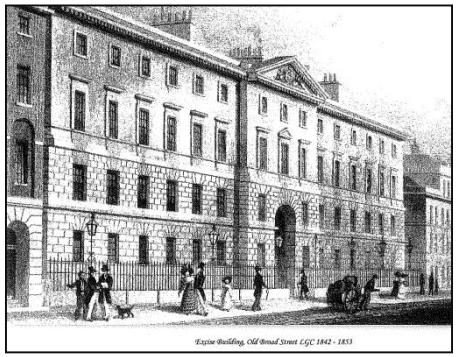
Does Proficiency Testing improve Laboratory Performance?

Science
for a safer world

History of LGC

- LGC are a leading international life sciences measurement and testing solutions company
- First founded in 1842 to detect adulteration of goods e.g. tobacco for HM customs & excise
- A public company until privatized in1996, with headquarters in Teddington, UK
- Currently 2000+ staff with offices in 22 countries



www.lgcgroup.com

History of Proficiency Testing (PT)

- 1880's - Collaborative studies to validate analytical methods were carried out by the Association of Agricultural Chemists (USA)
- 1930's – Collaborative studies carried out in UK under direction of Society for Analytical Chemistry
- 1947 - First formal PT study took place in USA on analysis of human tissue
- 1953 – First UK PT scheme on blood testing
- 1969 – National scheme in UK (UK NEQAS) for clinical chemistry testing of human serum was initiated funded by DHSS
- 1971 – First private sector (commercial) schemes
- Driven by legislation and accreditation

LGC Proficiency Testing (PT)

- LGC has organised PT programs for over 30 years.
- One of the first PT providers in UK to be accredited (UKAS 0001)
- Purpose built facilities in Bury, UK (2011)
- More than 1,500 PT exercises per year
- More than 10,000 laboratories serviced
- More than 250,000 samples produced and distributed per annum

Definition of PT

- Proficiency testing or EQA is the regular distribution of test materials to participating laboratories in order to independently compare the accuracy of their analytical measurements
  - with an external standard of quality
  - with the measurements of its peers
  - with past performance
- PT schemes may be sequential (same) or simultaneous (similar), quantitative, qualitative or interpretive.
- PT may relate to entire test procedure or just part.

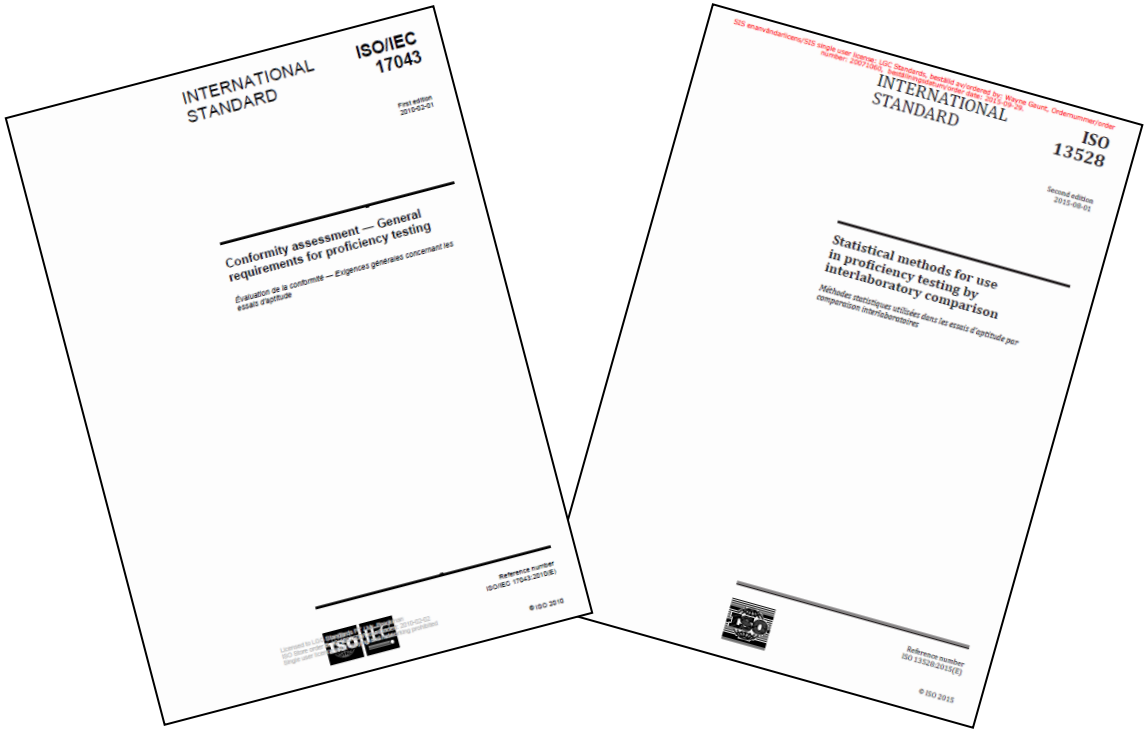
Aims of PT

- To assess the quality of a laboratory's results
- The promotion of improvements in the quality of analytical data
- To identify competent laboratories for regulatory, commercial or other purposes
- To provide support for laboratory accreditation activities
- Comparison of the performance of different analytical methods

Scheme Organisation

PT standards

- ISO/IEC 17043:2010
- ISO/IEC 17025:2005
- IUPAC International Harmonised Protocol (2006)
- ISO 13528:2015



The PT cycle

Sample preparation & despatch
✓Fit-for-purpose
✓Homogenous
✓Stable

Participants test and report results
✓ Realistic timescale
✓Clear instructions
✓Easy to return results
✓Use own methodology

Data analysis
✓Sufficient numbers
✓Statistically valid
✓Methodology

Participants receive summary of performance
✓Comprehensive reports
✓Fast reporting
✓Individual or full
✓Trend analysis

PT samples

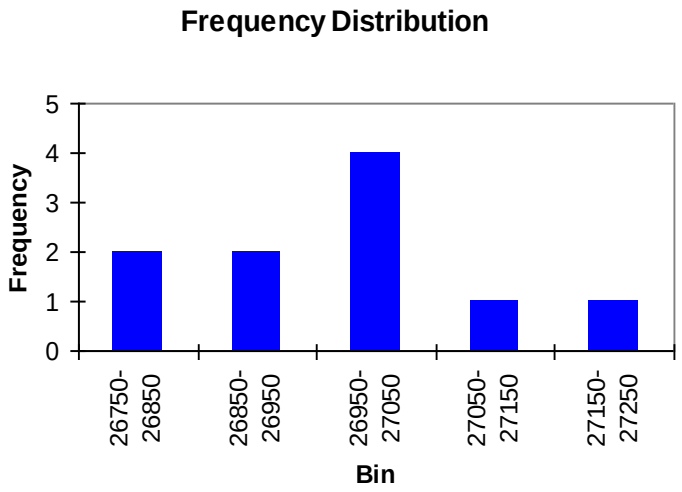
- Samples may be real products, adulterated products, or manufactured in-house to simulate ‘real’ products
- Samples are carefully selected to meet the needs of participants while still meeting the requirements for homogeneity and stability
- Sample contents will vary from round to round, not all analytes may be present every time
- ‘Real’ samples not always appropriate or available

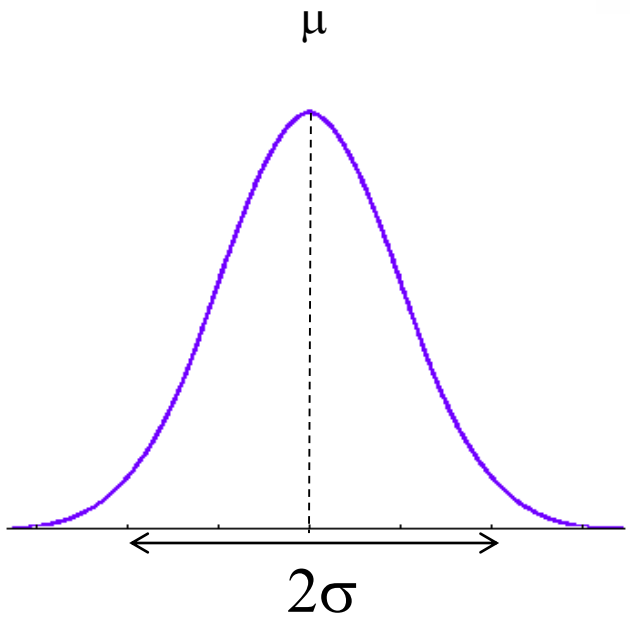
Performance Assessment

Statistical design

- According to ISO/IEC 17043;
  - PT providers shall use valid methods of evaluation
  - Reports should contain an indication of the performance of individual participants
  - Data analysis should generate summary statistics and performance statistics
- The assumption for most schemes that the underlying distribution of the data is ‘Normal’

The normal distribution





± σ	% population
1.00	68.3
1.64	90.0
1.96	95.0
2.00	95.4
2.57	99.0
3.00	99.7

Principle of performance assessment

To Compare…..

- Observed error
  - difference between laboratory result (x_i) and assigned value (x_{pt})
- ‘Allowed’ error
  - Defined by the standard deviation for proficiency assessment (σ_{pt}) or uncertainty

…. using a performance score

Assigned values (x_{pt})

- Best estimate of the actual concentration or level of the measurand in the PT sample

- Methods of determining the assigned value
  - formulation
  - certified reference values
  - reference values
  - consensus values from expert laboratories
  - consensus values from participants

Satisfactory range

- Defines the scale of acceptable variation among participants, using the standard deviation for proficiency assessment (σ_{pt})

- Methods of determining the SDPA (σ_{pt})
  - by perception of experts
  - by experience from previous PT rounds
  - results from an interlaboratory validation exercise (collaborative study)
  - general model e.g Horwitz
  - data obtained in each PT round

Performance scores

Key performance measures:

- *z*-scores, z'-scores and zeta-scores
- E_n numbers
- Estimates of laboratory bias (D ,$D_{\%}$ and Q-scores)

Examples of performance scores

- E_n numbers

$$(E_n)_i = \frac{x_i - x_{pt}}{\sqrt{U^2(x_i) + U^2(x_{pt})}}$$

- Laboratory bias estimates:

$$D_i = x_i - x_{pt}$$

$$D_i\% = \frac{100\left(x_i - x_{pt}\right)}{x_{pt}}$$

- Laboratory bias estimates: (*Q* scores)

$$Q_i = \frac{x_i - x_{pt}}{x_{pt}}$$

z-scores

$$z_i = \frac{(x_i - x_{pt})}{\sigma_{pt}}$$

x_i = submitted result
x_{pt} = assigned value
σ_{pt} = standard deviation for proficiency assessment

$|z| \leq 2$ Satisfactory performance (95%)
$2 < |z| < 3$ Questionable performance (5%)
$|z| \geq 3$ Unsatisfactory performance (0.3%)

Using PT to monitor & improve performance?

What affects performance?

Non-analytical errors:

- Calculation
- Transcription
- Units
- Reporting format
- Storage
- Sample
- Definitions

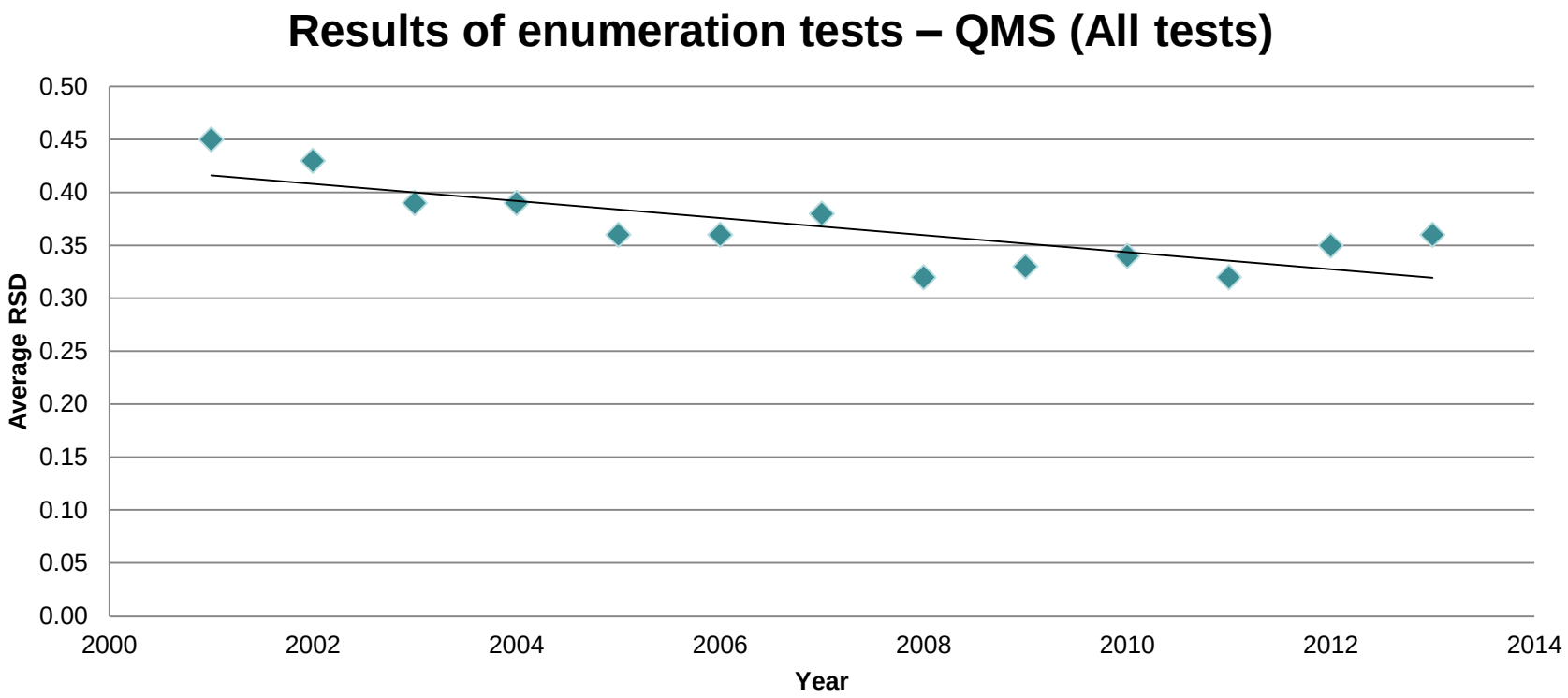
Analytical errors:

- Calibration
- Instrument problems
- Method performance
- Extraction/clean-up
- Interferences/matrix effects
- Diagnostic kits/reagents
- Analyst

Does PT improve performance?

- Proficiency testing is a self-help tool that enables participants to detect unexpected errors in their results
- It is not designed to be diagnostic
- Performance is highly dependant on the PT scheme;
  - Appropriateness of sample/matrix/analytes
  - How was assigned value and σ_{pt} set
  - Methods used
  - Other participants
- Like any scientific experiment, a single result is not very helpful

Does PT improve performance?



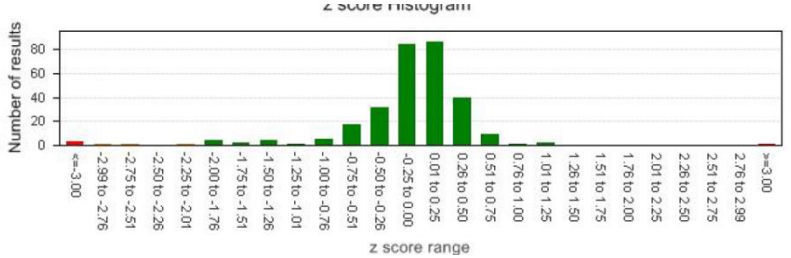
It was the sample…..!

Scheme: QWAS - Quality in Water Analysis PT Scheme

Sample: 412 - Potable Water Indicator combination

Analyte: Escherichia coli

WT9113	Other	8	-2.54
WT9134	MEMF Chromogenic agar 37	55	-0.15
WT9154	MEMF Chromogenic agar 37	48	-0.32
WT9154	MEMF mFC	47	-0.34
WT9154	MEMF TTC + Tergitol 7	68	0.11
WT9166	MEMF MLGA 37	22	-1.29
WT9191	Other	0	
WT9217	Colilert	107	0.68
WT9252	MEMF TTC + Tergitol 7	37	-0.64
WT9277	Other	55	-0.15
WT9282	MEMF TTC + Tergitol 7	46	-0.37
WT9321	MEMF Chromogenic agar 37	32	-0.82
WT9334	MEMF MLSB 37	53	-0.19
WT9348	Colilert	66	0.08
WT9348	MEMF Chromogenic agar 37	65	0.06
WT9357	Colilert	83	0.36
WT9358	MEMF Chromogenic agar 37	37,700	7.95
WT9361	MPN	120	0.82
WT9388	MEMF Chromogenic agar 37	69	0.13
WT9388	MEMF TTC + Tergitol 7	69	0.13
WT9389	MEMF Chromogenic agar 37	65	0.06
WT9389	MEMF TTC + Tergitol 7	58	-0.08
WT9390	MEMF Chromogenic agar 37	62	0.00
WT9391	MEMF Chromogenic agar 37	59	-0.06
WT9391	MEMF TTC + Tergitol 7	52	-0.22
WT9408	MPN	33	-0.78
WT9412	MPN	78	0.28
WT9418	Other	35	-0.71
WT9422	MEMF Chromogenic agar 37	52	-0.22
WT9424	Other	0	
WT9425	Other	38	-0.61
WT9428	MEMF Chromogenic agar 37	19	-1.47
WT9430	MEMF Chromogenic agar 37	69	0.13
WT9434	MEMF TTC + Tergitol 7	75	0.24



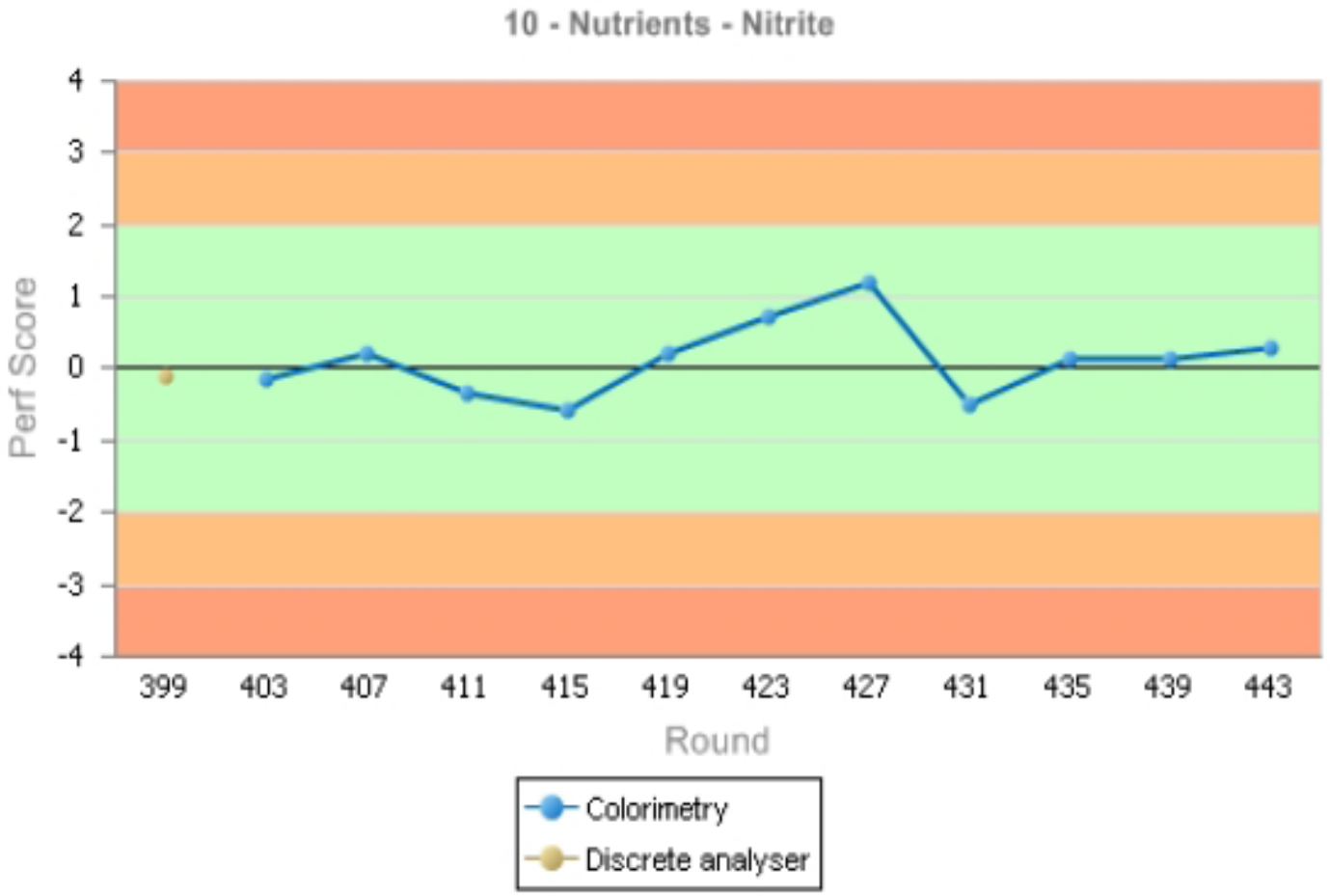
Data Statistics

	Value
Number of Results	308
Number of Excluded Results	15
Mean	1.75 log10
Median	1.79 log10
Standard Deviation	0.20 log10
Robust Standard Deviation	0.10 log10
Result Range	3 to 160 cfu/100ml

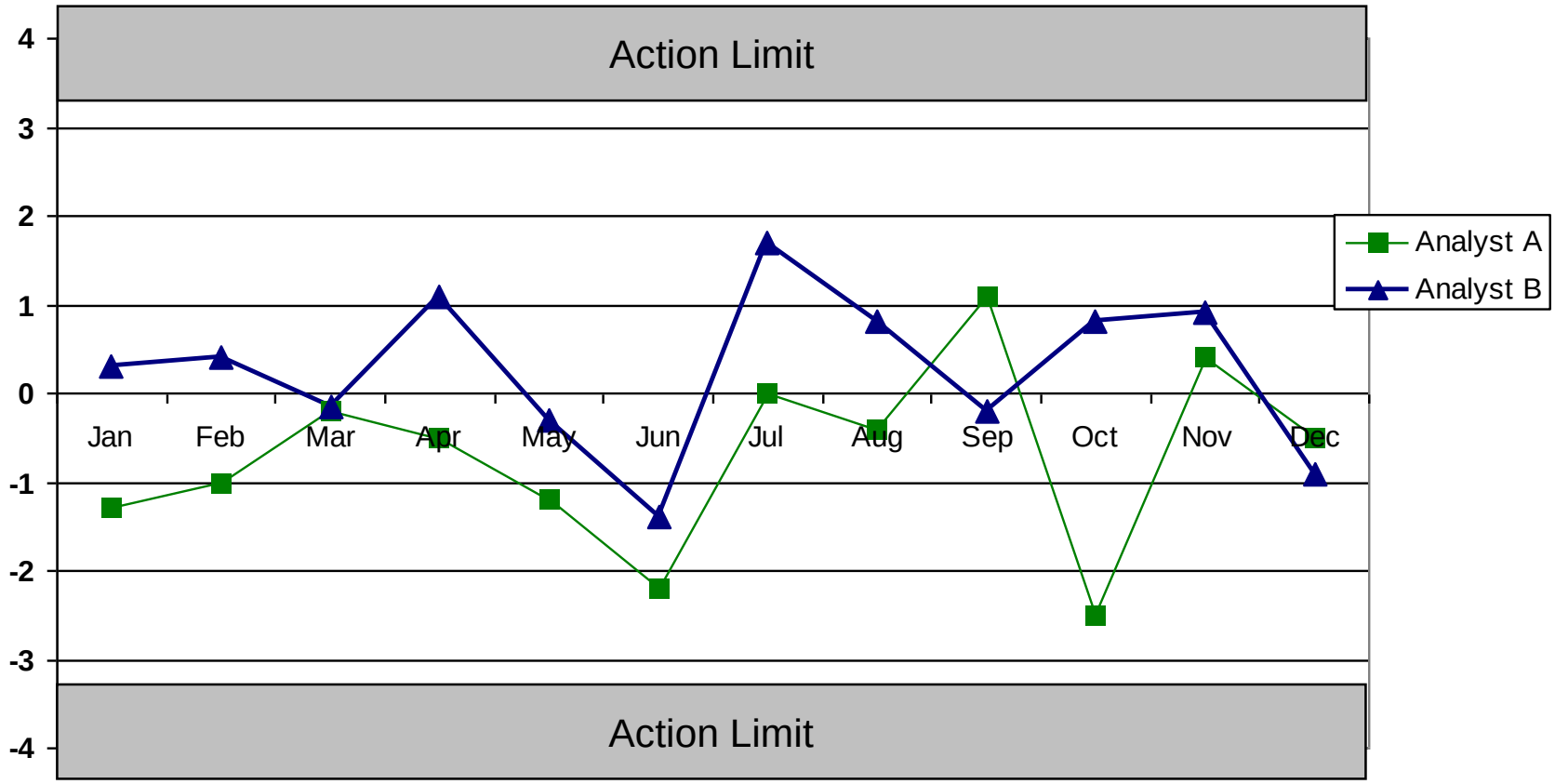
Performance Statistics

	Value
Assigned Value	62 cfu/100ml
Uncertainty of Assigned Value	0.01 log10
SDPA	0.35 log10
Satisfactory Range	12 to 311 cfu/100ml
Satisfactory z scores	97.6%
Questionable z scores	1.0%
Unsatisfactory z scores	1.4%

Compare performance over time



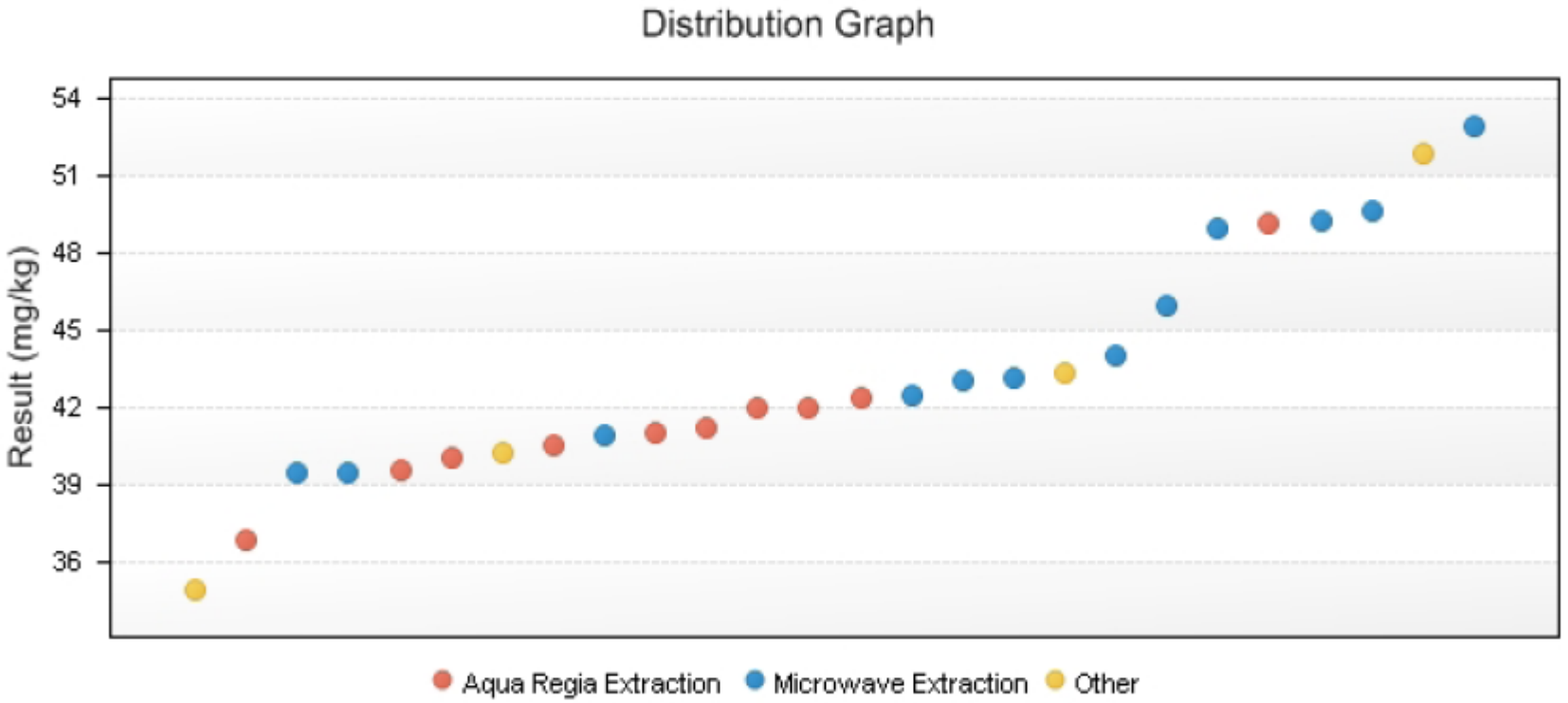
Compare analysts over time



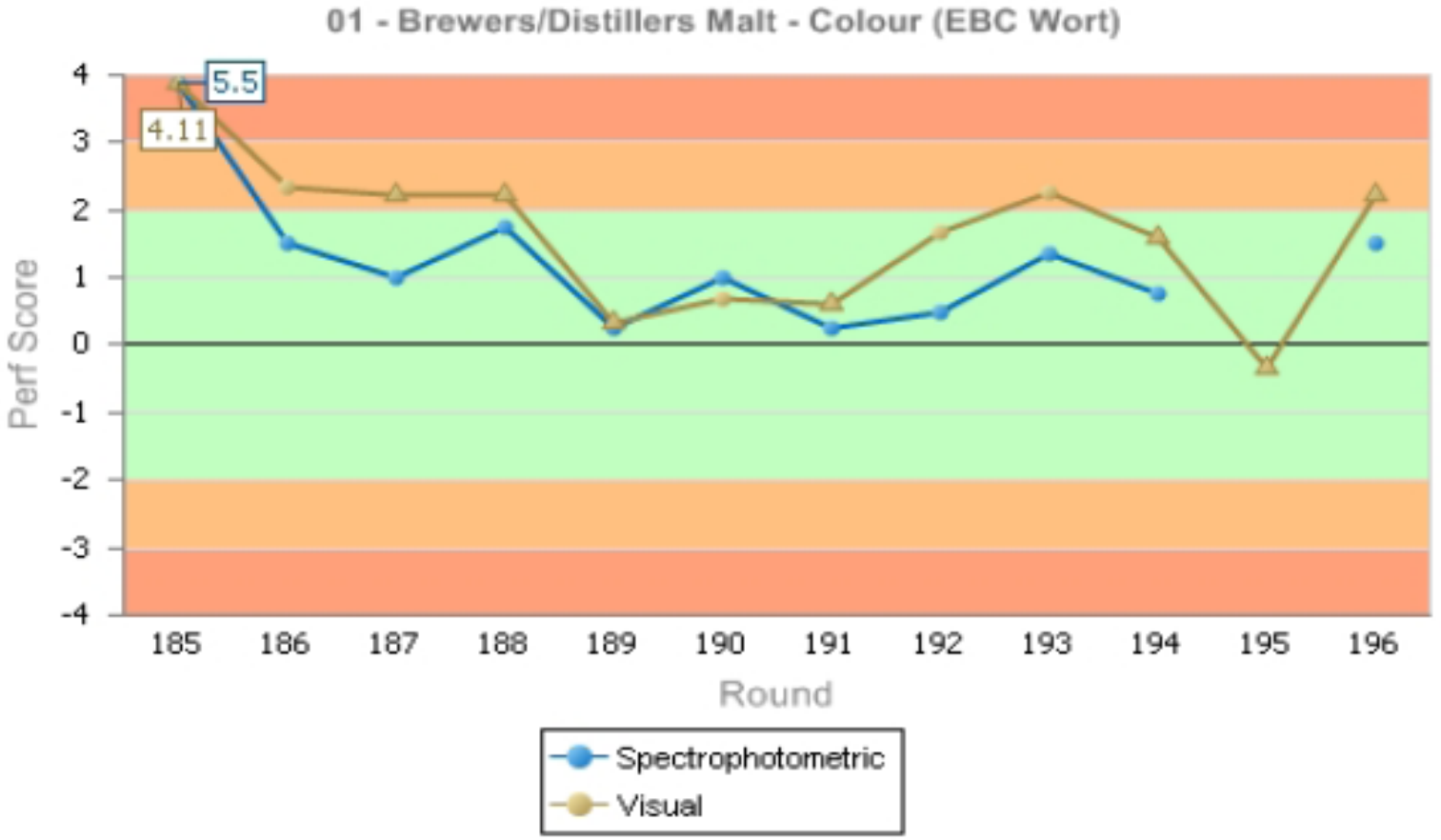
Compare method performance

Sample: 13 - Sewage Sludge Inorganics

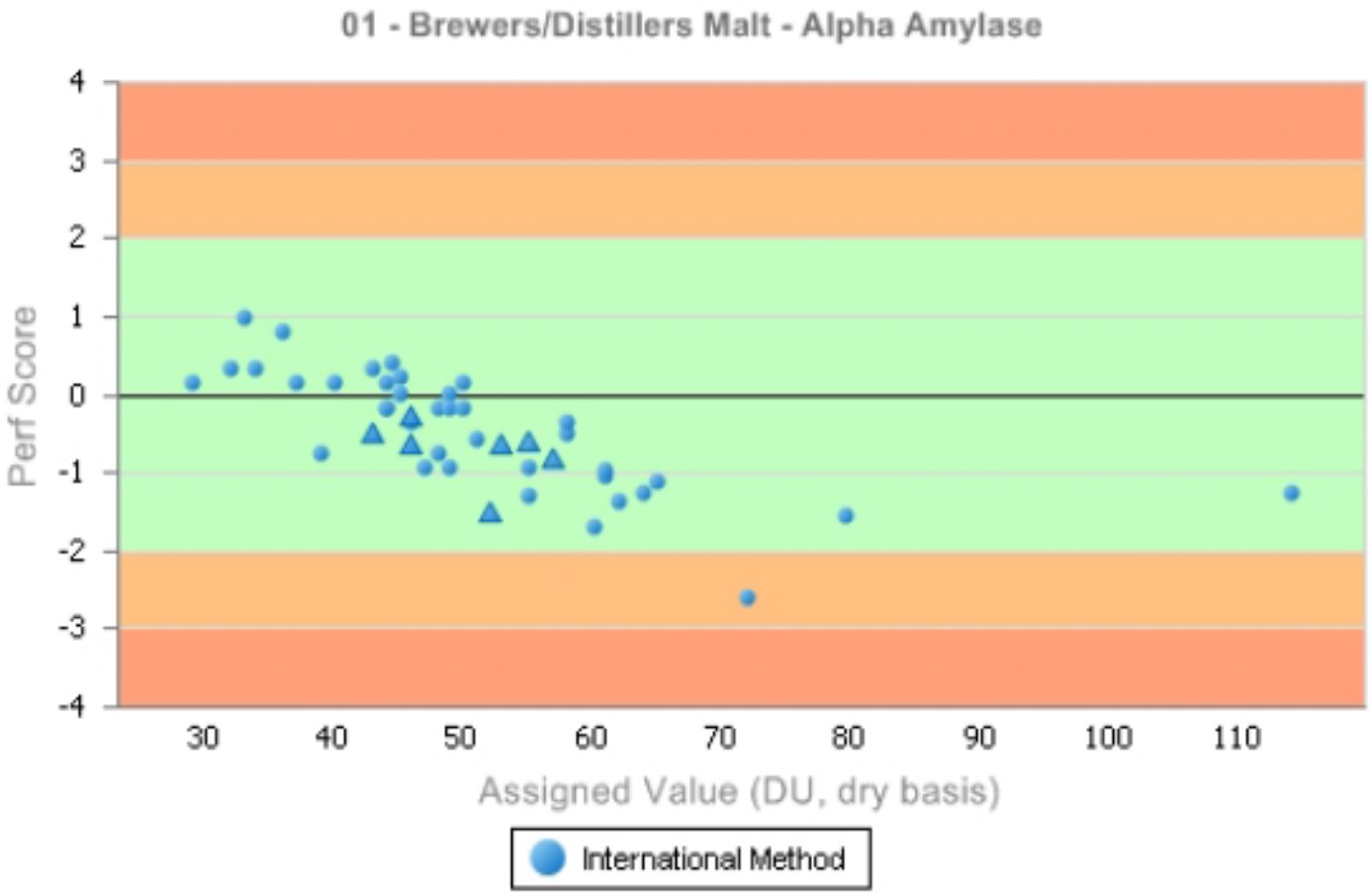
Analyte: Chromium



Method performance

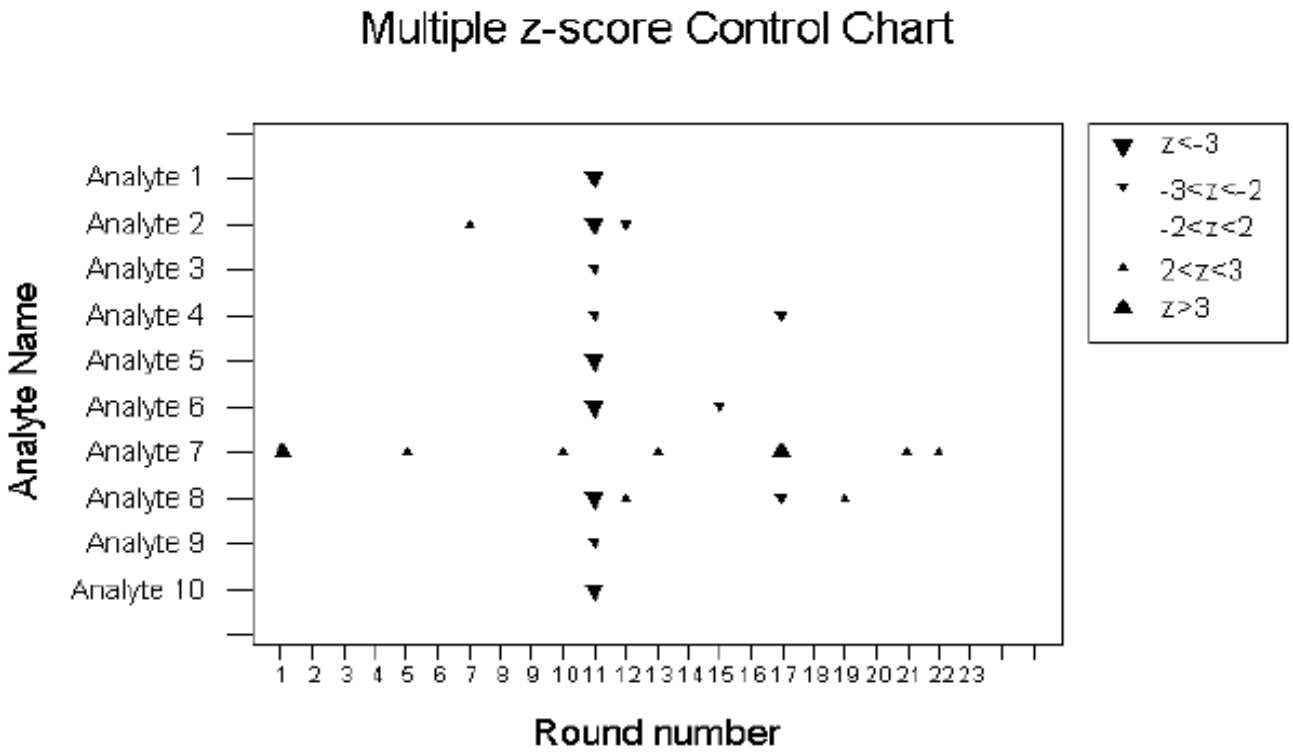


Monitor by concentration



Monitor by determinands

Fig A4.1. A control chart for z-scores.



How to get the best from PT

- PT is not about passing or failing, it is about learning from the result
- Compare the PT sample against your routine samples?
- Look at ALL the data and read the report
- Trend analyse using graphs, patterns are much easier to interpret than statistics….
- Don’t panic, everyone could be wrong except you!