

Establishment and Optimization of Detection Methods for Related Substances in Pharmaceutical Raw Materials

Wenqian Gu, Haili Zhao

Shanghai Wei Er Biopharmaceutical Technology Co., Ltd, Shanghai 201707, China

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This study is based on the idea that “quality comes from design,” that is, to systematically analyze the establishment of detection methods for related substances in pharmaceutical raw materials. At the beginning, hetero-mass spectrometry and the purpose of method development are summarized and explained, and suggestions on how to select key chromatographic parameters and how to do risk assessment are put forward. Then, the method design space is optimized by means of experimental design, and the intelligent development technology is used to make the method design more complete and efficient. The control of the chromatographic system, the influence of solvents, the integration conditions, and the detection wavelength are discussed in detail, and the corresponding control countermeasures are given. Finally, the whole life cycle of the detection of related substances is verified, and the forced degradation, sensitivity, and system applicability are taken into account. The results show that the standardized and complete development and verification ideas are conducive to ensuring the reliability of the detection method.

Keywords: Pharmaceutical raw materials; Related substances; Method establishment; Method optimization; Analysis quality comes from design; Method validation

Online publication: October 23, 2026

1. Introduction

Impurities, degradation products, and other related substances contained in pharmaceutical raw materials can have an impact on the safety, effectiveness, and stability of drugs. Therefore, establishing reliable, sensitive, and stable methods for detecting relevant substances is an important task in drug quality control. Traditional detection methods mostly rely on experience and repeated experimentation, which not only takes a long time but also lacks careful consideration of various complex influencing factors. With the continuous advancement of analytical technology, from the source design of method development to the final verification and confirmation, we pay full attention to the key technical parameters and control measures that affect the performance of detection methods, in order to improve the scientificity and application effectiveness of

detection methods.

2. Strategy for establishing substance detection methods based on AQbD concept

2.1. Determination of impurity mass spectrometry analysis and method development objectives

The first step of method development is to conduct a comprehensive impurity spectrum analysis of drug raw materials, and sort out the process impurities and potential degradation products by combining the synthesis process, chemical structure characteristics, and stability research data. On this basis, the general situation of the analysis target is determined, and the number of impurities to be separated, resolution, detection sensitivity, analysis time, and other key performance indicators are defined. The analysis of the target profile can provide direction for the development of the method and serve as an important basis for the subsequent method evaluation.

2.2. Screening and risk assessment of key method variables

There are many factors affecting chromatographic separation, mainly including mobile phase, chromatographic column, column temperature, gradient condition, and flow rate. The fishbone diagram, FMEA, and other tools can be used to screen the parameters, assess the risk, and identify the high-risk parameters that have a significant impact on the resolution, the number of theoretical plates, and the tailing factor. For impurities with large hydrophobicity differences, gradient time and organic phase ratio are more critical. Through single-factor experiments or partial factorial designs, the key variables can be quickly screened, and the optimization scope can be reduced.

2.3. Experimental design optimization and design space construction

After the key variables are selected, the experimental design (DOE) is used for system optimization, which can simultaneously investigate multiple factors and their interactions, and avoid the limitations of single-factor experiments. Taking gradient time and column temperature as examples, the central composite design or Box-Behnken design can be selected. **Table 1** shows the design matrix and partial results of the hypothetical experiment, with the separation of two key impurities from the main peak as the evaluation index.

Table 1. Two-factor, three-level center composite design example and resolution results

Experiment No.	GradientTime (min)	ColumnTemperature (°C)	Resolution (ImpurityAMainPeak)	Resolution (ImpurityBMainPeak)
1	20	25	1.2	1.8
2	40	25	1.9	2.3
3	20	35	1.1	1.6
4	40	35	1.8	2.5
5	15	30	1.0	1.5
6	45	30	2.0	2.4
7	30	22	1.6	2.0
8	30	38	1.5	2.1
9(Central Point)	30	30	1.7	2.2

The calculation formula for separation is:

$$R_s = \frac{2(t_{R2} - t_{R1})}{W_1 + W_2}$$

Where t_{R1} and t_{R2} respectively represent the retention times of two adjacent chromatographic peaks, W_1 and W_2 . This calculation formula is the core quantitative index for evaluating the chromatographic separation effect. Researchers can carry out mathematical modeling on the experimental results, draw contour maps or three-dimensional response surface maps reflecting the relationship between influencing factors and detection results, and then predict the parameter range that can stably meet all separation requirements, that is, the design space, based on these models. Carrying out experiments and method operations in the predetermined design space can ensure the performance of the detection method meets the expected requirements and improve the durability of the method.

2.4. Application of intelligent method development tools

In recent years, intelligent method development software has been widely used, integrating compound structure, chromatographic column database, and retention model, and inputting structure can predict retention time and generate virtual chromatogram^[1]. Through computer simulation, the optimal conditions can be quickly screened, and the number of experiments can be reduced; for example, the optimal pH range of the mobile phase can be quickly determined. This model of using calculation to guide experiments has significantly improved the efficiency and scientificity of method development.

3. Key factors affecting the accuracy and reproducibility of determination of related substances and control strategies

3.1. Chromatographic system factors and their control

The operating state of the chromatographic system will affect the analysis results, and the gradient delay volume is one of the factors that is easy to ignore. Different types of liquid chromatographs have different pipeline volumes from the mixer to the column inlet, which will make the actual start-up time of the gradient program deviate, thus changing the retention time and resolution of the early eluting impurities. In this regard, the model or delay volume range of the liquid chromatograph can be defined in the method; the delay volume difference can be compensated by using the initial isocratic stage, and equipment with column temperature control accuracy within $\pm 0.1^\circ\text{C}$ can be selected to ensure the stability of retention time. In addition, the preparation method of the mobile phase, the batch difference of the chromatographic column, and the decline of column efficiency in the process of use need to be monitored through strict system applicability testing.

The theoretical plate number n of a chromatographic column is an important parameter to measure column efficiency, and its calculation formula is:

$$N = 16 \left(\frac{t_R}{W} \right)^2$$

Where t_R is the retention time, and W is the peak width. N : The greater the value, the higher the separation efficiency of the chromatographic column.

3.2. Sample pretreatment and solvent effect

The preparation process of the sample solution is prone to analysis error. When selecting the preparation solvent, it is necessary to ensure that all substances to be tested can be fully dissolved, and at the same time, no additional degradation of the sample can occur. If the elution strength of the sample solvent is higher than the initial proportion of the mobile phase, the solvent effect will occur, resulting in the deformation, bifurcation, or retention time shift of the chromatographic peak, which is more obvious for impurities with strong polarity and early peaks. The solvent effect can be controlled by using a solvent with the same proportion or lower strength as the initial composition of the mobile phase, reducing the injection volume, and using the focusing effect of column enrichment. In the gradient method, the basic principle to avoid solvent effect is to ensure that the solvent strength of the sample is not higher than the initial composition of the mobile phase.

3.3. Integral parameter setting and data processing

The setting of integration parameters directly affects the accuracy and reproducibility of quantitative results, which is also the main reason for differences in laboratory data. The slope sensitivity, peak width, minimum peak area, baseline processing mode, and other parameters will change the peak area calculation results, especially for the shoulder peak with low content of impurities and incomplete separation. By establishing a standardized integral event table, the unified parameters are determined in the method validation phase and fixed in the method, which can effectively control the error. When dealing with baseline noise, it is necessary to reasonably set integral parameters, accurately identify impurity peaks, and eliminate baseline interference.

Another important aspect of data processing is the correction factor. When the principal component self-control method is used to calculate the impurity content, if the response factor of the impurity and the principal component is quite different, it is necessary to use the impurity reference substance, or complete the accurate quantification by measuring and introducing the correction factor, such as:

$$RRF = \frac{A_{imp} / C_{imp}}{A_{API} / C_{API}}$$

Wherein, A_{imp} , A_{API} , are the peak areas of impurities and main components, respectively, C_{imp} and C_{API} are their concentrations, respectively. When the relative response factor is beyond the range of 0.8~1.2, the correction factor method should be used for accurate quantification.

3.4. Wavelength selection and response consistency of detector

The detection wavelength of the UV detector will directly affect the detection sensitivity and impurity response factor. Generally, the wavelength with strong absorption of impurities and main components will be selected. In the determination of related substances, the general wavelength with sufficient sensitivity to key impurities is more commonly used. When the UV absorption of multiple impurities is quite different, a single wavelength is difficult to meet the sensitivity requirements of all impurities. A diode array detector can be used for multi-wavelength detection or wavelength switching, such as quantifying most impurities with the main wavelength, and then detecting impurities with special absorption with the exclusive wavelength. During verification, it is necessary to investigate the response consistency between the principal component and each impurity at this wavelength, or determine the response factor, so as to provide the basis for selecting a self-control method or a correction factor method ^[2].

4. Full life cycle verification of detection methods for related substances

4.1. Reasonable design of specificity and forced degradation research

Specificity is the primary content of method validation. It needs to be proved that the method can accurately distinguish the detected substance from impurities, excipients, degradation products, and other substances. Forced degradation testing is the core means to evaluate specificity. It needs to carry out systematic degradation treatment of the API under acid, alkali, heat, oxidation, light, and other conditions. The key is to control the degradation of principal components at about 10% to 20%, so as to avoid excessive degradation and the production of secondary products, so as to be close to the real stability^[3]. Through the analysis of degradation samples, it is necessary to ensure the effective separation between the main peak and each degradation product peak, as well as between degradation products. Generally, the purity of the main peak is required to be qualified, and the resolution RS of key impurities is greater than 1.5. As shown in **Figure 1**.

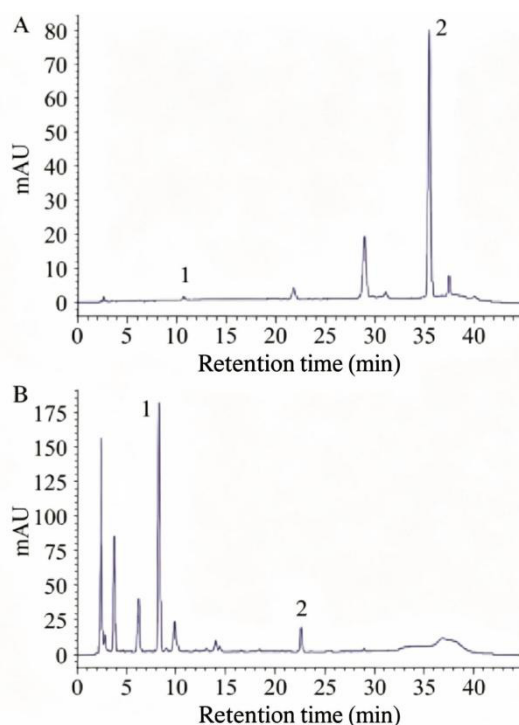


Figure 1. Chromatogram of a drug raw material oxidation degradation sample.

The peak purity factor can be calculated from the spectral data collected by the diode array detector, and its calculation formula is based on spectral similarity:

$$PurityFactor = \frac{\sum (\lambda_i - \lambda)(A_i - A)}{\sqrt{\sum (\lambda_i - \lambda)^2 \sum (A_i - A)^2}}$$

The closer the factor is to 1, the higher the spectral homogeneity of the peak and the higher the reliability of the peak purity.

4.2. Sensitivity and quantitative capability verification

The sensitivity is expressed by the limit of detection (LOD) and the limit of quantification (LOQ), generally

using the signal-to-noise ratio method, and the calculation formula is:

$$S / N = \frac{2H}{h}$$

Where h is the target peak height, and H is the maximum peak-to-valley height of baseline noise. The concentration is LOD when the signal-to-noise ratio is about 3:1, and LOQ when the signal-to-noise ratio is about 10:1. In the inspection of related substances, the LOQ concentration shall not be higher than the threshold of impurity reporting, and the quantitative ability of LOQ level shall be verified; that is, the LOQ concentration standard solution shall be continuously injected 6 times according to the formula:

$$RSD = \frac{\text{standard deviation (SD)}}{\text{Average}(x)} \times 100\%$$

To calculate the relative standard deviation of peak area, RSD is usually required to be less than 10% or 20% to ensure accurate quantification at low concentrations.

4.3. Linear range, accuracy and precision

For the linearity test, prepare standard solutions with at least 5 concentration levels within the range from LOQ to corrected concentration, and perform linear regression on the concentration based on peak area. The correlation coefficient r should generally not be less than 0.999, and the calculation formula of the correlation coefficient is:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

The accuracy is evaluated by the recovery rate test. Add a known amount of impurity reference substance into the blank solvent or test solution, prepare low, medium, and high concentration solutions for determination, and the average recovery rate should be between 90%-108%, and RSD should meet the requirements. The precision is divided into repeatability and intermediate precision, which is verified by measuring the impurity content and calculating RSD.

4.4. Solution stability and durability

The stability of the test sample and the reference sample at room temperature or in cold storage for different times shall be studied respectively to obtain the maximum allowable storage time (taking the impurity content not exceeding $\pm 10\%$ as the judgment standard). Various conditions for durability testing shall be designed according to minor adjustments: pH deviation of the mobile phase ± 0.2 . Column temperature ± 2 °C, flow rate ± 0.1 mL/min. The chromatographic columns of the same type but different brands were used to analyze the separation and numerical results. Finally, it was confirmed that the resolution, number of theoretical plates, tailing factor, and content did not lose acceptability when the parameters changed.

4.5. Reasonable setting of system applicability test

The system suitability test should be used as the necessary inspection for analyzing the system condition. The relevant items should be determined according to the key separation needs in the method and included in the routine test. The efficiency of the chromatographic column is represented by the number of theoretical plates,

n, and the tailing factor, t , is used to estimate whether the peak shape is symmetrical (i.e., whether there is tailing). The calculation formula is as follows:

$$T = \frac{W_{0.05}}{2f}$$

Where $W_{0.05}$ is the peak width at 5% peak height, and f is the length from the peak apex to the peak front. The tailing factor T should be within the range of 0.8–1.5.

The critical resolution R_s serves as a fundamental parameter. It shall be evaluated for the most difficult to separate pair of components, and a minimum acceptable value is specified, generally, $R_s \geq 1.5$ or 2.0.

Sensitivity determination shall be performed under specified signal-to-noise ratio conditions; the signal-to-noise ratio of the principal peak should preferably be greater than 10 times that of the reference.

All the abovementioned system suitability parameters shall be obtained through validation and documented in the standard operating procedure.

5. Conclusion

The establishment and optimization of material detection methods for pharmaceutical raw materials is a systematic task. The method is designed from the starting point of aqbd guiding ideology, controlling the elements of chromatography and sample processing, and taking the whole life cycle into consideration for validation. All points should be coordinated. The use of risk management and experimental arrangement together with control ideas is conducive to obtaining highly deterministic test results, which is of great significance to ensure the correctness of the quality analysis (i.e., actual determination) of pharmaceutical raw materials.

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Shi Y, Xiong J, Wei F, et al., 2019, Research on HPLC Retention Time Prediction Based on Machine Learning and External “Probe” Strategy. *Journal of Pharmaceutical Analysis*, 39(4): 716–721.
- [2] Zhang Y, Tian X, Wang L, et al., 2022, Determination of 7 Known Impurities in Omeprazole Sodium for Injection by Principal Component Self-Control Method. *Journal of Pharmaceutical Analysis*, 42(7): 1201–1208.
- [3] Chen Y, Zhang X, Wu J, et al., 2025, Establishment and Validation of Detection Methods for Related Substances in Carbamazepine Injection. *Journal of China Pharmaceutical University*, 56(1): 65–72.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.