WLTR



WHITE PAPER · CALIBRATION IN PRACTICE

REPLICATES · INSIGHTS 28

Averaged replicates or individual points: the setting that hides a failing level.

Triplicate standards can be fitted one injection at a time or averaged into one point per level. The line often comes out the same. The review does not.

Walter Hinchman, MBA

CO-FOUNDER & CEO, WLTR, INC.

NOVEMBER 2026

WLTR.TECH

SUMMARY

Many labs inject calibration standards in replicate, and most data systems can fit either every injection or the mean response at each level. This paper sets out what ICH M10 and EPA Method 8000D say about replicate standards, then works an illustrative seven-level triplicate curve both ways. The fitted line is identical. The acceptance picture is not: one level passes when averaged and fails ICH M10's replicate condition when read injection by injection.

THE SHORT VERSION

01 With complete replicates and concentration-based weights, averaging gives the same fitted line as individual points.

02 It does not give the same review: a mean can pass while most of the injections behind it fail.

03 ICH M10 judges replicates injection by injection, so choose the approach in the SOP and review at that level.

Two reviewers, one data set.

Two analysts review the same triplicate calibration. One sees seven points on the curve, every one inside its limit. The other sees twenty-one injections, and two of them at the same level read +16% and +17%.

Both are looking at the same data. The difference is a single processing option: fit the mean response per level, or fit every injection. Which one your lab uses decides what the reviewer is able to see.

What ICH M10 and 8000D say about replicates.

ICH M10 allows each calibration standard to be analysed in replicate. Data from all acceptable replicates should be used in the regression [1]. It then sets the acceptance test at the level of the injection:

01 **Accuracy.**
Back-calculated concentrations within ±15% of nominal, and ±20% at the LLOQ.

02 **Overall.**
At least 75% of the calibration standards, with at least six levels, meet the criteria.

03 **Per level.**
With replicates, the criteria should also be met by at least 50% of the standards tested at each concentration level.

04 **Range ends.**
If replicates are used and only one of the LLOQ or ULOQ standards fails, the calibration range is unchanged.

EPA Method 8000D is less prescriptive. When a standard fails its refit criterion, that one standard may be reanalysed immediately to rule out random error [2]. That is a replacement of a suspect injection, not a licence to average it away.

Two ways to build the same curve.

Most data systems offer a version of the same choice. Build the curve from the average response at each level, or from every calibration injection as its own point. Some also let a new injection replace an old one, or add to it.

The choice is often made once, when the processing method is first set up, and never looked at again. Our paper on what CDS calibration settings compute covers how settings like this hide in plain sight.

Seven levels in triplicate, read both ways.

Here is an illustrative curve: seven levels from 1 to 100 ng/mL, three injections each, fitted linear with $1/x^2$ weighting. It was processed both ways.

LEVEL (NG/ML)	EACH REPLICATE, %RE	REPLICATES WITHIN LIMIT	AVERAGED POINT, %RE
1	+3% −12% +11%	3 of 3 PASS	+0.5% PASS
2	−7% +2% −1%	3 of 3 PASS	−1.6% PASS
5	+1% −8% −1%	3 of 3 PASS	−3.0% PASS
10	+16% −3% +17%	1 of 3 FAIL	+10.1% PASS
25	−5% −0% +2%	3 of 3 PASS	−1.2% PASS
50	−1% −6% −0%	3 of 3 PASS	−2.5% PASS
100	−4% +1% −3%	3 of 3 PASS	−2.2% PASS

Illustrative data set constructed for this article, not instrument data. Limits are ±20% at the LLOQ and ±15% elsewhere, per ICH M10. Both fits use $1/x^2$ weighting.

+10.1%
AVERAGED POINT AT 10 NG/ML, INSIDE ±15%

1 of 3
REPLICATES WITHIN ±15% AT 10 NG/ML

19 of 21
INJECTIONS WITHIN LIMIT OVERALL

The two fits give the same slope and intercept, to every digit the CDS will report. That is expected: with the same number of replicates at each level and weights that depend only on concentration, averaging does not move the line.

What moves is the verdict. The averaged point at 10 ng/mL reads +10.1% and passes. Behind it, 2 of the 3 injections are outside ±15%. That level does not meet M10's 50% condition, even though 19 of 21 injections pass overall.

The statistics move too. r^2 is 0.9980 on the individual injections and 0.9990 on the averaged points. Averaging removes the scatter from the plot, so it removes it from the summary numbers as well.

Averaging does not change the line. It changes what the reviewer is able to see.

When averaging changes the line itself.

The two approaches stop agreeing as soon as the replicate count differs between levels. Drop one injection, for a documented reason, and the averaged fit still gives that level a full point. The individual fit gives it two thirds of one.

In the example, remove the +17% injection at 10 ng/mL. The two slopes now differ by 0.51%. That is small here. It grows with fewer levels, wider ranges and weights based on response.

The rule of thumb is simple. If anything about the replicates is unequal, the two settings give two different curves. The lab needs to know which one it reported.

Decide before the data, in the SOP.

ICH M10 says the selection of the regression model should be directed by written procedures [1]. How replicates enter the regression is part of that model. It belongs in the SOP, chosen before the data exist.

Whichever approach the lab fits, review at the injection level. Report every replicate's back-calculated concentration, count the passes per level, and keep the record of any injection that was replaced or excluded, with the reason.

A replicate standards checklist.

01 **Find the setting.**
Confirm whether the processing method averages, replaces or adds replicates.

02 **Write it into the SOP.**
One approach, chosen before the data, with the reason.

03 **Back-calculate every injection.**
Not only the mean at each level.

04 **Count passes per level.**
Against the 50% condition where M10 applies.

05 **Watch unequal replicate counts.**
They make the two approaches diverge.

06 **Report r^2 on the same basis.**
Averaged and individual r^2 are not comparable.

07 **Record every replacement.**
Which injection, why, and when.

Where WLTR fits.

A replicate set carries more information than its average. Seeing every injection back-calculated, under every model, is what keeps a failing level from passing as a mean.

WLTR is calibration data integrity infrastructure for regulated analytical laboratories. For each analyte, it derives nine calibration regression models: average response factor, plus linear and quadratic fits across four weightings each. It scores every model against fifteen ICAL parameters, including RSE. All of it is computed orthogonally to the CDS, under U.S. Patent 12,038,448 B2 [P].

CALIBRATION VERIFICATION

See every calibration standard back-calculated under every regression model and weighting, computed orthogonally to your CDS.

GET COMPLIANCE GAP ASSESSMENT

WLTR.TECH

SOURCES

1. ICH, M10: Bioanalytical Method Validation and Study Sample Analysis, final version adopted 24 May 2022, Secs. 3.2.4 and 3.3.2. www.pmda.go.jp/files/000246792.pdf
2. U.S. EPA, SW-846 Test Method 8000D, Determinative Chromatographic Separations, Revision 5, March 2018, Sec. 11.5.6.4. www.epa.gov/sites/production/files/2015-12/documents/8000d.pdf
3. U.S. Patent 12,038,448 B2, inventor Walter “Chip” Hinchman Sr. patents.google.com/patent/US12038448B2

Walter Hinchman, MBA

WLTR, INC. · NEVADA C-CORP

WLTR