HPLC Method Validations: Navigating the Pitfalls

Joe Page, Ph. D., President
Eurofins Advantar, San Diego

Method Validation and Pitfalls Outline

- Documents and Validation Parameters
- System Suitability
- Accuracy
- Sample and Standard Stability
- Precision
- Intermediate Precision
- Linearity
- Specificity
- LOQ/LOD
- Robustness
- Range

Method Validation Guidance Documents

ICH Harmonized Guideline:

- **Validation of Analytical Procedures: Text and Methodology Q2(R1)**

U.S. Department of Health and Human Services FDA, CDER, CBER:

- **Analytical Procedures and Methods Validation for Drugs and Biologics**

Method Validation Parameters



Type of analytical procedure characteristics	IDENTIFICATION	TESTING FOR IMPURITIES quantitat.	limit	ASSAY - dissolution (measurement only) - content/potency
Accuracy	-	+	-	+
Precision				
Repeatability	-	+	-	+
Interm.Precision	-	+ (1)	-	+ (1)
Specificity (2)	+	+	+	+
Detection Limit	-	- (3)	+	-
Quantitation Limit	-	+	-	-
Linearity	-	+	-	+
Range	-	+	-	+

- \- signifies that this characteristic is not normally evaluated
- \+ signifies that this characteristic is normally evaluated
- (1) in cases where reproducibility (see glossary) has been performed, intermediate precision is not needed
- (2) lack of specificity of one analytical procedure could be compensated by other supporting analytical procedure(s)
- (3) may be needed in some cases

Internal Method Validation Documents



Prior to starting a method validation, you need two documents:

- **QA approved validation protocol with acceptance criteria**

- **A detailed written method with version control**

Lack of version control can create confusion and potentially cause the validation to fail

System Suitability

System Suitability testing is an integral part of a GMP HPLC Method

Typical Data:

Standard injections (n=6), NMT 2% RSD.

%Recovery of Check Standard 98.0 to 102.0% (assay)

Resolution between two key peaks r ≥ 2.0

Tailing of main peak NMT 2.0

System suitability should be run at the start of every validation sample set. It's the only way to know that the system is suitable for testing. Its a validation parameter that is sometimes overlooked, but it can call your results into question (or failure) if it's not performed.

Accuracy

Also known as "Spike and Recovery". Your method should be able to quantitatively recover a known amount of standard or API spiked into your placebo

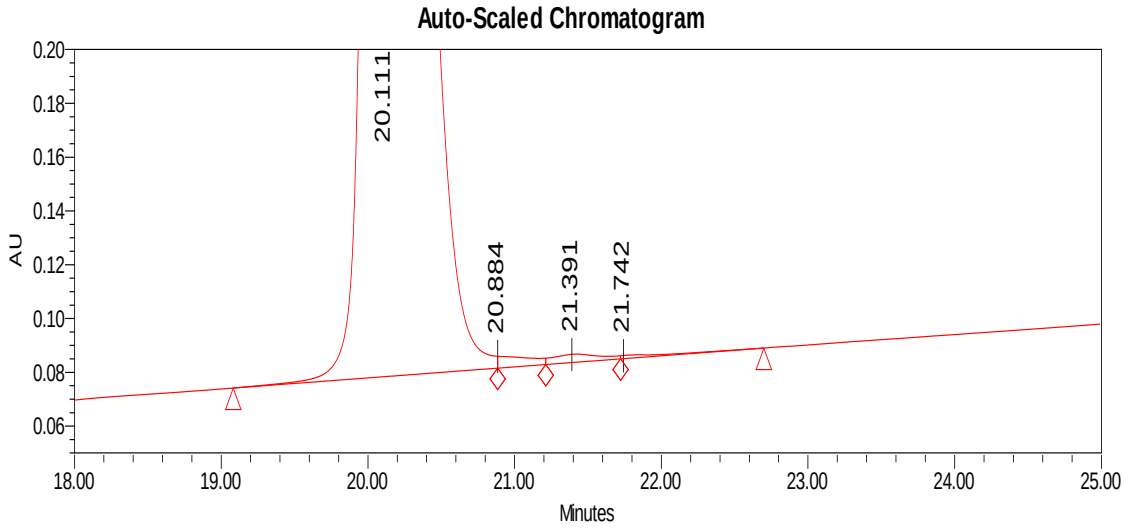
- Typical Assay Data: Spiking is typically performed at 80% (n=3), 100% (n=3) and 120% (n=3) of your drug's label claim
- Typical Assay Acceptance Criteria: % Recovery is within 98.0% to 102.0% of the amount spiked into the placebo.

- Typical Impurity Data: Spiking is typically performed near the LOQ, Specification, and 120% of the Specification
- Typical Impurity Acceptance Criteria: % Recovery is within 95.0% to 105.0% (or 90.0 to 110.0%) of the amount spiked into the placebo. Ranges vary depending on the capability of the method and toxicology results

Accuracy Pitfalls

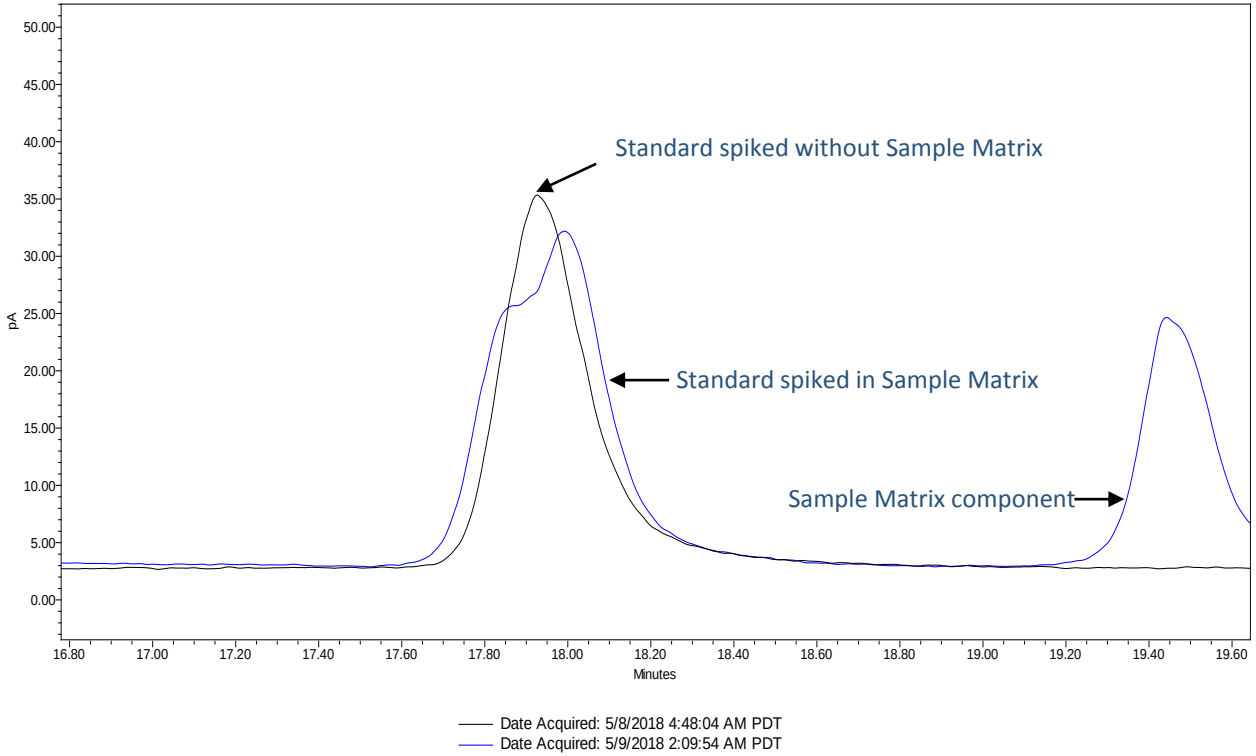
Reasons for not Achieving 100% Recovery

Method lacks specificity and peaks are not fully resolved



Accuracy Pitfalls

Method matrix interferes with recovery



Accuracy Pitfalls

Reasons for not Achieving 100% Recovery (continued)

More on method matrix interference

- Method diluent is slowly degrading the analyte (TFA, formic acid)
- Method diluent does not fully solubilize the analyte
- Sample has limited stability in the method diluent

Diluent evaporation results in over recovery (methanol).

Not accounting for differences in response factors (impurities)

Accuracy Pitfalls

Reasons for not Achieving 100% Recovery (continued)

- Method lacks solubility for the drug at the 120% spike.
- Low level impurity spikes are adhering to the glassware.
- Filtering of the sample results in a loss of analyte.
- Spiking the standard into placebo is not equivalent to drug product which may affect its solubility. This is especially true for:
  - Homogenized suspension formulations
  - Nanoparticle formulations

Sample and Standard Stability

The test samples and working standard solutions must be demonstrated to be stable over a defined period of time that the assay is run.

- Typical Data: Measure sample and standard at T=0, 1, 3, 5 days
- Typical Acceptance Criteria:
  - Longest period in which the Assay sample and standard is within 98.0% to 102.0% of the T=0 result.
  - Longest period in which the Impurity sample and standard is within 95.0% to 105.0% (or 90.0 to 110.0%) of the T=0 result.

Sample and Standard Stability Pitfalls

Run this test early in your validation!

I’ve seen sections of validations that had to be repeated because a parameter was determined with a 3-day old sample, and later it was determined that the sample was only stable for 1 day.

Precision (repeatability)

Precision of a method is the closeness of agreement between a series of measurements obtained from multiple samplings of a homogeneous sample

- Typical Assay Data: Precision is from the n=3 spiking replicates at 80%, 100% and 120% or n=6 at the 100% level
- Typical Assay Acceptance Criteria: RSD is NMT 2.0% at each level

- Typical Impurity Data: Precision is from the n=3 spiking replicates at LOQ, specification and 120% of the specification
- Typical Impurity Acceptance Criteria: RSD is NMT 10% at LOQ and NMT 5% to 10% at specification and 120% of the specification levels

Precision Pitfalls

Reasons for Failing Precision Criteria

Precision repeatability results are usually tied to the accuracy of the method. Problems with accuracy are also manifested in the precision results.

- For example, if your method poorly resolves the main analyte this is likely to show up during precision too as it may be difficult to reproducibly integrate the main analyte peak.

Intermediate Precision

Within the same lab, perform the accuracy experiment (assay at 80, 100, and 120% or impurities testing near LOQ, Spec, 120% of Spec) with a different analyst, with different preparations, different instruments, on different days.

- Typical Acceptance Criteria:
  - Must meet accuracy criteria
  - Each analyst must achieve precision criteria
  - Determine the precision of all results according to analysts, instruments, and days.
  - Compare the precision of each analyst, each instrument, and days.

Intermediate Precision Pitfalls

Intermediate precision failures are often the hardest to resolve due to the human component. Method directions may be vague or interpreted differently between Analyst 1 and Analyst 2

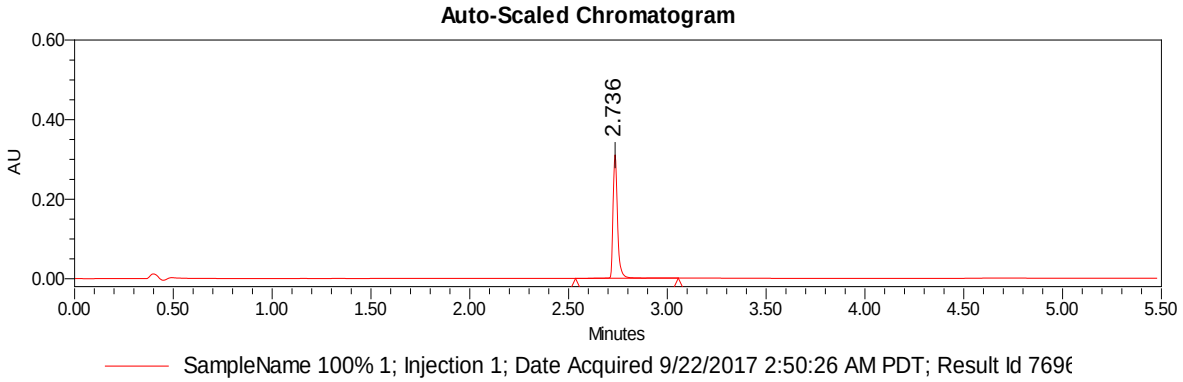
- Example 1:

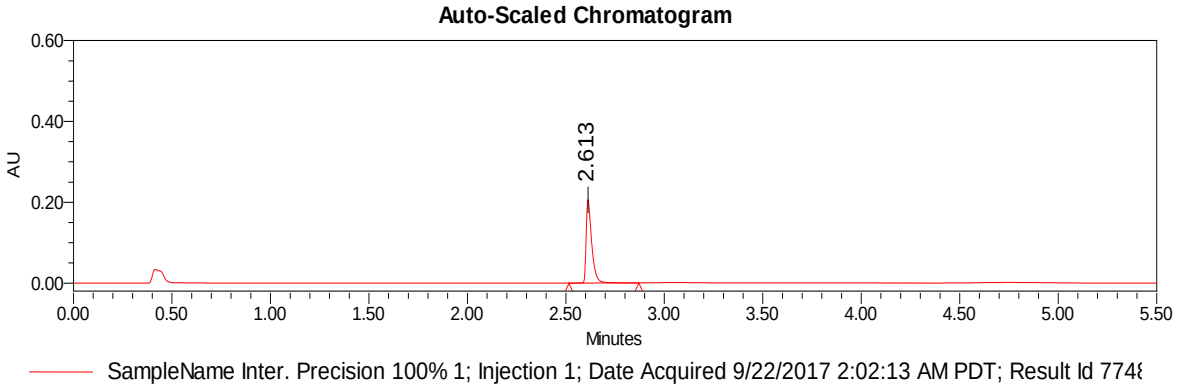
Analyst 1 aliquots the sample preparation into a glass HPLC vial whereas Analyst 2 aliquots into a plastic HPLC vial. The analyte in question has different compatibility with glass compared to plastic which skews the results.

Often times the only way to resolve intermediate precision issues is to closely observe both analysts in the lab.

Intermediate Precision Pitfalls

- Example 2:
  - Intermediate Precision Testing for Assay using the same DP on a column lots X and Y. System suitability passed on both systems, but the assay failed every time column lot Y was used.





Intermediate Precision Pitfalls

- Example 2 (continued)

Investigations showed that all the solutions preparations would pass on column X, but never with column lot Y. It was concluded that the method was sensitive to changes in column lots and that a control drug product standard would be needed to screen new column lots.

Specificity

The ability of the method to resolve the analytes of interest from impurities, degradation products, excipients, and matrix components.

- Typical Assay Data: Chromatographic resolution and peak purity of the peaks of interest in aged and force degraded (heat, acid, base, oxidation, light) samples.

- Typical Assay Acceptance Criteria:
  - All peaks of interest are resolved with a resolution ≥ 1.0
  - All peaks of interest are homogeneous as demonstrated by PDA or LCMS peak purity analysis

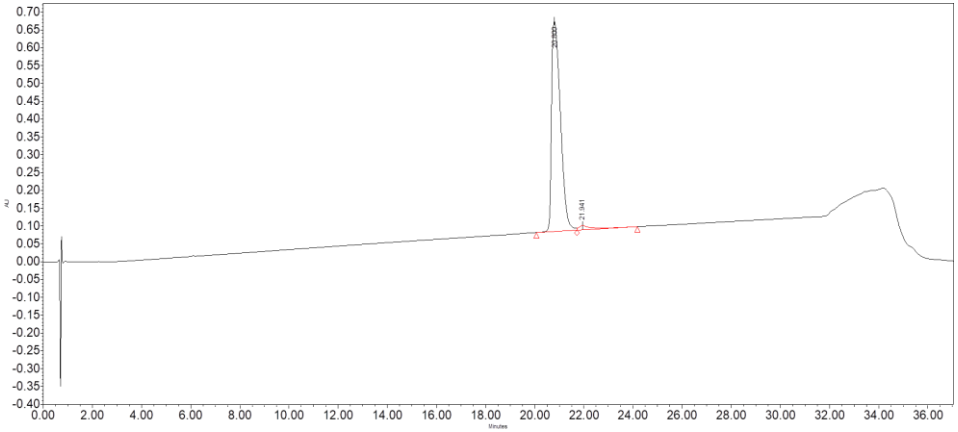
Specificity Pitfalls

Some separations are difficult and you may not be able to resolve all the impurities in a single method.

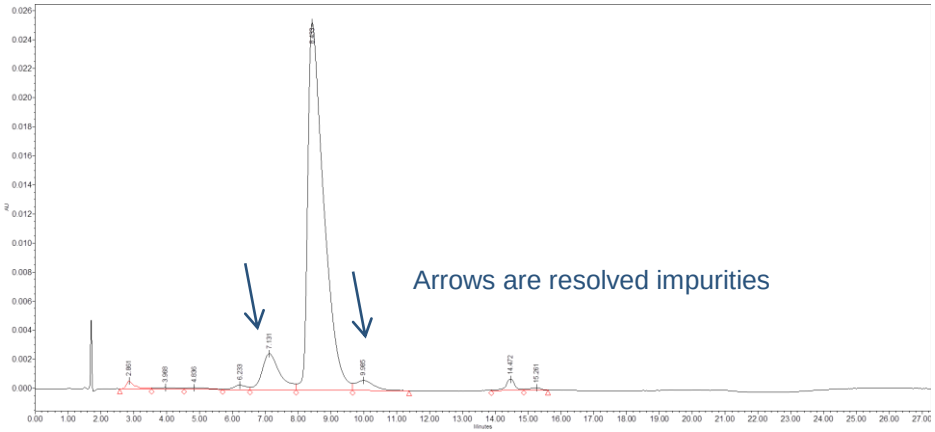
Example: A client in early phase developed an RP-HPLC assay method with seemingly good resolution and a homogeneous peak by PDA peak purity analysis. An orthogonal strong cation exchange (SCX) method was developed that showed the co-elution of related substances that was missed by the PDA analysis (due to similar UV spectra)

Specificity Pitfalls

Single RP-HPLC Peak, seemingly homogeneous by PDA Analysis



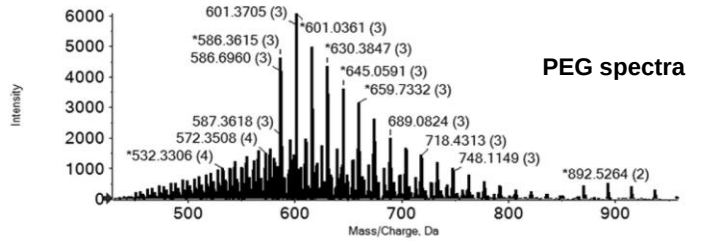
Multiple peaks under the main peak as resolved Ion Exchange-HPLC

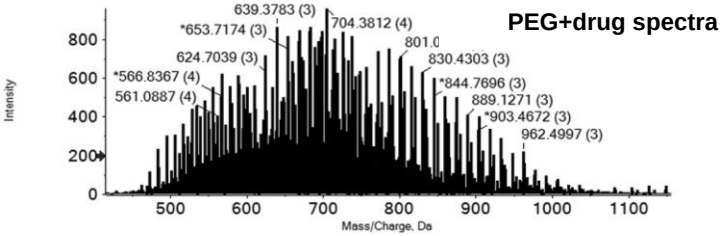


Specificity Pitfalls

For compounds that have impurities with closely related structures, the UV spectra are often the same which renders peak purity analysis by PDA useless. In these cases, you must go to LCMS for peak purity.

PEGylated and polymeric compounds have heterogeneous structures so even peak purity by LCMS may be impossible.





Linearity

The ability of the method to obtain results that are directly proportional to the concentration of the sample

- Typical Data:
  - Assay: 5 concentrations from 80% to 120% of label claim
  - Impurities: 5 concentrations from LOQ to 150% of specification.

- Typical Acceptance Criteria:

  Assay and Impurities: $r^2 > 0.98$

  Report the slope and intercept

  Show plot of linear regression line

Linearity Pitfalls

- An intercept value that is not close to zero is an indicator of bias in the method. Some validations require the intercept to be NMT 2% of the peak area of the 100% or assay sample.
- Slope values in the impurity range may differ from slope values in the assay range.
- Non-linear detectors should not use linear regression, but should use a polynomial model.

Limit of Quantitation

The lowest amount of a substance that can be measured with accuracy and precision.

- Typical Data: Spiking into placebo is typically performed at 2 or 3 levels that approach a signal to noise ratio of 10:1.

- Typical Acceptance Criteria:
  - Average S/N ≥ 10
  - % Recovery is within 90.0% to 110.0% of the amount spiked into the placebo. Ranges vary depending on the capability of the method and toxicology results.
  - The lowest concentration with a S/N ≥ 10 defines the LOQ. This can be reported as a percentage of the label claim.

Limit of Quantitation Pitfalls

Reasons for not Achieving Targeted LOQ Level

Method noise is decreasing the sensitivity of the method:

- Different system/detector being used than in development.
- Sample preparation adding noise.

An aged column is causing band-broadening.

At low levels the analyte is sticking to glass or plastic.

Limit of Quantitation Pitfalls

Reasons for not Achieving Targeted LOQ Level (continued)

Method lacks specificity and peaks are not fully resolved.

Not accounting for differences in response factors (if standard is different than analyte)

Limit of Detection

The lowest amount of a substance that can be detected but not necessarily quantitated

- Typical Data: Spiking into placebo is typically performed at 2 or 3 levels that approach a signal to noise of 2:1 or 3:1.

- Typical Acceptance Criteria:
  - Average S/N ≥ 2
  - % Recovery is within 70.0% to 130.0% of the amount spiked into the placebo.
  - The lowest concentration with a S/N ≥ 2 defines the LOD. This can be reported as a percentage of the label claim.

The same pitfalls as LOQ apply to LOD

Robustness

Robustness demonstrates that the method is reliable with respect to deliberate variations in the method parameters

- Typical Data: Monitor retention time, peak area, resolution with respect to changes in:
  - Mobile phase pH (higher and lower pH)
  - Mobile phase composition (+/- % organic phase)
  - Different column lots
  - Column temperature (+/-)
  - Flow rate (+/-)
- Typical Acceptance Criteria:
  - Retention time remains within X% or X min of the prescribed conditions.
  - Peak area of select analytes remains within X% of the peak area under the prescribed conditions.
  - Resolution of key peak separations is unaffected

Robustness Pitfalls

When a Robustness variation fails, one can narrow the range of the variation, or lock the parameter at a set value *(e.g.* a set pH, a set % organic, a set flow rate, etc.)

Setting ranges too wide can “over-challenge” the method forcing extra robustness testing, or forcing the method to be locked into set value that may not be necessary.

Range

The Range of a method covers sample concentrations where the method has been demonstrated to be linear, accurate and precise.

Typical Data: The range of an HPLC method is typically taken as a composite of the linearity, accuracy and precision results.



Thank You!