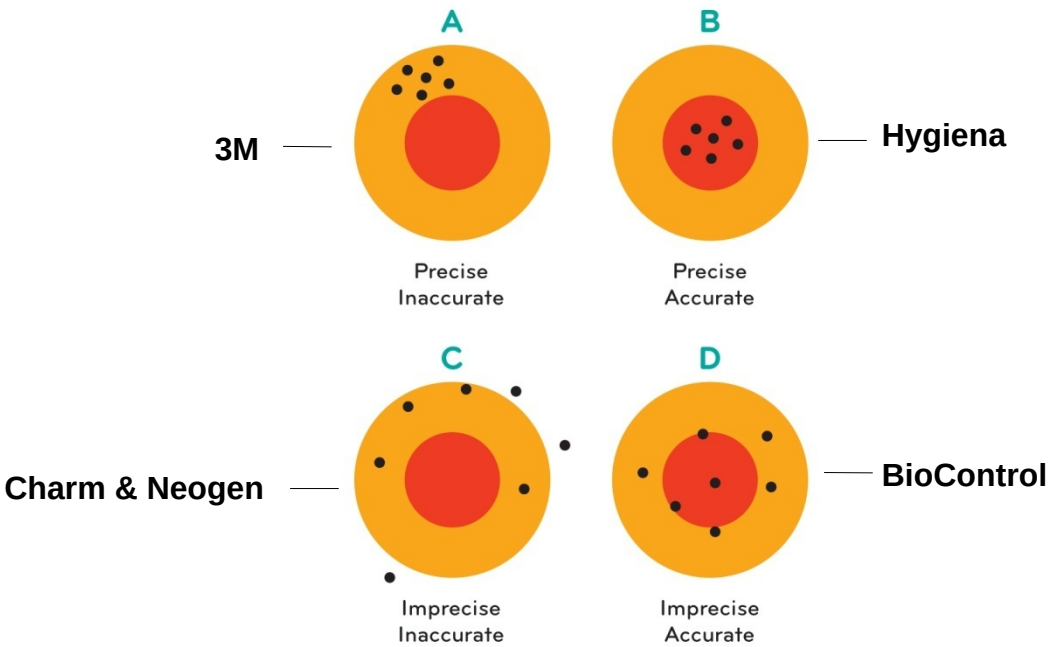
Average Performance
BioControl – 71 to 137% recovery (highly variable)

Worst Performance
3M – 52% recovery
Charm Science – 58% recovery

Data located in figure 2 page ‘26’ of the Silliker report

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Precision and Accuracy:
Recovery and detection of all available ATP to give consistent reliable results closest to the true value



Graph

The illustration opposite is used to describe the performance of a test and it shows how precision and accuracy are linked. It is important that they are considered together.
Better precision and accuracy means more consistent and reliable results that are closest to the true value.
The Hygiena system was the only system to be both precise and accurate.

Best Performance

Hygiena SystemSURE Plus & Ultrasnap displays precise, accurate results.

Average Performance

BioControl Lightning MVP & MVP swab displays accurate results, but is not precise.

3M NG & CleanTrace system is precise but only recovered 52% of the sample and is not accurate.

Worst Performance

Charm Novalum & Pocketswab Plus display imprecise and inaccurate results.

Neogen Accupoint & Accupoint swabs display imprecise and inaccurate results.

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Results Summary:
ATP detection looking at linearity, variability, repeatability & sensitivity

System	Linearity	Output (RLU)		Variability	Sensitivity
	(r)	Blank (Background at zero ATP)	Maximum (at 1000 fmols ATP)	Overall Average (CV%)	Limit of detection (fmols ATP)
BioControl Lightning MVP with Lightning swab	0.982	283	975,941	39	1.1
3M UniLite NG with CleanTrace swab	0.988	4.3	7386	26	1.3
Charm Novalum with Pocketswab Plus	0.949	0**	418,517 *	86	10.0
Hygiena SystemSURE Plus with Ultrasnap swab	0.988	0†	1589	28	1.0
Hygiena SystemSURE Plus with Supersnap swab	0.987	0†	4949	9	0.17
Neogen AccuPoint with Accupoint swab	0.976	0**	15,649 *	123	10.0

High RLU output does not give better sensitivity or performance

* does not detect below 10 fmols at which level the instrument shows 0 RLU.

** not a genuine zero reading (limited instrument output)

† SystemSURE Plus is the only system with genuine low background that is linear to zero RLU

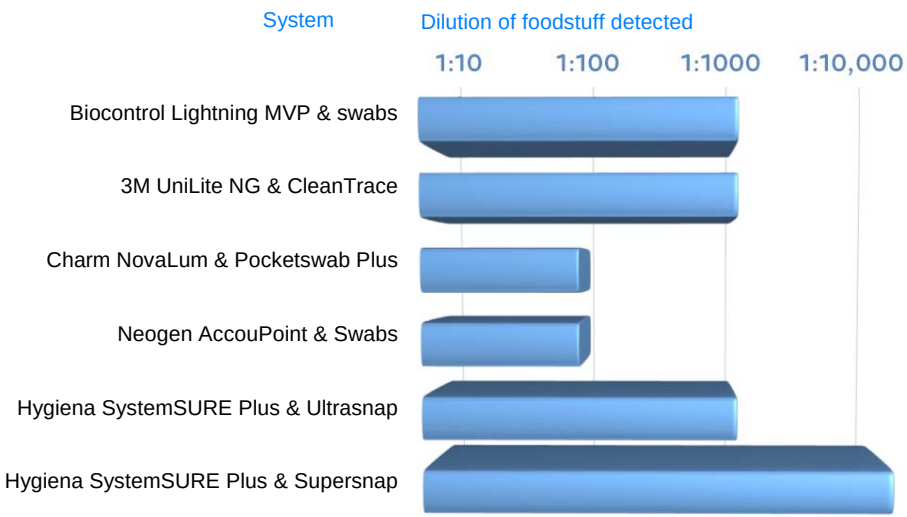
Performance	Linearity	Sensitivity	Repeatability	Accuracy
Best Performance	Hygiena BioControl 3M Charm Neogen	Hygiena	Hygiena	Hygiena
Average Performance		BioControl 3M	3M BioControl	BioControl
Worst Performance		Charm Neogen	Charm Neogen	Charm Neogen 3M

New Technology, better system design = BETTER PERFORMANCE

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Section 2:
Detection of Foodstuffs

Food Type: Ground beef or Pasteurised Milk



Ground Beef Milk

Similar results obtained for orange juice and mixed green salad

Evaluation of ATP systems in a simulated Food & Beverage environment. This is just one industry that ATP systems are used to monitor hygiene and contact surfaces.

Background Information:
4 food items were chosen for this section of the study.
•Pasteurised milk
•Ground beef
•Orange juice
•Mixed green salad

4 food items were made into liquid suspensions and also dried on to a stainless steel surface to test for pickup and extraction capabilities.

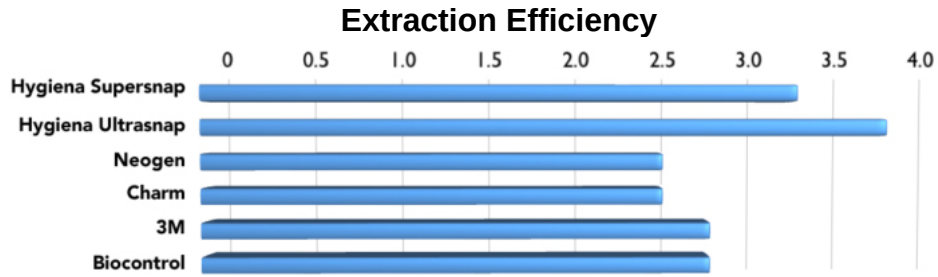
Results:

Performance	Linearity	Sensitivity	Repeatability	Accuracy
Best Performance	Hygiena BioControl 3M Charm Neogen	Hygiena	Hygiena 3M	Hygiena
Average Performance		Hygiena BioControl 3M	BioControl Charm	BioControl
Worst Performance		Charm Neogen	Neogen	Charm Neogen 3M

Data located on page ’15,16 & 31’ of Silliker report and Appendix C

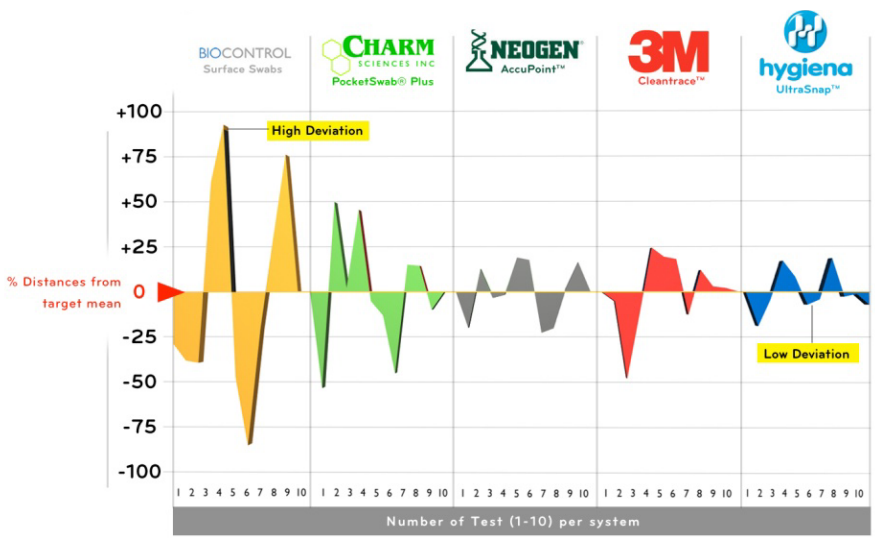
A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Detection of Foodstuffs:
Variation and extractability



Serial Dilution of Foodstuffs

Variation between systems with 1-10 dilution in Milk



Performance qualification:
The detection of small amounts of food residues on surfaces after cleaning is dependent on the ability to detect all the available sample collected.
The extraction efficiency of the reagent swab device determines the smallest amount of sample that can be detected. The graph compares the extraction efficiency of each ATP system and shows that the larger the dilution factor detected then the more sensitive the systems.

Similarly the less variation in the system then the more consistent and reliable the result. The graph and table below shows the variation for pasteurised milk. Neogen Accupoint did not detect 1 in 100 dilution or lower. Charm did not detect 1 in 1000 dilution .

•Best Performance
•**Hygiena** SystemSURE Plus & Ultrasnap with x10 higher extraction efficiency and greatest consistency across the whole dilution range.

•Average Performance
•**BioControl** Lightning MVP & MVP
•**3M** NG & CleanTrace.

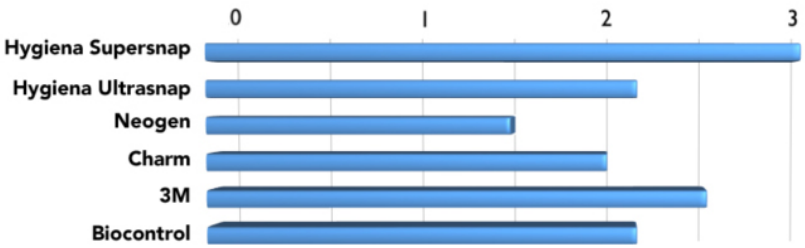
•Worst Performance
•**Charm** Novalum & Pocketswab Plus
•**Neogen** Accupoint & Accupoint swabs

ATP system	Variation (CV%) in dilutions of milk			
	None	1 : 10	1 : 100	1 : 1000
BioControl Lightning MVP with swabs	26	57	27	61
3M UniLite NG with CleanTrace swabs	30	20	16	10
Charm NovaLum with Pocketswab Plus	24	32	39	ND
Neogen AccuPoint with swabs	22	15	ND	ND
Hygiena with UltraSnap swabs	15	11	10	15
Hygiena with SuperSnap swabs	16	11	10	43

Data located in Table 16 of Silliker report and Appendix C

Section 3:
Detection of Micro-organisms

Extraction Efficiency



Serial Dilution of Micro-organisms

All organisms were detected and yeasts showed the highest amount of ATP because they were the largest. Staph aureus showed lowest amount of ATP. There was large variation in results from all organisms but good linearity was shown by all systems. ATP systems can detect these micro-organism but the limit of detection was shown to be 10,000 to 100,000 cfu/ml. This is generally not sensitive enough for most cleaning or hygiene programs. Microbes multiply in organic matter and ATP systems verify that organic residues have been removed thus reducing the risks. Therefore ATP testing is not a replacement for microbial testing, but an additional proactive step in support of a complete food safety or hygiene safety program.

Background :
ATP systems are intended to detect very small amounts of organic material on surfaces or in liquid samples. Micro-organisms are organic and contain ATP but at much smaller amounts than foodstuffs. Micro-organisms contain different amounts of ATP dependent on the type of microbes, size and state of health. The test cannot differentiate food ATP from microbial ATP.

Serial dilutions of 5 bacteria and 1 yeast culture were prepared and tested in each system to determine the limits of detection.

Results from Section 3: Detection of Micro-organisms

Performance	Linearity	Sensitivity	Repeatability	Accuracy
Best Performance	Hygiena BioControl 3M Charm Neogen	Hygiena	Hygiena 3M	Hygiena
Average Performance		Hygiena BioControl 3M Charm	Hygiena BioControl Charm	BioControl 3M Charm
Worst Performance		Neogen	Neogen	Neogen

Snapshot:
Universal ATP test device that makes other systems better

Background:
The performance of Hygiena's liquid stable reagent technology in other luminometers was evaluated to see if Hygiena's new technology improved the performance of instruments of other systems. Snapshot was tested against all the same criteria using ATP, foodstuffs and microbes. Results from the Silliker study showed that Hygiena's test device technology is the most superior currently on the market.

"The Hygiena swabs collectively are more sensitive to ATP and better at detecting low level food and cultures than all other systems." (Page 17 of Executive Summary report)

Results from ATP, Foodstuffs & Micro-organism

Performance	Linearity	Sensitivity	Repeatability	Accuracy
Improvement in Performance	☑	☑	☑	☑
Same Performance				

Snapshot was consistently shown to provide the following benefits;
- Improvement in linearity
- Improvement in accuracy
- Increased sensitivity (2x to 30x fold)
- Reduction in background noise
- Improved repeatability and consistency (lower c.v.)
- Increased extractability of ATP

Sensitivity Results

System	Sensitivity (Limit of Detection: fmols ATP)	
	Suppliers own swab	Hygiena Snapshot
BioControl MVP	1.1	0.04
3M UniLite NG	1.3	0.42
Charm Novalum	10.0	5.0

There is no snapshot device suitable for the Neogen Accupoint system

Data located in tables 7 – 17 and Appendices A,B and C of Silliker report

A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.

Summary

- 5 commercial ATP detection systems were compared for the detection and measurement of ATP, foodstuffs and micro-organisms.
- All systems were shown to give a good linear response to all 3 sample types, however there was a difference in the sensitivity, repeatability and accuracy between systems.
- The most precise, accurate and repeatable systems was the Hygiena SystemSURE Plus with either Ultrasnap or Supersnap swabs.
- The least sensitive and most variable systems were the Neogen AccuPoint and Charm Science Novalum systems with their respective swabs.
- The Snapshot Universal swab improved the repeatability and sensitivity of the 3M UniLite NG, BioControl MVP and Charm Novalum luminometers.
- ATP hygiene monitoring is a cleaning verification test so accuracy and consistency at low ATP levels is critical. Study shows that Charm and Neogen have the poorest sensitivity and highest variability.
- Data confirms that the SystemSURE Plus photodiode based system is equal to or better than the other commercial photomultiplier tube based system.
- Data shows that each system displays a different RLU for any given sample. This could be confusing to users comparing systems. It is clear that a larger RLU number does not mean a more sensitive reading. Charm displayed the highest RLU value result for each sample type and is one of the least sensitive systems.
- All ATP systems can detect micro-organisms in the absence of other organic matter, but at a level of detection was 10,000 – 100,000 cfu/ml. This is generally a lot higher then cleaning standards allow. Detection of micro-organisms showed a large variation based on species and quantity. Hygiena Supersnap had the best extraction efficiency for the recovery and detection of microbes.

Helpful Links & Documents

- How to setup an ATP hygiene monitoring program
- The difference between Photodiode and Photomultiplier Tube systems
- How to set thresholds
- Understanding Relative Light Units (RLU)
- Visit Hygiena at www.hygiena.net




A Comparative Study of Commercial ATP Hygiene Monitoring Systems
Based on Data Generated by Silliker Group Inc.