An Introduction to UK NEQAS Guildford Peptide Hormones ABC Reports

20th July 2017

'ABC of EQA' is an ISO 17043:2010 compliant framework for the assessment of a laboratories analytical performance in a particular assay which meets and surpasses the utility of existing systems. The main benefit for participants, EQA Organisers, Steering Committees, Specialist Advisory Groups and the NQA Advisory Panels alike, is that it is a single system, which can allow meaningful comparisons to be made between analytes, schemes and disciplines.

External quality assessment (EQA) is intended to give an independent and objective assessment of the performance of the clinical assays undertaken in the participant's laboratory. This requires effective scheme design, including

- A reliable basis for assessment, with reliable specimens and valid target values.
- Sufficient recent data, achieved by frequent distributions and rapid turnaround of reports
- Effective communication through informative, intelligible reports and a scoring system showing performance over time

The reports for the UK NEQAS Guildford Peptide Hormone' scheme are structured so as to best utilise the 'ABC of EQA' scoring system, so you are able to see at a glance if your laboratory is performing well. If performance is acceptable, no further action is required. If performance is poor, you can probe further into the data presented. Similarly, you can see if you are performing in keeping with other users of your method and whether the method itself is performing well.

Definitions

There are three scores A, B and C

A is for Accuracy (total error)

B is for Bias

C is for Consistency of bias

They are conveniently referred to as the 'A score', 'B score' and 'C score', or simply A, B and C. Every laboratory in the scheme will have an A, B and C score for each analyte they measure and all 3 should be used when reviewing performance. Each of the 3 scores is calculated over a **rolling time-window** and thus comprises results from many specimens. They are always being updated with fresh current data, while older data drops out of the 'time-window'. The time-window has been set at 6 distributions. One of the main purposes of a performance score derived from many samples is to 'smooth out' the natural variation in deviations from target values over a number of distributions, by trimming extreme values and deriving a robust estimate of the central tendency for overall bias together with an index of its consistency. Thus when interpreting the performance score elements of

reports, it is important to note that a small number of atypical results are unlikely to affect overall scores, and aberrant results which are numerous enough to affect performance scores will take time to work their way out of the scoring 'window'.

For all UK NEQAS centres, **a low score is 'good', a high score is 'bad'**.

The A Score — Accuracy

The A score is weighted as part of a transformation process to take into account factors such as 'degree of difficulty' and normalised (median set at 100). The A score is primarily used as a quick 'comparator' or 'screening tool' for performance across all analytes. An A score of '100 is 'average', but this may of course be 'better' or 'worse' than what is required clinically, depending on the analyte. As more UK NEQAS schemes adopt the 'ABC of EQA' approach, the more useful the A score becomes in allowing broad comparisons to be made between analytes.

The **A score** tells you, on average, how good your overall performance is. This takes into account such factors as **bias, consistency of bias, degree of difficulty** etc. It has been transformed to ensure that **A scores** are broadly equivalent across analytes. For example, if you have an **A score** of 85 for Insulin and you also have an **A score** of 85 for Gastrin, this would indicate that you are performing both, on average, equally well.

The **A score** is an estimate of **accuracy [total error]** and is derived by taking the **Specimen % bias** and transforming it by a 'degree of difficulty' factor (see below) to get a **Specimen transformed bias** [this can be positive or negative]. The modulus of this **Specimen transformed bias** is then taken to give the **Specimen Accuracy Index** [as it is a modulus it has no sign]. Finally, the **'A score'** is calculated as the **trimmed mean** of all of the **Specimen Accuracy Indices** in the rolling time-window.

Because the A score is an across-analyte comparator, the limits used for the A score are common across all analytes, namely:

- Up to and including 100 (green on report)
- From 101 up to and including 200 (yellow on report)
- Greater than 200 is (red on report)

Degree of difficulty factor: The transformation itself has been empirically derived separately for each analyte and is based on modelling of data dependent on the concentration (target value) for the individual specimen. An examination of **the relationship between CV and target value** for the analyte was conducted to **derive an equation** for this relationship. This yielded the concentration-dependent factors used. **Normalisation** of the factors to yield a median (average participant) A score of 100 was then carried out.

The B and C Scores Bias and Consistency

The B and C scores (which have not been transformed) should be looked at together and provide analytical data on average bias and its consistency (pattern). The **B score** is Bias and therefore shows, on average (across the 6 distribution window), how far from the target you are and if you are running high or low.

The specimen % bias calculation **(specimen %bias)** is at the heart of all calculations:

$$\text{Specimen \% bias} = \frac{(\text{Result} - \text{Target})}{\text{Target}} \times 100$$

lf the target is 10 and you get a result of 11, then your bias is +10%; if the target is 10 and you get a result of 8, then your bias is -20%; if the target is 10 and you get a result of 10, then your bias is 0%, and so on. The **'B score'** is then calculated, (ie average bias), as the **trimmed mean** of all individual **'specimen %biases'** (including the sign) in the rolling time window.

The Consistency of bias or **C score** indicates, on average, if you usually have the same bias pattern. It is also not transformed and can assist in answering the following questions. 'Do you have different bias depending on the concentration of analyte in the sample?' 'Does your bias vary depending on the specimen matrix?' 'Has your bias changed during the time window?' 'Are you imprecise?' A high (poor) **C score** does not necessarily mean that you are imprecise, though if you **are** imprecise, it is impossible for you to have a very good (low) C score. Poor consistency of bias is **not** the same as imprecision.

The **'C score'** is simply the **standard deviation** (adjusted to take into account the degree of trimming) of the data which make up the B score.

The Standard Uncertainty

The Standard Uncertainty (SU;'u') statistic has now been added into our reports. The inclusion of this statistic is a requirement for UKAS ISO17043 Accreditation. The SU can be found to the right of the histograms.

The SU is calculated as 1.25*[SD/SQRT n]. The 'n' used in this calculation relates to a post trimming value (where appropriate) not the 'n' value listed on reports. The target is considered valid if 'u' is less than 0.3*SD. It is our reading of the algebra that, when you re-arrange the equations, if n<18 it is impossible to pass.

Limits of Acceptable Performance

A Score	
<100	Better than average performance (Green)
101 - 200	Worse than average performance (Amber)
>200	Worse than average performance (Red)

B Score	
<±25%	Acceptable performance
>±25%	Unacceptable performance

C Score	
<25%	Acceptable performance
>25%	Unacceptable performance

Participants will be defined as poor performers under the following circumstances:

- Failure to return for one distribution unless valid reason for non-return has been communicated to the Scheme Organiser
- Having an average B Score out- with the stated limits
- Having an average C Score out-with the stated limits

Participants will be defined as **persistent poor performers** under either of the following circumstances:

- A poor performer as defined above, compounds the errors by failing to make more returns or continues with a **B Score** out-with the limits
- **C Score** remains out-with the limits over further distributions.