

WLTR · WHITE PAPER · BIOANALYSIS



When incurred sample reanalysis fails, check the calibration.

An ISR pair compares two results from two analytical runs, and each run has its own calibration curve. When failures cluster near the LLOQ and all point the same way, compare the two curves before you investigate the samples.

ISR

8 MIN READ

OCTOBER 2026

PUBLIC

DOCUMENT	VERSION	DATE
WLTR Insights 16	1.0	October 2026
PREPARED BY	ORGANIZATION	CLASSIFICATION
Walter Hinchman, MBA	WLTR, Inc. · Nevada C-Corp	Public
REV 1.0 · OCTOBER 2026 · FIRST PUBLICATION · WLTR.TECH		

THE SHORT VERSION

- 01 ICH M10 requires at least two-thirds of chromatographic ISR repeats to fall within $\pm 20\%$, with the percent difference calculated against the mean of the original and repeat results.
- 02 Published reviews attribute many ISR failures to human error and procedural issues and name instability, back-conversion and inhomogeneity among the other causes, yet every pair also rests on two calibration curves.
- 03 A low-end model or weighting bias can fail pairs near the LLOQ while both runs pass their calibration criteria, so WLTR recommends comparing the two runs' curves at the levels that bracket the failing samples.

01 • ISR TESTS THE REPORTED RESULT, NOT THE VALIDATION

ISR tests the reported result, not the validation.

M10 states the purpose: “ISR is intended to verify the reliability of the reported sample analyte concentrations” (Section 5)¹. Calibration standards and QCs are prepared by spiking blank matrix, and study samples can differ through protein binding, metabolite back-conversion, inhomogeneity, concomitant medications or unique sample components¹.

Under Section 5, ISR should be performed at least in nonclinical studies (in general, at least once per species), all pivotal comparative BA/BE studies, the first clinical trial in subjects, pivotal early patient trials (once per population) and the first or pivotal trial in patients with impaired hepatic and/or renal function¹. Samples are reanalyzed “in separate (i.e., different to the original) runs on different days using the same bioanalytical method,” within the stability window but not on the day of the original analysis. Selection is predefined and as random as possible, covers C_{max} and the elimination phase, and excludes pooling. The subset is at least 10% of up to 1,000 study samples, plus 5% of those above 1,000.

The percent difference is (repeat value – initial value) ÷ mean value × 100. “For chromatographic methods, the percent difference should be within $\pm 20\%$ for at least 2/3 of the repeats” (Section 5); for LBAs the limit is $\pm 30\%$ ¹. The Bioanalytical Report includes an ISR table with original and reanalysis values, run IDs, percent difference and percent passed (Section 8, Table 1)¹. FDA issued M10 as final guidance in November 2022².

02 • A FAILED ISR REQUIRES AN INVESTIGATION

A failed ISR requires an investigation.

“If the overall ISR results fail the acceptance criteria, an investigation should be conducted and the causes remediated” (Section 5)¹. An SOP directs how investigations are triggered and conducted; if no cause is found, the potential impact on study validity goes into the Bioanalytical Report. When ISR passes but “shows large or systemic differences between results for multiple samples,” Section 5 advises further investigation, for example when all ISR samples from one subject, or one run, fail¹. Individual flyers (above 50%, for example) need not be investigated, and ISR data should not replace the original data¹.

FDA's M10 Questions and Answers (June 2024) add that the investigation should take into account “the entire process, including sample handling, processing, and analysis,” and method issues “such as interferences and instability” (Q&A on Section 5)³. In WLTR's view, both runs' calibration curves belong in that scope.

03 • ISR BEGAN AS A WORKSHOP RECOMMENDATION

ISR began as a workshop recommendation.

The third AAPS/FDA Bioanalytical Workshop met in Arlington, Virginia, on May 1 to 3, 2006^{4,5}. Its report recorded “a major concern and issues raised with respect to stability and reproducibility of incurred samples” and left the approach to scientific judgment (Conclusion, item 3)⁴. Rocci and colleagues then proposed sample numbers, sample types and data evaluation⁶.

An AAPS workshop on assay reproducibility for incurred samples followed in 2008⁷. Its report calls the 2006 meeting Crystal City III, notes that it endorsed demonstrating assay reproducibility with incurred samples, and records the steering committee's ISR recommendations⁷. The European Bioanalysis Forum issued recommendations in 2009 for integrating ISR across a method's lifecycle⁸, and the Global Bioanalysis Consortium reviewed repeat analysis and ISR practice for harmonization in 2014⁹.

04 • PUBLISHED ROOT CAUSES ARE PROCEDURAL OR METHODOLOGICAL

Published root causes are procedural or methodological.

At Eli Lilly, fewer than 5% of ISR samples failed, mostly through human error, though ISR also exposed stability and ruggedness issues¹⁰. At GSK, 26 studies from more than 1,000 pharmacokinetic end points failed ISR; five datasets (three partial, two full) were rejected for instability or preanalytic contamination, and other investigations found unexpected within-study assay variability, an inappropriate assay range and sample homogeneity¹¹. The 2018

Workshop on Recent Issues in Bioanalysis agreed that most issues were procedural rather than scientific¹². In simulations, pass rates fell significantly at higher incurred sample imprecision, and larger sample sizes did not restore them¹³.

ISR FAILURE CAUSE	WHAT YOU WOULD SEE	WHAT TO CHECK
Metabolite back-conversion ^{1, 12}	One-directional differences, larger where metabolite levels are high	Metabolite stability, processing conditions, separation
Analyte instability in incurred samples ^{3, 10, 11}	Repeat lower, worse with storage time	Incurred sample stability, storage and freeze-thaw history
Sample inhomogeneity and mixing ^{1, 11, 12}	Scatter in both directions, unrelated to concentration	Thawing, mixing and aliquoting steps
Interference from concomitant medications or unique sample components ^{1, 3}	Failures confined to certain subjects, periods or treatments	Chromatograms and IS response
Human error, procedural issues, preanalytic contamination ^{10, 11, 12}	Isolated failures tied to one analyst, batch or step	Sample tracking and aliquot records
Inappropriate assay range ¹¹	Results clustered at one end of the range	Result distribution against the range; M10 Section 3.3.3 ¹
Calibration fit in one run (this paper)	One-directional differences near one end, tied to a run	Both runs' model and weighting, bracketing %RE, slope ratio, intercepts

Causes are from the cited sources except the last row, this paper's argument; other columns are WLTR's guidance.

05 • EVERY ISR PAIR RESTS ON TWO CALIBRATION CURVES

Every ISR pair rests on two calibration curves.

M10 requires a calibration curve “for each analytical run” (Section 3.2.4), and the repeat runs on a different day, so a pair's two results come from two curves¹. Both can pass ($\pm 15\%$ of nominal, $\pm 20\%$ at the LLOQ; Section 3.3.2) and still disagree, because the criteria are checked only at the standards and allow a band of bias at each¹.

Low-end model or weighting bias. When variance grows with concentration, the highest standards dominate an unweighted fit, and weighting improves accuracy at the low end¹⁴. Gu and colleagues found curves with incorrect weighting factors unstable, though concentrations from passing curves were little affected¹⁵. Still, a passing curve can be off by up to 20% at the LLOQ. M10 states that the regression model, weighting scheme and transformation “should be determined during the method validation” (Section 3.2.4), so a run processed otherwise departs from the method¹.

Samples near the LLOQ. M10 accepts precision up to 20% CV at the LLOQ (Section 3.2.5), so random error alone can split a pair¹.

Slope differences between runs. Each run's curve absorbs a proportional change in response. Differences arise when the curves also differ in intercept or shape, or when response drifts within a run; M10 requires QCs interspersed through the run (Section 3.3.1)¹.

Dilutions and carry-over. A diluted sample inherits the curve's bias at its as-read concentration, and different dilution factors put the two readings at different points on different curves. M10 requires validated dilution integrity and dilution QCs (Sections 3.2.7 and 3.3.2) and caps carry-over after the highest standard at 20% of the LLOQ response (Section 3.2.6), which matters most for low samples run after high ones¹.

06 • AN ILLUSTRATIVE ISR FAILURE TRACED TO THE CALIBRATION

An illustrative ISR failure traced to the calibration.

Illustrative data. An LC-MS/MS assay in human plasma covers 1 to 1,000 ng/mL with eight non-zero levels, validated with linear regression and $1/x^2$ weighting. Run A, the original analysis, was processed unweighted, a deviation that went unnoticed because every standard passed. Run B, the repeat on a later day, was processed as validated; its slope was 0.95 times run A's. Five ISR samples come from near Cmax and five from the late elimination phase. Sample random error is held at 2% or less for clarity.

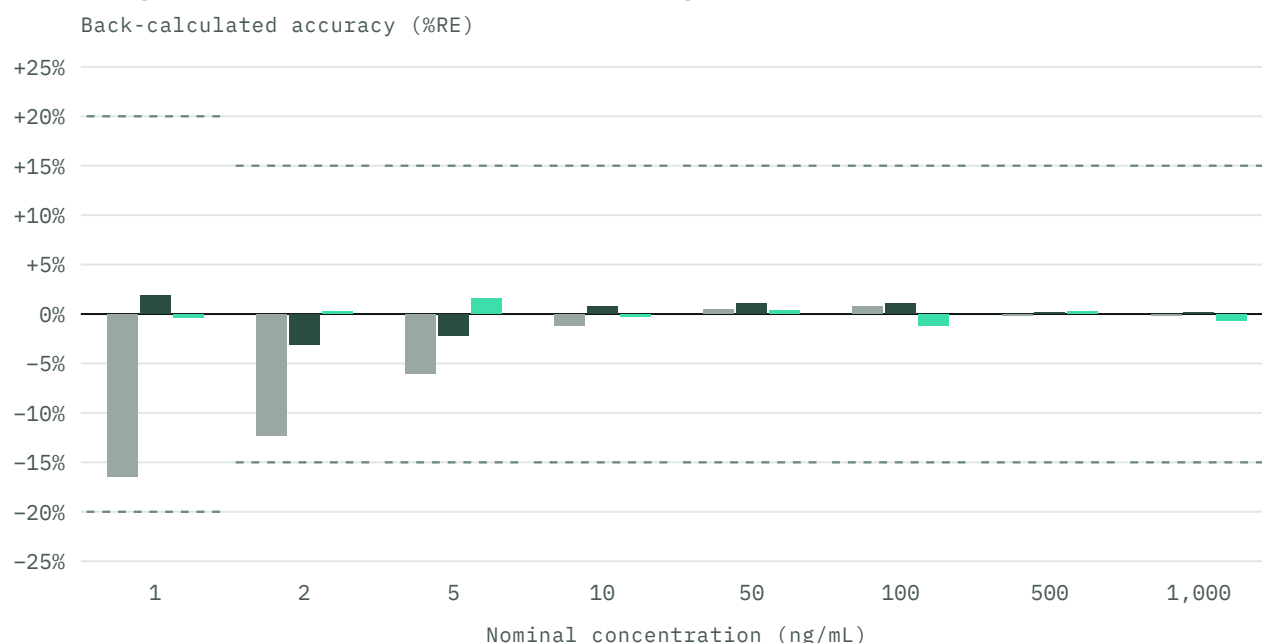
SAMPLE	PHASE	ORIGINAL, RUN A (NG/ML)	REPEAT, RUN B (NG/ML)	% DIFFERENCE	WITHIN ±20%
S01	Cmax	653	632	-3.3	Yes
S02	Cmax	551	567	+2.9	Yes
S03	Cmax	477	462	-3.2	Yes
S04	Cmax	388	397	+2.3	Yes
S05	Cmax	318	308	-3.2	Yes
S06	Elimination	2.39	2.58	+7.6	Yes
S07	Elimination	1.20	1.55	+25.5	No
S08	Elimination	1.14	1.41	+21.2	No
S09	Elimination	1.07	1.43	+28.8	No
S10	Elimination	1.06	1.32	+21.8	No

Illustrative data. Percent difference = (repeat - original) ÷ mean × 100, from reported values (three significant figures).

Six of ten pairs pass, one short of the seven needed. The Cmax pairs agree within $\pm 3.3\%$; four of five elimination-phase pairs differ by +21.2% to +28.8%, all in one direction. Run A's low QC at 3 ng/mL read -7.7%, well inside $\pm 15\%$.

FIGURE 1 Calibration standard accuracy in runs A and B

Run A, unweighted (as reported) Run A, $1/x^2$ refit (diagnostic) Run B, $1/x^2$ (as reported)



Illustrative data. Run A's unweighted fit passes at every level but trends negative toward the LLOQ; the $1/x^2$ refit and run B do not.

Run A's standards read -16.4%, -12.3% and -6.0% at 1, 2 and 5 ng/mL: passing, but trending. Run B's are within $\pm 1.6\%$ at every level. In an unweighted fit the top of the curve sets the intercept; in this design, a 1% error in the 1,000 ng/mL standard moves it by 0.55 ng/mL, more than half the LLOQ, against 0.003 ng/mL under $1/x^2$.

Refit at the validated $1/x^2$ as a diagnostic only, run A brings all ten pairs within $\pm 20\%$ (elimination-phase differences 0.0% to +12.6%). The reported results stand; any reprocessing is for the investigation and the SOPs to decide.

07 • COMPARE THE TWO RUNS BEFORE YOU BLAME THE SAMPLES

Compare the two runs before you blame the samples.

This check is WLTR's recommendation, not a regulatory requirement. The ISR table gives each result's run ID, so both curves can be retrieved¹.

01 Confirm the model. Both runs should use the validated model, weighting and transformation; review any excluded standards.

02 Compare the bracketing levels. For each failing pair, compare the %RE of the standards that bracket the as-read concentration in each run: here, -16.4% and -12.3% in run A against -0.4% and +0.3% in run B.

- 03 **Compute the slope ratio.** A ratio away from 1 with matching %RE patterns is a change the calibration absorbed. Here it was 0.95, and the Cmax pairs agreed within $\pm 3.3\%$.
- 04 **Compare intercepts in concentration units.** Divide each intercept by its slope: 0.358 ng/mL in run A (36% of the LLOQ), 0.044 ng/mL in run B. The 0.314 ng/mL difference alone predicts +22.9% to +26.4% at the four failing pairs; +21.2% to +28.8% was observed.
- 05 **Refit as a diagnostic.** Refit both runs under the validated model and the alternatives, recalculate the failing pairs and document the outcome.
- 06 **Look across runs.** A low-end pattern that follows runs, instruments or sites points to calibration or processing; one that follows subjects points to the samples.

08 • WHERE WLTR FITS

Where WLTR fits.

WLTR does not calculate ISR results. It verifies the calibration curves those results rest on. WLTR evaluates the calibration data your CDS already produces, checks back-calculated accuracy at every level so that bias at the LLOQ and ULOQ surfaces before it reaches a study report, and shows the candidate regression and weighting choices side by side against your lab's own acceptance criteria¹⁶.

Each evaluation is documented and reproducible, including how the reported model compares with alternatives and how accurately each standard back-calculates, so the curves behind both results of an ISR pair can be reviewed on the same terms. WLTR also trends calibration performance across batches, instruments and sites, so a recurring low-end pattern can be seen early, before it becomes an ISR investigation. WLTR verifies. Your lab validates.

SOURCES

1. ICH, M10 Bioanalytical Method Validation and Study Sample Analysis, Step 4, May 24, 2022, Secs. 3.2.4, 3.2.5, 3.2.6, 3.2.7, 3.3.1, 3.3.2, 3.3.3, 5 and 8 (Table 1). database.ich.org/sites/default/files/M10_Guideline_Step4_2022_0524.pdf
2. U.S. FDA, M10 Bioanalytical Method Validation and Study Sample Analysis, Guidance for Industry, November 2022 (ISR in Sec. V). www.fda.gov/media/162903/download
3. U.S. FDA, M10 Bioanalytical Method Validation and Study Sample Analysis: Questions and Answers, Guidance for Industry, June 2024, question on trends of concern or ISR failure (M10 Section 5). www.fda.gov/media/179296/download
4. C. T. Viswanathan, S. Bansal, B. Booth, A. J. DeStefano, M. J. Rose, J. Sailstad, V. P. Shah, J. P. Skelly, P. G. Swann, R. Weiner, "Workshop/conference report: Quantitative bioanalytical methods validation and implementation: Best practices for chromatographic and ligand binding assays," The AAPS Journal 9(1), 2007, E30-E42, Conclusion, item 3. DOI 10.1208/aapsj0901004. doi.org/10.1208/aapsj0901004
5. C. T. Viswanathan, S. Bansal, B. Booth, A. J. DeStefano, M. J. Rose, J. Sailstad, V. P. Shah, J. P. Skelly, P. G. Swann, R. Weiner, "Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays," Pharmaceutical Research 24(10), 2007, 1962-1973, Abstract. DOI 10.1007/s11095-007-9291-7. doi.org/10.1007/s11095-007-9291-7

6. M. L. Rocci, V. Devanarayan, D. B. Haughey, P. Jardieu, "Confirmatory reanalysis of incurred bioanalytical samples," The AAPS Journal 9(3), 2007, E336-E343, Abstract. DOI 10.1208/aapsj0903040. doi.org/10.1208/aapsj0903040
7. D. M. Fast, M. Kelley, C. T. Viswanathan, J. O'Shaughnessy, S. P. King, A. Chaudhary, R. Weiner, A. J. DeStefano, D. Tang, "Workshop Report and Follow-Up: AAPS Workshop on Current Topics in GLP Bioanalysis: Assay Reproducibility for Incurred Samples, Implications of Crystal City Recommendations," The AAPS Journal 11(2), 2009, 238-241, Abstract and Introduction. DOI 10.1208/s12248-009-9100-9. doi.org/10.1208/s12248-009-9100-9
8. P. Timmerman, S. Luedtke, P. van Amsterdam, M. Brudny-Kloeppel, B. Lausecker, et al., "Incurred Sample Reproducibility: Views and Recommendations by the European Bioanalysis Forum," Bioanalysis 1(6), 2009, 1049-1056, Abstract. DOI 10.4155/bio.09.108. doi.org/10.4155/bio.09.108
9. E. Fluhler, F. Vazvaei, P. Singhal, P. Vinck, W. Li, J. Bhatt, T. de Boer, A. Chaudhary, M. Tangiuchi, V. Rezende, D. Zhong, "Repeat Analysis and Incurred Sample Reanalysis: Recommendation for Best Practices and Harmonization from the Global Bioanalysis Consortium Harmonization Team," The AAPS Journal 16(6), 2014, 1167-1174, Abstract. DOI 10.1208/s12248-014-9644-1. doi.org/10.1208/s12248-014-9644-1
10. L. B. Lee, E. R. Wickremesinhe, "A Decade of Incurred Sample Reanalysis: Failures, Investigations and Impact," Bioanalysis 10(21), 2018, 1767-1772, Abstract. DOI 10.4155/bio-2018-0214. doi.org/10.4155/bio-2018-0214
11. S. G. Summerfield, M. Barfield, S. A. White, "Incurred Sample Reanalysis at GSK: What have we Learned?" Bioanalysis 10(21), 2018, 1755-1766, Abstract. DOI 10.4155/bio-2018-0204. doi.org/10.4155/bio-2018-0204
12. J. Welink, Y. Xu, E. Yang, A. Wilson, N. Henderson, et al., "2018 White Paper on Recent Issues in Bioanalysis: 'A Global Bioanalytical Community Perspective on Last Decade of Incurred Samples Reanalysis (ISR)' (Part 1: Small Molecule Regulated Bioanalysis, Small Molecule Biomarkers, Peptides & Oligonucleotide Bioanalysis)," Bioanalysis 10(22), 2018, 1781-1801, section "What ISR/ISS is telling the global bioanalytical community about its methods/procedures." DOI 10.4155/bio-2018-0268. doi.org/10.4155/bio-2018-0268
13. S. Subramaniam, D. Patel, B. M. Davit, D. P. Conner, "Analysis of Imprecision in Incurred Sample Reanalysis for Small Molecules," The AAPS Journal 17(1), 2015, 206-215, Abstract. DOI 10.1208/s12248-014-9689-1. doi.org/10.1208/s12248-014-9689-1
14. A. M. Almeida, M. M. Castel-Branco, A. C. Falcão, "Linear regression for calibration lines revisited: weighting schemes for bioanalytical methods," Journal of Chromatography B 774(2), 2002, 215-222, Abstract. DOI 10.1016/S1570-0232(02)00244-1. doi.org/10.1016/S1570-0232(02)00244-1
15. H. Gu, G. Liu, J. Wang, A.-F. Aubry, M. E. Arnold, "Selecting the Correct Weighting Factors for Linear and Quadratic Calibration Curves with Least-Squares Regression Algorithm in Bioanalytical LC-MS/MS Assays and Impacts of Using Incorrect Weighting Factors on Curve Stability, Data Quality, and Assay Performance," Analytical Chemistry 86(18), 2014, 8959-8966, Abstract. DOI 10.1021/ac5018265. doi.org/10.1021/ac5018265
16. U.S. Patent 12,038,448 B2, inventor Walter "Chip" Hinchman Sr. patents.google.com/patent/US12038448B2



Walter Hinchman, MBA

CO-FOUNDER & CEO

CALIBRATION INTELLIGENCE INFRASTRUCTURE. · INSIGHTS 16 · OCTOBER 2026

WLTR, INC. · NEVADA C-CORP

WLTR, INC. · NEVADA C-CORP · 2026

WLTR INSIGHTS 16