Statistical analysis used in the PTS

The applied statistical calculations are based on the standard ISO 5725.

Analyses and statistics

For the statistical analyses a distinction is made based on the number of participants usin the same test system.

Less than 6 participants using the same test system
If there are less than 6 participants using the same test system, only the results as reported (quantitative and qualitative) are shown and the within-lab reproducibility sd is calculated for each participant.
Calculation of the within-lab reproducibility sd is only possible if the participant has reported quantitative results per sample in duplicate.

6 or more participants using the same test system
If there are 6 or more participants using the same test system, the results as reported (quantitative and qualitative) are shown.
The statistical calculations includes:

- table with z-scores,
- scatter plot of the z scores
- box plot
- trueness of the test results,
- within-laboratory reproducibility
- between-laboratory reproducibility

Before the statistical calculations were done for a group of laboratories using a particular test system, the values of outliers were adjusted using robust analysis (algorithms A and S). Outliers (CG-outliers) were not excluded from the calculations of the mean results of a group of laboratories using a particular test system, but taken into account using the adjusted value. Z-score outliers are presented in bold red in the tables.

Box plots are used to present results from all laboratories and for all samples for a specific test. The box plot consist of a box for each sample in which 50% of the results of all laboratories are contained, from Q1 (point below which 25% of the data) to Q3 (point below which 75% of the data). The difference between Q1 and Q3 is called Inter Quartile Range (IQR). Q2 is the median result from all participants. The vertical lines that rise and descend from the box represent the area with the data that are within Q3+ 1.5xIQR or Q1-1.5xIQR. Data that are between 1.5xIQR and 3xIQR away from the box are presented as a dot and laboratory code. Data that are more than 3xIQR away from the box are presented as an asterisk and laboratory code.
See for an example Fig 1.

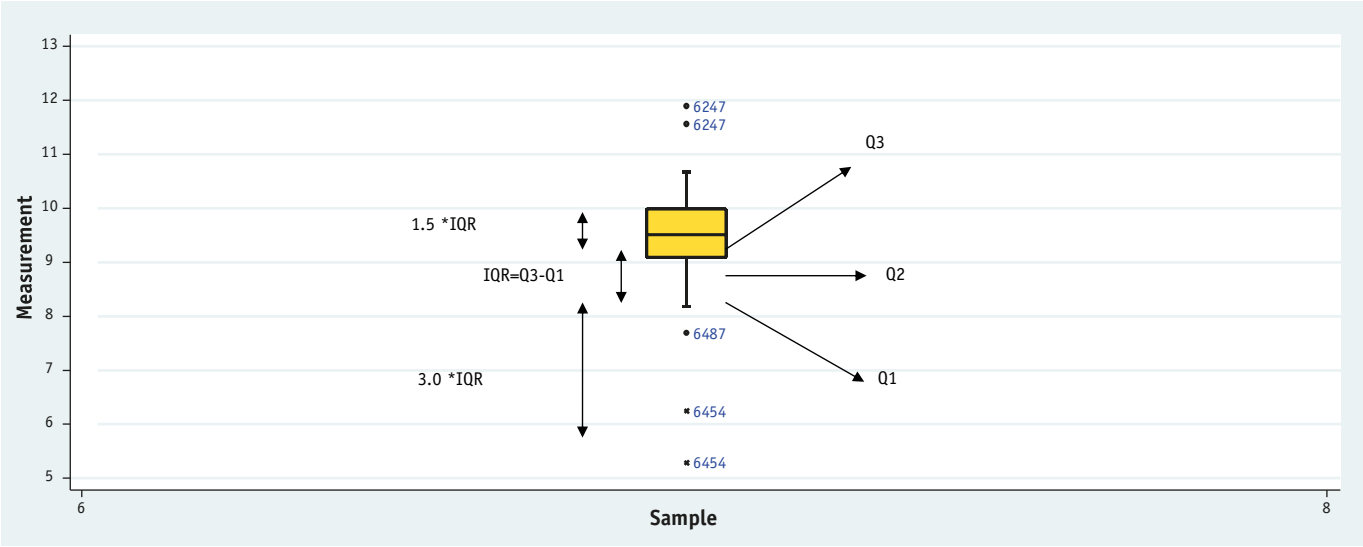

Fig 1. Example of Box plot

Definitions

The total mean result of a test system is the mean of all reported semiquantitative results for all samples and laboratories together.

The total mean value of a sample is the mean of all reported semiquantitative results for that sample by all laboratories using the same test system

The z-scores were calculated using the formula below (i):

(i) $z_{ij}=(y_{ij} - m_i) / s_i$ where z_{ij} = z-score for sample i, duplicate j
y_{ij} = measured value for sample I, duplicate j of laboratory X
m_i = mean value for sample i (all laboratories)
s_i = standard deviation for sample i (SD; all laboratories)

A z-score of the value of the serum below -2 or above 2 means that the result is more than 2 standard deviations different from the mean result of all laboratories for that serum using the same test system. Such a z-score is considered to be a z-score-outlier. Individual z-score-outlier results are presented in bold red color in the z-score tables. It is important to realize that the z-score is related to the SD value. If a test shows little variation between different laboratories (small SD) scoring a z-score of more than 2 (or less than -2) can be caused by a limited difference from the mean result. On the other hand, if a test shows a large difference between the results of different laboratories (high SD), scoring a z-score of more than 2 (or less than -2) will be caused by a substantial difference from the mean result.

The average z-scores (mean of the individual z-scores for all sera reported by a laboratory for a test system) are presented in XY diagrams or scatter plot. The scatter plot shows the difference expressed in z-scores for the average value of the first test run (z1) and the second test run (z2) compared to the average value of all laboratories using the same test system.

The trueness of the test results of a laboratory is defined as the difference of their total mean value (of all sera) compared to the total mean value of all participants using the same system. If the average of the reported values of a laboratory was significantly lower or higher than the total mean value for the same test, it is mentioned in the text and results are in **bold red**.

The within-laboratory reproducibility sd (sR_{within}) (ISO 5725-2) of test results of a laboratory is defined as the square root of the mean variances of the duplicates of the test results. This means that a lower sR_{within} represents a higher within laboratory reproducibility!

The between-laboratory reproducibility sd ($sR_{between}$) is defined (ISO 5725-2) as the square root of the mean variances in reported results by all laboratories using the same test system. This means that a lower $sR_{between}$ represents a higher between laboratory reproducibility!