

VARIANCE COMPONENTS. ESTIMATION OF REPEATABILITY, BETWEEN RUN IMPRECISION AND TOTAL IMPRECISION.

2018-07. Anders Kallner. Dept Clin Chem Karolinska university hospital, Stockholm Sweden.
anders.kallner@ki.se

Table of Contents

Summary.....	1
Scope	1
General	2
Experimental design	2
Data input.....	2
Results and reports.....	4
Graphics	4
Imprecision	4
Trueness	5

Summary

A reference material or patient sample is measured repeatedly in many "runs". This experimental design allows estimation of within run imprecision, equivalent to repeatability, the between run imprecision (intermediate) and the combined or intra-laboratory imprecision, reproducibility. Two or more concentrations should be measured in a verification procedure. If an assayed material or any other material with a known concentration is used this information can provide information about the bias at the given concentrations.

Short instruction

- A sample dataset is provided (*2018-07 ACB Precision (imprecision) and trueness (bias) - example data.xlsx*) comprising results from measurements of an analyte with a low and one with a high concentration; a typical setup for a verification according to the accreditation standard EN/ISO 15189.
- Open the spreadsheet program and copy all or parts of the results and "Paste special", "Values (123)" into the spreadsheet. It is important that the special pasting option is used to preserve the format of the spreadsheet.
- You can now delete, add or change any of the results.
- The highest (red) and lowest (green) results are marked.
- The graphs have now been populated. Make sure that the Y-axes are optimized, the default is set to "automatic". Choose features of the graphs e.g. display SD or SEM.
- Enter analyte identification in cells C7 and M7. Unless this has been made no calculation will be initiated.
- If any potential outlier has been identified a warning is issued in cells Y41 and Z41.
- ID of the runs is optional and specifications of the study (rows 2 to 5).
- Depending on the aim of the experiment target data can be added.

Scope

This spreadsheet program is intended for verification of imprecision based on repeated measurements of the concentration of a defined sample. It allows up to ten repeats in ten runs for two samples (two reference materials). It handles unbalanced designs, i.e. different number of repeats in each run. It performs the estimation of within group, intermediate and intra-laboratory variation, based on analysis of variance components (ANOVA). Each run is evaluated and the

combined variance calculated. If the value of the reference material is known the bias can be estimated and reported. It also compares the outcome with specifications and evaluates the result.

General

Save an “archive” copy of the spreadsheet in a safe place!

Calculations will not be performed unless the name of the component is properly entered on the *Indata* sheet (cells C7 and M7).

All formulas and calculations are protected and only cells with a blue or violet border or blue text can be changed.

Note that any entry can be copied or deleted and new data entered, whereas ‘Cut-and-paste’ is not allowed, nor ‘Drag-and-drop’. These operations may jeopardize the correct function of the spreadsheets. Copy, paste and “clear contents” (i.e. use the “delete” button) are acceptable and will do no harm. In case data are pasted, use the ‘Paste special’ and ‘Value function’. This will ensure that the lay-out e.g. lines and colours remain unchanged.

Experimental design

The setup requires measurements of the same sample in several different series. A series is usually defined as measurements made within a certain time, a work shift or between calibrations. The results will then be within- and between series and the total represents the intra-laboratory variation.

A minimum of five repeats in five series is usually required for acceptable results. A series can also be defined as repeated measurements made in one laboratory if an inter-laboratory comparison or checkup is made. The results will then represent within-, between- and inter-laboratory variation. In this case the experimental setup needs to be carefully evaluated before the analysis of variance components is evaluated.

The template allows up to 10 repeated measurements in up to 10 series.

If the value of the material is known, the bias can be evaluated, e.g. by using assayed control material or patient material measured by a reference procedure.

Data input

Select a stable material of which there is enough for many measurements. It is recommended to pool or otherwise prepare enough material, which is then distributed into separate vials. This design should be suitably customized for the component under study with regards to stability and homogeneity.

There are two sectors of input representing two concentrations. The data are treated similarly in both sectors. All cells with blue borders are open for data input. The name of the component should be entered in cells C7 and M7. No calculations will take place unless the component is identified. The runs may optionally be identified (e.g. by date) in the blue cells in row C8 to V8 (Figure 1).

Operator: XYZ Date(s): 2011 Reagents: xxx Calibrator: xxx										Instrument: ABC Comments:										Operator: Date(s): Reagents: Calibrator:										Instrument: Comments:									
Replicate	DATE	Group (day, run or instrument)																																					
		FT4 Lvl1										FT4 Lvl2																											
		09-06-20	06-21	06-22	06-23	06-24						09-06-20	06-21	06-22	06-23	06-24	06-25																						
	1A	2A	3A	4A	5A	6A	7A	8A	9A	10A	1 B	2 B	3 B	4 B	5 B	6 B	7 B	8B	9B	10B																			
1.	9.50	9.58	8.98	9.88	9.49						43.70	45.53	43.50	44.20																									
2.	10.33	9.65	9.04	10.60	10.25						42.90	44.50	43.00	43.63																									
3.	9.70	9.00	9.96	10.06	9.73						43.95	44.40	46.90	43.85																									
4.	9.85	10.50	10.00	10.60	10.23						43.00	43.90	45.48	43.64																									
5.	9.76	9.66	9.66	10.20	9.59						44.40	44.70	40.00	44.12																									
6.																																							
7.																																							
8.																																							
9.																																							
10.																																							

Figure 1. Data input matrix. The maximum (red) and minimum (green) values are marked.

Target imprecision values are entered in Y27 and Z27 for the first and second component (Figure 2), if available. If the target imprecision is expressed in relative terms (CV%) then that should be entered in Y29 and Z29. The relative imprecision will be recalculated to a standard deviation. The standard deviation takes precedence over the relative standard deviation. The rejection rate (α) should be entered. This defaults to 5 %.

The spreadsheet checks for outliers using Grubbs' test (see Figure 2). The Grubb's value can be set to 1 % or 5 %, corresponding to a probability of 99 % and 95 % for identification of the outlier. Therefore 5 % is the more stringent condition. Groups may contain different number of observations – an unbalanced design is thus accepted.

Claimed repeatability (SD):		0.80		FT4 Lvl1	FT4 Lvl2
Calc SD	0.32	1.10			
Claimed repeatability (CV%):	3.3%	2.5%	Target value:	10.0	20.0
False rejection rate (α %):	5.0	5.0	U_{target} (SEM):		0.5
Verification value:	0.42	1.07	Calc SEM:	0.3	
Conclusion:	Accept	Review!!	U_{target} (CV%):	3.0%	
Claimed intra-lab (SD):			Bias:	0.17	-23.96
Calc SD	0.69	1.49	U_{bias} :	0.09	0.58
Claimed intralab (CV%):	7.0%	3.40%	z-score:	-1.84	41.15
False rejection rate (α %):	5.0	5.0	α %	5	5
Verification value:	0.90	1.98	t-value (k):	2.064	2.093
Conclusion:	Accept	Accept	High ver value:	10.6	21.2
Critical Grubbs value (1/5)	5	5	Low ver value:	9.4	18.8
Upper outlier			Conclusion:	Accept	Reject!
Lower outlier		Check!			

Figure 2. Input and evaluation of specifications for imprecision and for bias evaluation

If the quantity value (concentration) of the used material is known the collective data can be used to estimate the bias and evaluate acceptability in view of the specifications. The target value and its uncertainty (SEM) are entered in AC7 and AC8 and AD7 and AD8, respectively (figure 2). The uncertainty of the target value can be made in absolute or relative terms. As for the imprecision evaluation, the standard deviation takes precedence over the relative standard deviation (CV). Entered data can be deleted or entered at any time and in any position and the results will be continuously updated.

Results and reports

Graphics

The graphs are generated simultaneously with the input of results (Figure 3). The X-axis will be adjusted to fit the number of runs entered.

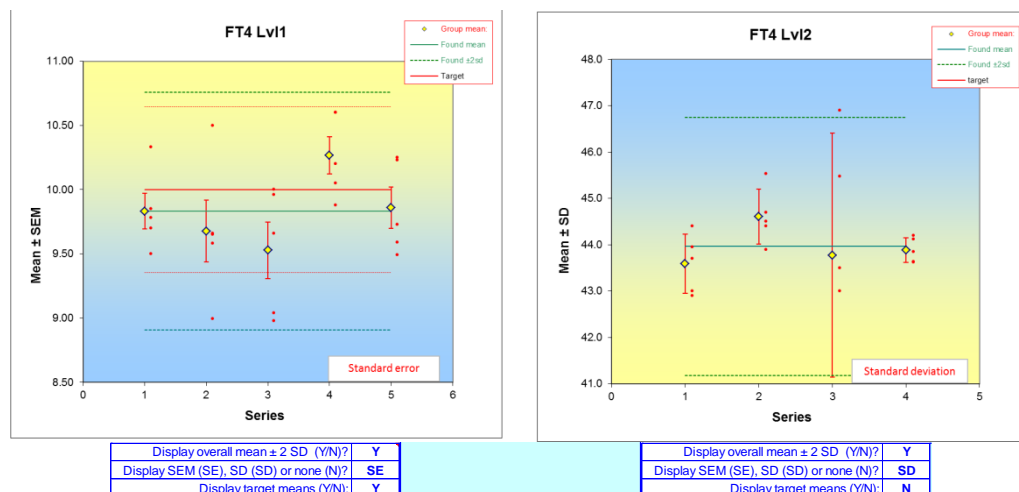


Figure 3. Graphical display of the data. Averages and SD deviations and all individual observations and their overall average and standard deviation are shown. The graph shows the target data for the samples and the allowable variation, see figure 2.

There are several options for the display: the individual results and their average for each group are always shown. The imprecision can be shown as SD or SEM, the average of the observations and ± 2 SD can be chosen and finally the target value and acceptance limits can be displayed, provided a target value has been given. The options are toggled between using the cells below the graph.

Imprecision

The data is evaluated by analysis of variance components. The calculation is based on sums of squares and “mean squares”, MS, that can be obtained by a one-way ANOVA. This procedure is designed to estimate the within- and between-group variances and the sum of these constitute the intra-laboratory variance.

As shown in Figure 4, the results of measurement of two concentrations will be presented in separate columns. Basic variables are calculated, and the within-run (repeatability) variance, standard uncertainty and relative standard uncertainty are shown, followed by the between run (intermediate) and intra laboratory variances, standard uncertainty and relative standard uncertainty.

The analysis of variance components requires adjustment of the between-run variance with respect to within-run variance. In cases when the *Within* variance is larger than the *Between* variance, the intra-laboratory variance is equal to the within-run variance, thus the between-variance is conventionally set to zero. In those cases, the mean square of the between groups is shown in brackets in the columns. Details of the calculations will be presented elsewhere.

	FT4 Lvl1	FT4 Lvl2
Between group df:	4	3
Within group df:	20	16
Number of observations:	25	20
Mean:	9.83	43.96
SEM:	0.09	0.30
Average of average of series	9.83	43.96
Average of series' SD	0.41	1.39
Repeatability variance:	0.1718	1.932
Intermediate variance:	0.0420	(.3)
Intralaboratory variance:	0.2138	1.932
Repeatability (SD):	0.4144	1.390
Intermediate imprecision (SD):	0.205	
Intralaboratory imprecision(SD):	0.462	1.390
Repeatability (CV%):	4.2	3.2
Intermediate imprecision (CV%):	2.1	
Intralaboratory imprecision (CV%):	4.7	3.2

Figure 4. Results of the analysis of variance components. Note the number in brackets corresponding to the intermediate variance in the second column. This indicates that the MS_{bet} was less than the MS_{with} and the intra laboratory variance is then equal to the MS_{with} .

The method of calculation allows an unbalanced design, i.e. different numbers of observations are possible in the groups. Only if the number of observations in the group differs the displayed average will differ from the overall average.

If the claimed specifications are entered, the estimated repeatability and intra-laboratory variation are tested against this data using χ^2 evaluation. The outcome of the comparison will be reported as “Accept” or “Reject” (Figure 2).

At the bottom of the panel in Figure 2, the results of the Grubb’s test for outliers are presented. In this test the maximum and minimum values of the entire study are identified and their deviation from the average evaluated. “Check!” is displayed in case the calculated Grubbs’ value exceeds that of the table value. The maximum and minimum values are flagged in the input area and may be deleted. It is not advisable to perform Grubb’s test again after removing a suspect value.

Trueness

If the target value of the used material is known together with its uncertainty, the trueness can be estimated. The target value of the reference material should be entered in cells AC7 and AD7, respectively. The uncertainty of these values is equivalent to the standard uncertainty of the average, i.e. the SEM. This is the characteristic that manufacturers are supposed to supply but it can also be estimated from PT/EQA values or measurements by a reference procedure. The high and low allowable limits of the results of the measurements are estimated from a t -distribution. If the target value is within these limits the result is *Accepted* otherwise *Rejected* (Figure 2).

The uncertainty of the target value can be entered either as the absolute value or relative. As described above the absolute value takes precedence over any relative value and therefore if the relative value is the relevant then the absolute must be deleted. The target value and its standard deviation can be displayed in the graphs and the bias visualized.