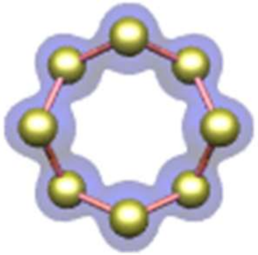
Errors in Chemical Analysis

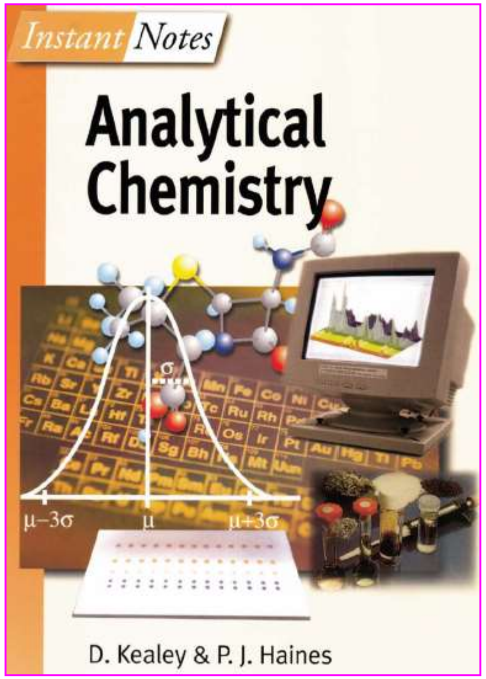
Syllabus:

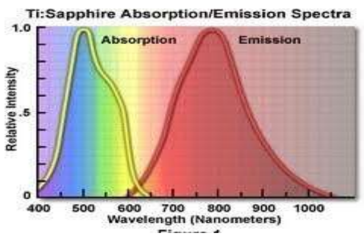
Errors in chemical analysis

Classification of errors

Determination of accuracy of methods – improving accuracy of analysis



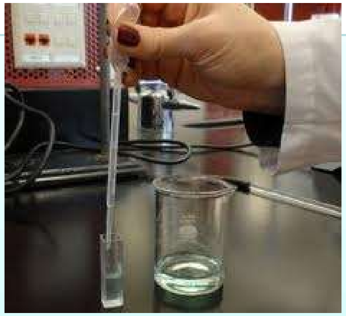




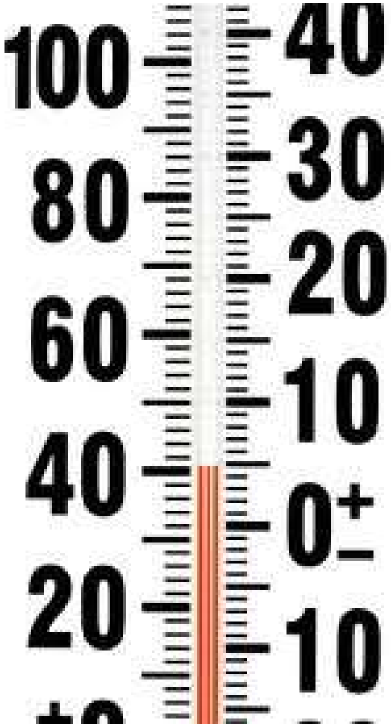
Dr. K. SIVAKUMAR
Department of Chemistry
SCSVMV University
chemshiva@gmail.com

Measurement Errors:

Measurement errors:



- are numerous and their magnitudes are variable
- leads to uncertainties in reported results
- can be minimized and some types eliminated altogether by careful experimental design
- effects can be assessed by the application of statistical data analysis and chemometrics
- can be avoided by proper equipment maintenance and appropriate training of personnel



Absolute & Relative errors:

Absolute Error: is the numerical difference between measured value and the accepted (True) value

Simply subtract the measured value from the accepted answer $E_A = x_M - x_T$

E_A – Measurement error

x_M – Measured value

x_T - True or accepted value

Relative Error (also, percent error): is the ratio between absolute error and true value

$$E_R = (x_M - x_T)/x_T$$

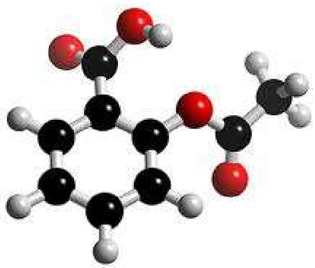
E_R – Relative error

x_M – Measured value

x_T - True or accepted value

Example: 200 mg aspirin standard has been analyzed a number of times.

- absolute errors range from -4 mg to +10 mg.
- absolute error may be used to express the inaccuracy in a measurement.



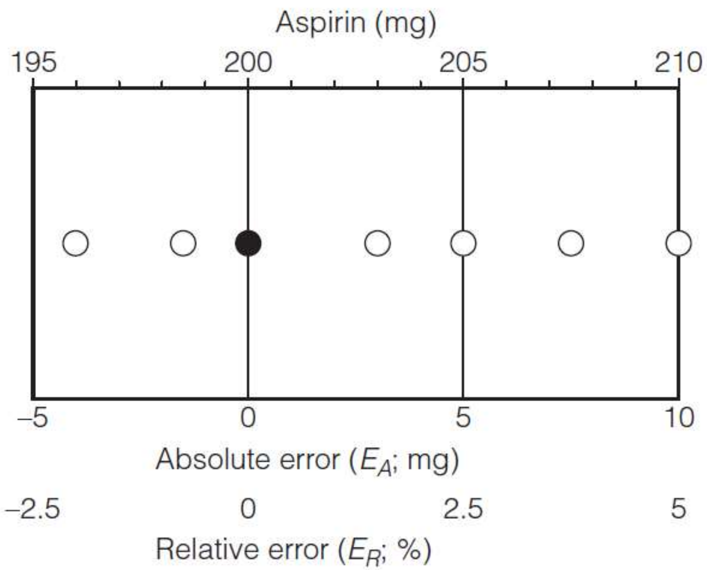
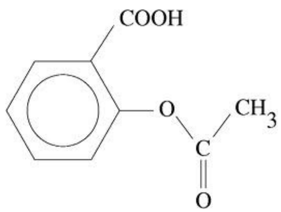


Fig. 1. Absolute and relative errors in the analysis of an aspirin standard.



- the relative error ranges from -2% to +5%.
- relative errors are particularly useful for comparing results of differing magnitude.

Determinate Errors:

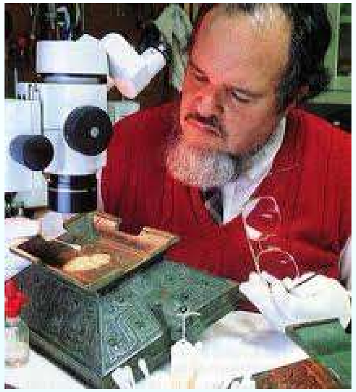
The basis of determinate or systematic errors arises from three sources:

- the analyst or operator;
- the equipment (apparatus and instrumentation) and the laboratory environment;
- the method or procedure.

It can be rectified by

- careful observation and record keeping
- equipment maintenance
- training of laboratory personnel

Determinate Errors: Reasons…



Operator errors: can arise through
- Carelessness
- insufficient training
- disability

Equipment errors: are due to
- substandard volumetric glassware
- faulty or worn mechanical components
- incorrect electrical signals
- poor or insufficiently controlled laboratory environment

Method or procedural errors: are caused by
- inadequate method validation
- Improper methods and concentrations that affect measurements.

Determinate errors that lead to a HIGHER value than a true value are said to show a positive bias;

Determinate errors that lead to a lower value than a true value are said to negative bias.

Determinate errors can be proportional to the size of sample taken for analysis.

Determinate Errors: Example

Cu^{2+} can be determined by titration after reaction with potassium iodide to release iodine,

$$\text{i.e., } 2Cu^{2+} + 4I^{-} \longrightarrow 2CuI + I_2$$

However, the reaction is not specific to Cu^{2+}, and any Fe^{3+} present in the sample will react in the same way.

Results for the determination of copper in an alloy containing 20%, but which also contained 0.2% of iron are shown below, ***for a range of sample sizes***.

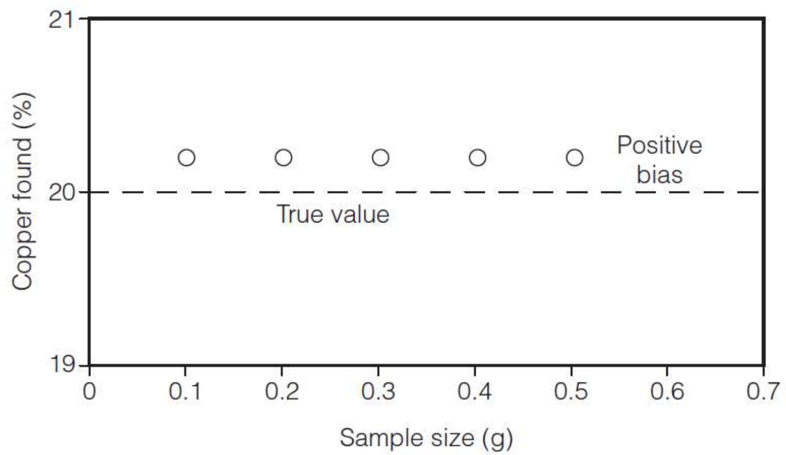


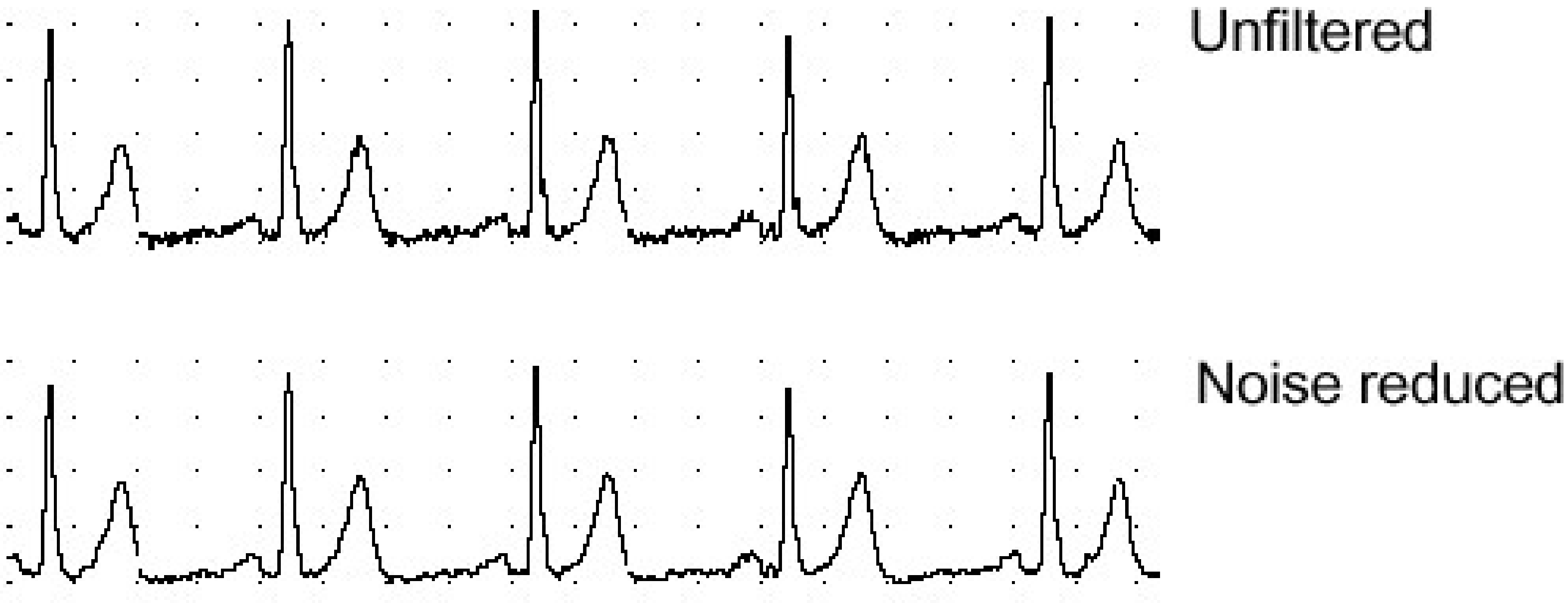
Fig. 2. Effect of a proportional error on the determination of copper by titration in the presence of iron.

The same absolute error of +0.2% or relative error of 1% (i.e. a **positive bias) occurs regardless of sample size, due to the** presence of the iron.

This type of error may go undetected unless the constituents of the sample and the chemistry of the method are known.

Indeterminate Errors: Reasons

- This type of error arises **random fluctuations in** measured quantities which always occur even under closely controlled conditions.
- It has an EQUAL chance of being POSITIVE or NEGATIVE
- It can be **REDUCED** by careful experimental design and control but never **ELIMINATED**.
- Environmental factors such as temperature, pressure, & humidity
- Electrical properties such as current, voltage, & resistance are some of the factors producing random variations described as NOISE**.**
- These contribute to the overall indeterminate error in any physical or physico-chemical measurement, but no one specific source can be identified.



Accumulated Errors: (Also called as propagation errors)

- Errors are will be aggregated in the final calculated result.
- The accumulation is treated similarly for both determinate (systematic) and indeterminate (random) errors.
- Determinate (systematic) errors can be either positive or negative
- Cancellation of errors is likely in computing an overall determinate error
- In some instances this may be zero.

- The overall error is calculated using one of two alternative expressions, that is

Overall ABSOLUTE determinate error E_T

$$E_T = E_1 + E_2 + E_3 + \ldots\ldots.$$

E_1 & E_2 etc., being the ***ABSOLUTE determinate errors in the individual*** measurements

Overall RELATIVE determinate error, E_{TR}

$$E_{TR} = E_{1R} + E_{2R} + E_{3R} + \ldots\ldots.$$

E_{1R} & E_{2R} etc., being the ***RELATIVE determinate errors in the individual*** measurements

ACCURACY

Degree of agreement between measured value and accepted true value.

i.e., Being near to true value

PRECISION

Degree of agreement between replicate measurements of same quantity.

i.e., Repeatability of a result

Measures of accuracy is,

Absolute Error:

$$E_A = x_M - x_T$$

E_A – Measurement error

x_M – Measured value

x_T - True or accepted value

&

Relative Error:

$$E_R = (x_M - x_T)/x_T$$

E_R – Relative error

x_M – Measured value

x_T - True or accepted value

Relative error is more useful in practice

Useful for measuring deviation from mean

$$d_i = |x_i - \bar{x}|$$

Where,

d_i - deviation

x_i - any individual value in the population

$\bar{x}$ - Experimental mean

Accuracy & Precision

ACCURACY

Degree of agreement between measured value and accepted true value.

i.e., Being near to true value

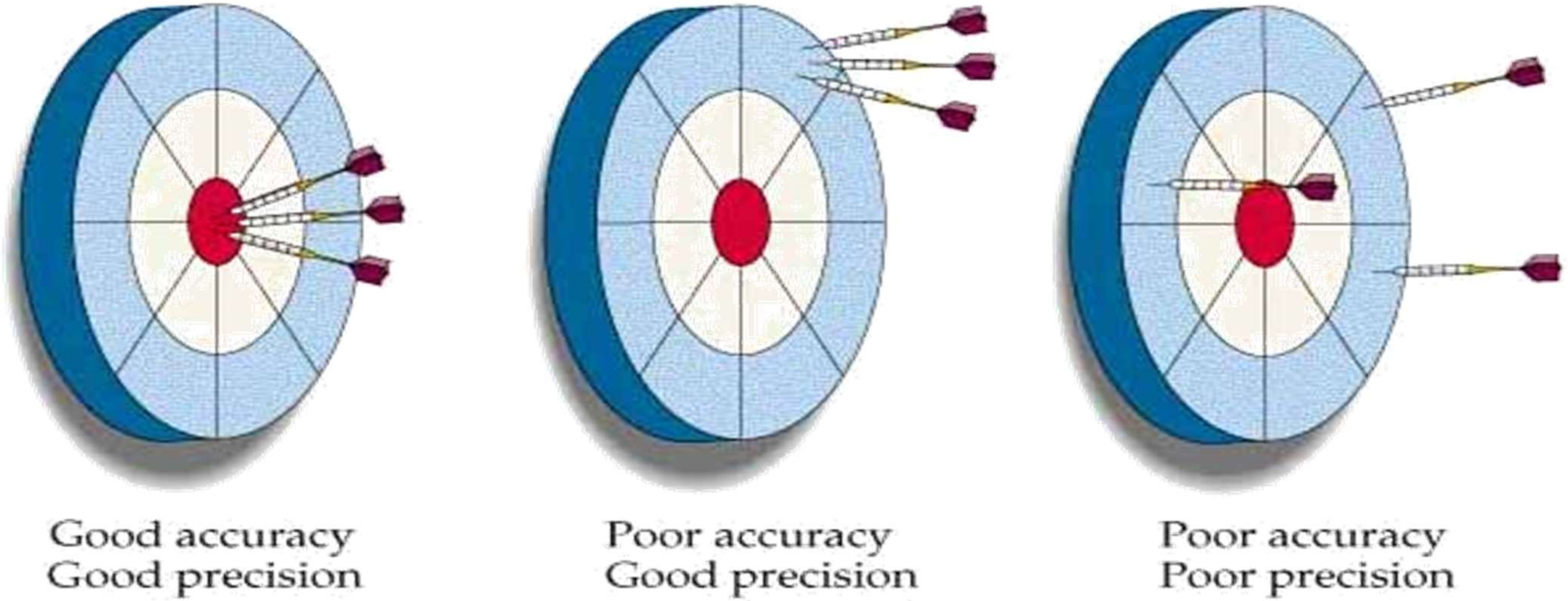
PRECISION

Degree of agreement between replicate measurements of same quantity.

i.e., Repeatability of a result

Measurements can be

Accurate & Precise | Precise but Inaccurate | neither accurate nor precise



Accuracy & Precision: Example

- Four analysts have each performed a set of five titrations for which the correct titer is known to be 20.00 cm3. The titers have been plotted on a linear scale, and inspection reveals the following:

- the average titers for analysts B and D are very close to 20.00 cm^3 – these two sets are therefore said to have **good accuracy**;
- the average titers for analysts A and C are well above and below 20.00 cm^3 respectively – these are therefore said to have **poor accuracy**;
- the five titers for analyst A and the five for analyst D are very close to one another within each set – these two sets therefore both show **good precision**;
- the five titers for analyst B and the five for analyst C are spread widely within each set – these two sets therefore both show **poor precision**

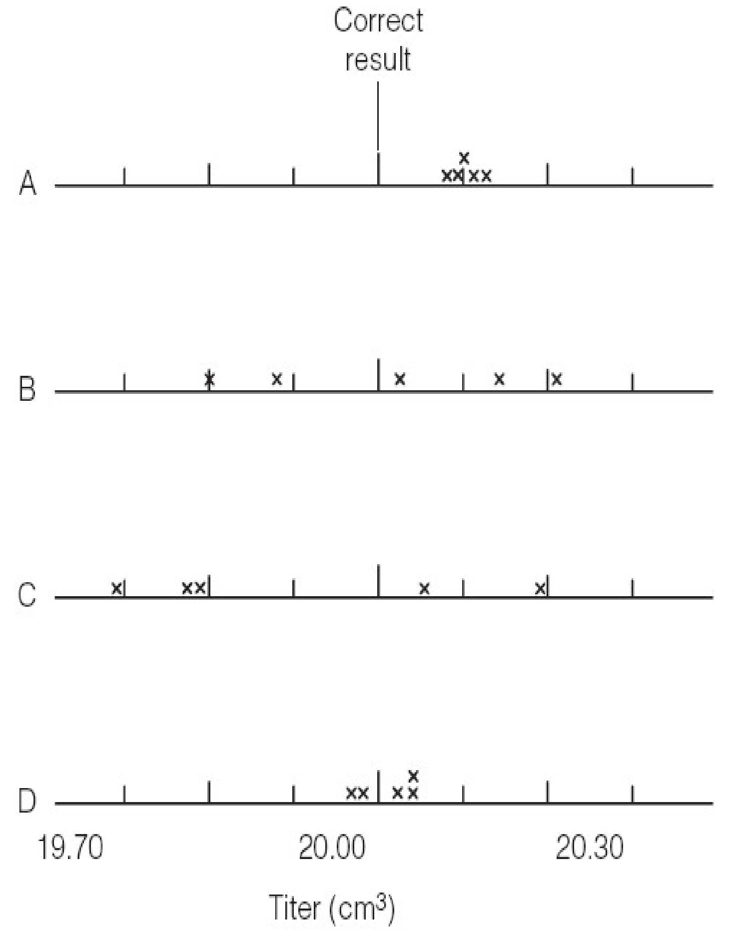


Fig. 1. Plots of titration data to distinguish accuracy and precision.

Standard Deviation:

- A measure of the width of the distribution
- Small standard deviation gives narrow distribution curve

$$s = \sqrt{\frac{\sum_{i=1}^{i=N}(x_i - \bar{x})^2}{N-1}}$$

Where,

s	- Standard deviation
x_i	- any individual value in the population
$\bar{x}$	- Experimental mean
N	– total No. of values
N-1	– No. of degrees of freedom is defined as the number of independent deviations $(x_i - x)$ used to calculate s
Σ	- Sum for i = 1 to i = N

Calculation of an estimated standard deviation, *s*, involves the following steps:

- calculation of an experimental mean;
- calculation of the deviations of individual x_i *values from the mean;*
- squaring the deviations and summing them;
- dividing by the number of degrees of freedom, *N - 1, and*
- taking the square root of the result.

Example 1

	x_i/cm^3	$(x_i - x)$	$(x_i - x)^2$
	20.16	−0.04	1.6×10^{-3}
	20.22	+0.02	4×10^{-4}
	20.18	−0.02	4×10^{-4}
	20.20	0.00	0
	20.24	+0.04	1.6×10^{-3}
Σ	101.00		$\mathbf{4 \times 10^{-3}}$
$\bar{x}$	**20.20**		

$$s = \sqrt{\frac{4 \times 10^{-3}}{4}} = 0.032\ \text{cm}^3$$

$$s = \sqrt{\frac{\sum_{i=1}^{i=N} (x_i - \bar{x})^2}{N - 1}}$$

Relative Standard deviation (RSD) or Coefficient of Variation (CV):

Measure of relative precision and is normally expressed as a percentage of the mean value

$$s_r = (s/\bar{x}) \times 100$$

$$s_r = \frac{0.032}{20.20} \times 100 = 0.16\%$$

Example 1

	x_i/cm^3	$(x_i - x)$	$(x_i - x)^2$
	20.16	−0.04	1.6×10^{-3}
	20.22	+0.02	4×10^{-4}
	20.18	−0.02	4×10^{-4}
	20.20	0.00	0
	20.24	+0.04	1.6×10^{-3}
Σ	101.00		$\mathbf{4 \times 10^{-3}}$
$\bar{x}$	**20.20**		

$$s = \sqrt{\frac{4 \times 10^{-3}}{4}} = 0.032\ \text{cm}^3$$

Pooled Standard deviation:

Where replicate samples are analyzed on a number of occasions under the same conditions, an improved estimate of the standard deviation can be obtained by pooling the data from the individual sets.

$$s_{pooled} = \sqrt{\frac{\sum_{i=1}^{i=N_1}(x_i-\bar{x}_1)^2+\sum_{i=1}^{i=N_2}(x_i-\bar{x}_2)^2+\sum_{i=1}^{i=N_3}(x_i-\bar{x}_3)^2+\ldots+\sum_{i=1}^{i=N_k}(x_i-\bar{x}_k)^2}{\sum_{i=1}^{i=k}N_i = k}}$$

Where, $N_1, N_2, N_3 \ldots\ldots\ldots N_k$ are the numbers of results in each of the k sets

$\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots\ldots\ldots \bar{x}_k$ are the means for each of the k sets.

Variance:

Variance

- Is the square of the standard deviation

Variance = s^2

Overall Precision:

- Random errors accumulated within an analytical procedure contribute to the overall precision.
- Where the calculated result is derived by the addition or subtraction of the individual values, the overall precision can be found by summing the variances of all the measurements so as to provide an estimate of the overall standard deviation, i.e.

$$s_{overall} = \sqrt{\left(s_1^{\,2} + s_2^{\,2} + s_3^{\,2} + \ldots\right)}$$

Median, Range & Probability:

Median:

- The middle number in a series of measurements arranged in increasing order.
- The average of the two middle numbers if the number of measurements is even

Range:

- The difference between the highest and the lowest values

Probability

- Range of measurements for ideal Gaussian distribution
- The percentage of measurements lying within the given range