CombiStats Online – Training
Q&A session
18-22 November 2024

Module 1: Introduction to the online application

Question Asked	Answer Given
Do you plan to do this online training again?	Yes, in Q1 2025.
Would you please give me an e-mail address I can send questions to in future?	You can use the EDQM HelpDesk.
A pop-up message says that CS-Desktop will run for 28 more days after 31 October. Will CS-Online be available only after this period?	If you renew your licence, you will get CS-Desktop until 31 March 2025. CS-Desktop will not be supported after this date.
Is there a user guide providing information about features implemented in CS-Online?	There are 24 notes for guidance, which can be accessed via the CombiStats Help page.
How will you inform users about version updates?	A note will appear 1 month before the update in a yellow banner at the top of the homepage. Once implemented, the changes will be listed in the release notes.
Is there a limit to the amount of data that can be excluded?	You can exclude an entire data table if you want. There are no limitations (except that the model selected may not be meaningful anymore).
I am only a manager of user accounts (no need to have access to the data). Would it be possible to not pay a licence?	Thank you for your comment. However, it won't be possible to implement such a change in the next releases.
Can a user create an assay template, or just the manager?	Anyone can create a template.
Can we import templates from CS-Desktop to CS-Online?	Yes. Note that the protection level will be removed and that you will have to redefine a protection level in CS-Online (L0 to L3).
Can you modify a template (for example the standard has a new assigned value), and have that tracked through the audit/revisions?	No, you need to create a new template (from a copy of the current template, for example). This new template will have new metadata (e.g. creation date)
Are templates also deleted after 3 years?	No, templates have to be deleted manually.
Is there an easy way to see how the protection level is set on a template?	Yes, on the right of the page, in the list of revisions.

How is the data protected from unauthorised third-party access, since it is based online?	Please read the security notice for further information, and also accessible from the Help page.
I think the report is too long.	Thank you for your comment. This will be improved in a future version.
I thought there was GMT time zone in the profile options; why is it not working?	This has not yet been implemented in CS-Online.
Is the date and time shown on the report based on the creation of report or when an assay was run?	It is the date of creation of the report. The last page contains the list of revisions with their corresponding dates.
Can we create a report without publishing the assay?	No, it needs to be published (saved) first.
Is it possible to include the graphics in the report?	Yes, they are included by default.
How can you export to PDF?	You just need to click the Report button.
Is it possible to print results directly from CS-Online?	No, you need to export the PDF report first.
Is it possible to export the results as plaintext (.txt) file?	No, you need to export the PDF report.
Can you copy and paste within the Value field of the Potency columns?	No, values have to be entered one by one.
Is it possible to customise the data tables?	Yes, data tables can have a different number of rows and/or columns.
Should you always write ? IU/ml for an assumed potency, or can one add an approximate value?	For an assumed potency, an approximated value can work. It will not interfere with the potency calculation.
Why not showing the audit trail in the exported .epax file?	The exported file may be modified outside CS-Online.
Will (parts of) the audit trail be seen on the printout?	The printout includes the entire audit trail.
Is it possible to perform a blank subtraction (blank average) automatically?	No, the subtraction needs to be done manually.
How can I report a result like "no calc" (e.g. OD > 4.00)?	You cannot enter a string character in the data tables. In this scenario, there are two options. Option 1: let the cell empty and put a comment in the remark field. Option 2: enter 4.00 and exclude the value (double click on the cell).
Is the now available R^2 calculated from weighted or unweighted r?	Both R^2 estimates can be calculated, depending on the type of regression model used.

What does SOP mean?	Standard operating procedure.
If we have several samples in one assay, can we copy/paste the standard data from one assay file to another?	Yes, you can use the Ctrl+C and Ctrl+V keyboard shortcuts. Another option is to test all the samples against the standard in a single assay file.
Does the option "Dose = automatic" produce equidistant doses?	Yes.
Can doses reported in ng/mL, for example, be reported as inverse dilutions instead?	Yes.
Can you explain how to enter pre-dilutions?	Please refer to note for guidance 05. If you need more help, please do not hesitate to send a request via the EDQM HelpDesk.

Module 2: Assays based on quantal response

Question Asked	Answer Given
Is the linearised plot the same as the parallel lines analysis?	The sigmoid curves are linearised using Probit transformation. The linearised plot represents the data after transformation. The interpretation of these plots is similar to the parallel lines analysis for the evaluation of parallelism and linearity.
Why are non-parallelism and non-linearity calculated instead of parallelism and linearity?	This is due to the statistical testing approach. In the ANOVA table, lack of parallelism and lack of linearity are tested.
If non-linearity or non-parallelism fail, what could be the reason for it?	It could be a true lack of linearity or parallelism in the selected dose range (another range may be more appropriate), an outlier disturbing the regression analysis, a bias due to a lack of randomisation of the trials or, for non-parallelism, a difference in the nature of the standard and sample tested.
Can you use the option "theoretical variance" instead of "observed residuals"?	For quantal responses, "theoretical variance" is the only available option.
Can the statistical calculation be run even when reported results are the same for the standard and test preparations?	Yes.

Module 3: Assays based on quantitative response

Question Asked	Answer Given
What does “Treatments” mean in the ANOVA table?	A treatment is a combination of preparation and dose. If you have 2 preparations x 3 doses each, you have 2 x 3 = 6 treatments. If the signal-dose relationship is significant (significant slope), a low *p*-value will be observed in the ANOVA table. Information on treatments is provided by default in the ANOVA table, with no specific added value. Please refer to note for guidance 11 for more information.
Why would Preparations be significant (*p*-value <= 0.05) for some assays and not significant for others (*p* > 0.05)?	In a parallel lines analysis, Preparations (in the ANOVA table) is about differences between regression intercepts (i.e. relative potency). If you test a batch vs a standard and it is expected that the batch is more potent than the standard, then Preparations should be significant. If you assay 2 batches of the same product, they should have more or less the same potency, so Preparations should not be significant.
Why are the analysis options fixed starting from the first level of protection, so it is not possible to change the data transformation on a protected template?	This is because a given data transformation is already foreseen in the template for routine use. If another transformation needs to be performed, you can make a copy of the .epa file.
How do you get the quadratic curvature in the ANOVA table?	By selecting “extended ANOVA” in the drop-down list.

Does quadratic fit always fit x^2 (meaning, y = intercept + x + x^2), and Lack of Quadratic fit always fit x^3 (is it y ~ intercept + x + x^2 and x^3 in the model)?	Quadratic fit tells whether x^2 is significant, and lack of quadratic fit indicates an even more complex trend might exist. If the number of doses allows it, it may not be limited to x^3 – it can also contain x^4, x^5, etc. Of course, such higher polynomial models are usually unlikely... Note. x = dose for SRA; log(dose) for other models. When you have many dilution points, you can fit - A linear reg: $y = a + b\,x + error$ - A quadratic model: $y = a + b\,X + c\,x^2 + error$ - A cubic model: $y = a + b\,x + c\,x^2 + d\,x^3 + error$ - Etc… The non-linear contrast tells you if all these additional polynomial terms (x^2, x^3, x^4) are significant or not. If the contrast is significant, it is common to look at x^2 (quadratic curvature) first, because it is unlikely that higher degrees (x^3, x^4) would be practically meaningful. If it is significant, then $y = a + b\,x + c\,x^2 + error$ would be an option and "Lack of quadratic fit" tells you whether further terms (x^3, x4, etc.) would improve the fit even more. If not (Lack of quadratic fit *p*-value > 0.05), then you can stop at $y = a + b\,x + c\,x^2 + error$. But, in a parallel lines analysis, the expected model is $y = a + b\,x + error$. A data transformation can help improving linearity (if x^2 is significant occasional). Otherwise, another range of doses where the dose-response relationship is more linear could be selected.
In a 4PL model, after log transformation of data, the residual plot still shows a pattern (a wave pattern, so not completely random). What else can be done to fit a better model to the data?	It is hard to say. A data transformation is not very common for sigmoid curves. A closer look at the curves would help. It could be that the method needs some improvement.
The 4PL model cannot have 0 and 1 as lower and upper asymptotes. How does the linearising work? Does it somehow transform the data to a 0 to 1 range?	The general 4PL equation is y=D+(A-D)/(1+exp(B*(ln(dose)-C))), where D and A are the lower and upper asymptotes. The estimation of all 4 parameters is done in an iterative process. Asymptotes are equal to 0 and 1 for quantal responses only and can take on other values for quantitative data. In addition, following the linearisation, the y-axis is not limited to [0 - 1].
When applying log transformation to the 3-parameter exponential model, does it always assume D=0 and A=1?	Not necessarily.
Are potency estimates and confidence limits based on the constrained model or unconstrained model?	They are based on the constrained model. There are no final results calculated using the unconstrained model in CombiStats.

What do LCL and UCL represent for assumed and assigned potencies?	For assumed potencies: LCL/UCL% relative to the calculated potency are reported, providing information about the precision of the calculated potency. For assigned potencies: LCL/UCL% relative to the assigned potency are reported, providing information about the recovery of the calculated potency.
In the parallel lines model, why do the confidence limits change slightly depending on the number of samples analysed?	This is linked to the fact that the variance used to calculate the confidence limits is estimated from the sample data. If the number of samples changes, the dataset, and thus the variance estimate, changes.
What is difference between non-linearity table 1, non-linearity table 2, vs. just non-linearity? what is non-linearity measuring?	The non-linearity is tested for each preparation. Table1 refers to the preparation data entered in table 1, and so on. Non-linearity measures the lack of fit for linearity and is a pool of the non-linearity assessed on the various tables.
In case of non-linearity due to a low residual error, is there a way to recalculate the *p*-value?	Yes, you could enter a known residual error in the variance drop-down list (e.g. "user-defined variance s^2 = ").
A significant non-linearity can be observed when the confidence interval of the potency estimate is very narrow. What is the reason for this?	This is likely to be due to a low residual error. However, for quantal responses, a lack of linearity is not very common.
Sometimes we do not comply the non-linearity criteria but we do not have curvature in the data. How can we improve it? Theorical variance, maybe?	Yes, you could enter a known residual error in the variance drop-down list (e.g. "user-defined variance s^2 = ") if you know that it is practically meaningless and due to a low residual error. Alternatively, the non-linear trend may be clearly visible and improved by a data transformation (e.g. sqrt or log). The flowchart in the following document (CS-Desktop) may also help.
Is the logarithmic transformation log10 or ln? In some figures, I see ln, but in the options, I see log.	The default transformation y' = log(y) calculates logarithms with base 10. The User-defined transformation can be used for all other logarithm bases. For example: ln(y) or log2(y) for natural or base-2 logarithms.
For equivalence testing, the ratio or difference must fall into a defined range. Has the CI also to be inside this range?	Yes, if not, you are only 50% confident that equivalence is demonstrated. To be 95% confident, the confidence limits should be within the margins of equivalence too.
How can I narrow the confidence interval of the equivalence test (non-parallelism assessment)?	Assuming proper selection of the dose range (doses resulting in 0% and 100% are of limited interest) and proper randomisation, an increase in the sample size (e.g. more replicates per assay or more than one assay) can make a real difference.
Are there commonly accepted values for the margins of equivalence?	There is no indication in Ph. Eur. The margin should be defined in advance by the user.

Can I change the confidence level of the equivalence test?	No, the default is 90%.
If the confidence level of the equivalence test is 90%, how could you say that you have 95% confidence in the decision taken?	A two-sided 90% CI is calculated in CS-Online, which is the same as two one-sided 95% CIs. Our recommendation would be to read the standard operating procedures (SOPs) again to check whether the 95% confidence level applies to the lower or upper confidence limit only. If this is the case, CS-Online provides the confidence limits requested by the SOPs.
Does the subset analysis allow different dose ranges for the standard and test samples?	Yes, if you select *subset = shift*.
If 4 doses must be used according to the Ph. Eur. method, is it OK to report data from 3 doses (subset analysis)?	No, I do not think so.
If we are running an assay based on parallel lines analysis that requests 4 different doses, is it possible to remove 1 or 2 doses, applying the subset analysis and still maintain the compliance of the test with the Ph. Eur.?	It is not possible to specify the number of doses you would like to keep in the final assay when applying the subset analysis. If the monograph requires a minimum of 4 doses, then a subset analysis starting with say 5 doses could be envisaged, provided that, if the 5-dose assay is invalid, it is highly likely that one of the 4-dose assays tested as part of the subset analysis will be valid.
Is the subset of doses based on non-linearity and non-parallelism?	Yes, both.
What is the impact of the use of non-equidistant doses in a model (both parallel lines and slope ratio)? It is a matter of complexity of calculation or the use of an inaccurate approximation by the software for the calculation?	Equidistant doses usually result from serial dilutions using multichannel pipets (constant dilution step). There is no added calculation issues for statistical software nor approximation made in case of non-equidistant doses. On the other hand, when assessing non-linearity on 3 doses, for example, it would make sense from a statistical viewpoint to use equidistant doses. So non-equidistant doses are possible but may require a bit more justification than equidistant doses, especially for indirect assays.
For an 8x8 Latin square design in a plating bioassay, replicates are not independent. Can parallel lines analysis be used? Should an ANOVA or equivalence testing method be used?	Can we say that they are not independent? Let's assume, for example, that you did 8 independent preparations of each of the 8 samples. While it is true that they will be tested on the same gel, the Latin square design should mitigate the potential bias due to heterogeneity of the substrate in the gel. So yes, parallel lines analysis can be used. The choice between ANOVA or equivalence does not depend on the above-mentioned considerations, but rather on their ability to detect non-parallelism of practical relevance.
What is meant by variance = observed residuals?	The variance is calculated using (squared) distances from data points to the regression line. Such distances are called "residuals".

Is replacing the test variance with a theoretical variance as recommended by Ph. Eur. general chapter *5.3* a good approach or not?	Our recommendation would be to use the test variance on a routine basis and to use a theoretical variance only in some specific situations (such as a run with a low variance).
For parallel line and slope ratio models with 3 replicates per dose, is it possible to exclude one replicate from each dose to meet the requirement for parallelism?	Options to exclude replicates should be described (anticipated) in your standard operation procedure. For an assay you are experienced with, outliers should not be that common. Although outliers may happen at any dose, an assay with 1/3 of outliers would thus be unlikely, unless they all belong to the same series. You should discuss with your QA representative if such an assay would still be valid.
Can we replace outliers?	I do not think so: we can delete them or keep them (and run a robust regression to alleviate their potential negative effect on the regression analysis).
Should outlier(s) identified in the residual plot be removed, even though the assay is valid?	An outlier may inflate the residual error such that statistical tests in the ANOVA table may lack sensitivity (an assay wrongly declared as valid). Possible actions in presence of outliers should be anticipated in the method standard operating procedure. Several options are possible, including deletion and robust analysis.
Can values of a sample be excluded automatically when the corresponding CV is above a certain threshold?	No, this should be done manually.
How many replicates are needed to evaluate the data distribution?	The number of replicates is usually not enough in an assay for a formal assessment of the data distribution. The evaluation is usually performed using the residual plot.

Module 4: Single-dose assays, combination of assay results

Question Asked	Answer Given
Is it possible to use the semi-weighted combination in a 4PL model in CS-Online?	Yes.
In CS-Desktop, the unweighted combination was indicated when there were enough assays (at least 6), whereas in CS-Online, it is indicated when assays are not independent. Can you clarify please?	This statement can be found in the user manual of CS-Desktop and is most likely motivated by the fact that when there are less than 6 assays, the t-value used to calculate the confidence limits of the mean potency is rather large (e.g. 12.7 for 2 potency values, 4.3 for 3 potency values), resulting in a wide confidence interval (that may not fulfil some requirements). But in fact, the selection of a combined approach does not depend on the number of assays to be combined.
If assays are independent, do you only choose between weighted and semi-weighted combinations regardless of the number of assays?	Yes.
Which combination (geometric or arithmetic) can be applied for the slope ratio model?	The arithmetic combination would be logical, because confidence intervals of potency results to be combined are symmetrical. However, if the distribution of these potency results would be skewed, a geometric combination may be a better option.
Is the *p*-value of homogeneity of test results displayed?	Yes, the *p*-value appears above the table of calculated means.
If the *p*-value of homogeneity is equal to the significance threshold (e.g. 0.10), are the test results homogeneous or not?	When the *p*-value is equal to the significance threshold, the null hypothesis is rejected. In that case, homogeneity would be rejected.