Journal of Chromatography A, 1218 (2011) 2712–2721

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

# Validation of thin layer and high performance thin layer chromatographic methods

B. Renger[a,*], Z. Végh[b], K. Ferenczi-Fodor[b]

[a] Bernd Renger Consulting, Fritz-Reichle-Ring 2, 78315 Radolfzell, Germany
[b] Gedeon Richter Plc., P.O.B. 27, H-1475 Budapest, Hungary

ARTICLE INFO

*Article history:*
Available online 27 January 2011

*Keywords:*
TLC
HPTLC
Validation
Uncertainty

ABSTRACT

Analytical validation is a key requirement to asses and to prove a method's reliability and suitability for an intended use. Planar chromatographic procedures are used in different applications ranging from simple screening tests to sophisticated instrumental quantitative assays of analytes in complex matrices. This paper intends to give guidance on how to adopt international accepted formal requirements and guidelines for validation of these different TLC/HPTLC procedures. In addition, some selected parameters for robustness testing and for on going quality assurance of analytical performance based on control charts are reported.



## 1. Introduction

The suitability of an analytical method for a given application is determined by its reliability and proven by its validation. The prerequisite of reliable analytical results is an analytical method with known performance parameters, based on an appropriate sample representing the entire batch or lot to be analysed, qualified equipment, standard substances and reagents, and an analyst having adequate theoretical and practical knowledge [1]. There are different guidelines dealing with validation in general [1] or in pharmaceutical applications [2–5], all of them initially generated in the mid 1990s. But beside these more basic approaches there are only a limited number of publications giving more detailed instruction on how to perform a validation in practice like Ermer and Miller [6] or Shabir et al. [7,8] for procedures in pharmaceutical analysis. However, interestingly the first guidance on validation analytical procedures ever published dates back to 1985, but due to its title addressing water testing and being only available in German this pioneering work by Funk et al. never got the recognition it deserved [9], even after being published with a more general view addressing quality assurance in analytical chemistry [10,11].

Thin layer chromatography and in particular its high performance application is still a widely used analytical technique in analysis of pharmaceuticals, botanicals, foodstuff, environmental and clinical samples. First examples on how to apply general validation requirements on specific HPTLC procedures have been published in 1993–1995 [12–15]. However, more general guidance on validation of this technique has been published considerably later [16] with special focus on botanicals [17] and pharmaceuticals [18]. Articles reporting analytical methods usually include validation data and give information about the strategy of validation, but a considerable amount of manuscripts do not meet international accepted standards. As a consequence, recently a review was published about the recurrent failures of validation of TLC methods in the field of pharmaceutical analysis [19].

Unfortunately the published discussion on correct validation of planar chromatographic procedures has been dominated by application in pharmaceutical analysis. However, pharmaceutical analysis is a very unique application, as amounts of analyte to be expected or limits are known and thus working ranges and ranges of validation are narrow. As a consequence, validation requirements are extremely stringent and inflexible. Validation is often not considered an iterative process but as a separated, additional step in method implementation. Other fields of application like foodstuff, environmental or herbal analysis with rather unknown or widely varying expected ranges of analyte concentration may need more flexible approaches how to perform validation including more integrated and iterative approaches.

In this work the authors try to give a general survey of validation of TLC procedures independently of the area applied for.

## 2. Prerequisites of validation: verification of analytical equipment

In analytical laboratories, including those using TLC, only regularly controlled, qualified and verified analytical equipment and processes – from the simplest to the most sophisticated ones – should be used. This paper will not cover any aspects of equipment and tools like analytical balances or appropriate glassware;

* Corresponding author. Tel.: +49 160 96 96 32 92.
*E-mail addresses:* bernd.renger@br-consult.eu (B. Renger), z.vegh@richter.hu (Z. Végh), k.fodor@richter.hu (K. Ferenczi-Fodor).

0021-9673/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2011.01.059

only equipment and instrumentation specially used for TLC analysis will be covered.

### 2.1. Sample application

Although semi-automatic and automatic sample application devices of different manufacturers are commercially available, many laboratories still prefer the manual alternative—solutions of analyte and reference standards are applied onto the TLC sorbent layer using Hamilton syringes. Volume accuracy of these syringes should be checked gravimetrically, by weighing the deliverable volume of the thermostatted solvent. Water or other appropriate slow volatile solvents of high density, like bromoform, can be used for this purpose. This process should be repeated at least six times, and a mean value is calculated. In case the difference between the measured and nominal volume is higher than 10% of the theoretical nominal value, a correction factor should be used in calculation of the chromatographic result. Only when using the same volume for all samples the volumetric bias can be neglected.

The performance of semi-automatic and automatic sample applicators is tested by applying at least eight times the same sample solution of a standard substance onto a chromatographic plate. After development the chromatogram is evaluated by spectrodensitometry and the relative standard deviation (RSD) of the readings (peak area or peak height) is calculated. The RSD should not be higher than a given value, about 1–2%, depending on the instruments' specification.

### 2.2. Detection

UV-lamps, predominantly used for semi-quantitative evaluation of chromatograms must be qualified on a regular basis. In $UV_{254}$ detection is based on light-absorption, in $UV_{366}$ the fluorescent light-emission of the substances is observed. The sensitivity of evaluation highly depends on the light intensity of the lamps, which may decrease over time (because of solarization of the lamps' filter and ageing of the fluorescent tubes). The European Pharmacopoeia (Ph. Eur.) specifies a simple and useful functional test for checking the light-intensity of the UV-lamps [20]. Small and known quantities of sodium salycilate are applied onto non-fluorescent silica gel sorbent layer. The sensitivity of the UV-lamp can be accepted if the spots can be detected definitely both in 254 and 366 nm wavelengths by visual inspection.

Similarly, a non-developed "chromatogram", containing sequences of decreasing quantities of a fluorescing standard substance applied in a special pattern, can be used for testing sensitivity and illumination homogeneity in the video densitometric documentation systems.

Up-to-date spectrodensitometers include self-test-routines for the daily check of stage-positioning and their electronics, which is automatically started with each power-on of the equipment. Also scanner validation/qualification packages are used for testing correct working conditions of measuring electronics, lamps, detectors, monochromator and the measuring geometry. Alignment of the lamps and monochromator is automatically performed if necessary. Results are imported by the system software.

These routine checks are to be performed by the analyst independent from regular maintenance and calibration routines.

## 3. Prevalidation processes

### 3.1. Stability of analyte during the chromatographic investigation

When developing a new TLC procedure the most important and therefore the first check is to test the stability of all substances investigated during all steps of the procedure. TLC is an open method; separations are performed mainly on normal phases on highly active polar surfaces. Therefore substances sensitive to humidity or oxygen might undergo hydrolytic or oxidative degradation on the surface of the sorbent layer, or as an effect of the solvent used for sample preparation, pre-treatment or development. Therefore the following parameters should be tested to assure stability of the analyte prior to starting the time-consuming validation process [17,18].

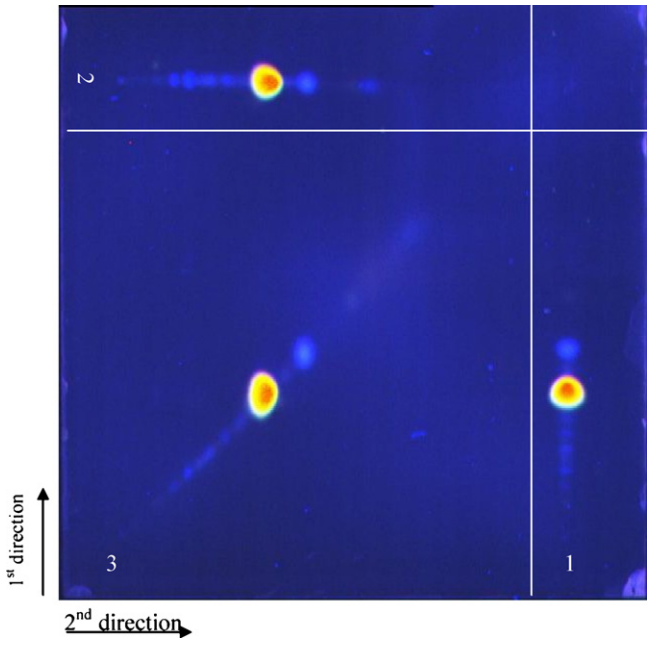


**Fig. 1.** Two-dimensional separation for testing stability of nandrolone decanoate during development according to Ph. Eur.; (1) control spot developed only in the first direction; (2) control spot developed only in the second direction; (3) test-spot developed in both directions. There was no degradation during the development.

(a) Stability of the analyte in solution: application of solutions (analyte, reference substances) stored for different times (0.1–24 h) onto the same plate side-by-side and consecutive development and assessment.
(b) Stability of the analyte adsorbed to sorbent layer before development: plotting analytical solution onto the same sorbent layer at different times (4–<0.1 h) prior to development and assessment.
(c) Stability of the analyte on the sorbent layer *during* development: two-dimensional separation using the proposed eluent system in both directions. In case of no degradation the spots are situated on the diagonal of the chromatogram (Fig. 1), whereas spots originating from degradation in the first run will be observed out of the diagonal.
(d) Stability of the analyte when standing on the sorbent layer after development and before and after visualisation by densitometry or post-chromatographic derivatisation: checking spots of analyte periodically and recording peak areas or peak heights. In this case the degradation cannot cause secondary spots but may induce changes or fading in light-absorption/emission, mainly in case if the calibration substance and the analyte are different.
(e) Stability of the analyte during sample-preparation steps, e.g. in case of long extraction times of active principles from botanicals [17]. This could be part of the optimisation of the sample preparation.

Possible strategies in case of observable degradation of the analyte or reference material during sample preparation or the chromatographic process includes changing the method, switch to

other solvents or sorbents, using antioxidants, excluding day-light or overlay using a nitrogen or $CO_2$ atmosphere during development, until the stability of the analyte is evaluated suitable.

### 3.2. Robustness

There has been an ongoing confusion concerning the terminology when it comes to the parameters ruggedness and robustness. For pharmaceutical applications, there are two guidelines important to any method validation—USP Chapter <1225>, "Validation of Compendial Methods" [5]; and the International Conference on Harmonization (ICH) Guideline "Validation of Analytical Procedures: Text and Methodology Q2 (R1)" [2].

For a long time, ruggedness was defined in the USP as the degree of reproducibility of test results obtained by the analysis of the same samples under a variety of conditions, such as different laboratories, analysts, instruments, reagent lots, elapsed assay times, assay temperatures, and days. The term ruggedness, however, is not used by ICH, but is addressed in guideline Q2 (R1) under intermediate precision (within-laboratory variations; different days, analysts, equipment, and so forth) and reproducibility (between-laboratory variations from collaborative studies applied to the standardization of the method). However, as the USP is considered the sole legal document in the eyes of the FDA and thus pharmaceutical companies in the United States, the term "ruggedness" was used in favour of intermediate precision or reproducibility. In addition, tests to address the method's capability to remain unaffected by deliberate small changes in the procedure were not covered by the USP approach.

Luckily, the USP Convention, as evident in the recent revisions to Chapter <1225> in USP 30 [22], has decided to harmonize more closely with ICH, using the term "intermediate precision", delete all references to ruggedness and introduce the concept of robustness.

Unfortunately the confusion between the two terms is still not totally eliminated, as the Eurachem Guide [1] uses the two at least partially synonymous. To avoid any further irritation, the manuscript will focus on the ICH and USP approach, making "robustness" the topic of this chapter.

The usual approach in pharmaceutical analysis to perform robustness test as part of method validation separately after developing the new method must be heavily challenged. If one or more tested method parameters will show significant effects on the result or analytical performance it will be unavoidable to change the method until it proves to be sufficiently robust and to start validation again. To avoid these unnecessary loops, our approach considers robustness testing as an integral part of method development and the pre-validation.

Based on a risk assessment the experimental and environmental conditions considered having major impact on the chromatographic performance are tested. One has to decide: what has an effect on what? Variables that may have critical impact on TLC procedures are summarized in Table 1.

The robustness testing for a quantitative OPLC determination of vinblastine in *Catharantus roseus* was performed by saturated fractional factorial experimental design [21]. The factors investigated, their levels and calculation of their effects were reported in details.

More general guidance may be found in the papers of Vander Heyden et al., including case studies, describing robustness tests using multifactorial design that are critically reviewed and discussed [23,24].

**Table 1**
Experimental and environmental conditions with potential input on TLC performance characteristics.

| Conditions, having effect on selectivity | Conditions, having effects on quantitative result |
|---|---|
| Relative humidity | Sample-preparation: extraction time |
| Temperature | Stability considerations: standing on the sorbent before and after development |
| Eluent composition | Post-chromatographic derivatisation: heating time, heating temperature |
| Chamber saturation time | Densitometry: detection wavelength |
| Running distance | |
| Amount of the mobile phase | |
| Sorbent quality (TLC/HPTLC) | |
| Sorbent of different manufacturers | |
| Parameters to be tested: Rf, Rs, *T* | Parameters to be tested: DL/QL, quantity of analyte |

## 4. Validation

### 4.1. Validation plan

Before starting validation experiments it is essential to develop a validation plan. This plan should outline the purpose of the proposed procedure to allow proper planning of validation activities.

1. What type of TLC method is going to be validated: a qualitative test for identification or a quantitative assay or determination—limit test, semi-quantitative purity test or quantitative densitometric procedure for purity determination or assay?
2. What assessments and what decisions will be based on the TLC procedure: in R&D stage or for screening purposes or as preliminary method prior to other more precise techniques results might be acceptable showing a higher uncertainty than in official routine QC applications e.g. for testing of foodstuff, drug substances or finished drugs?
3. What is the specification limit to be controlled by the new TLC procedure: defining meaningful specification limits and choosing or developing a method appropriate to control them is an iterative process [1]. Limit should be based on the purpose of the proposed test and the method used; but unfortunately limits expectations by the costumers and formal requirements by authorities may be driven by ambition rather than science, demanding tight specification limits and consequently high precision of the proposed methods.

The ICH guideline to be followed when validating an analytical procedure to be used in pharmaceutical industry [2] gives clear guidance on what validation characteristics to be determined in the course of the validation. This useful approach may be accepted as a basis for all other areas of TLC analysis (Table 2).

In addition, the validation plan should include all standard substances, reagents, solvents, samples, and equipment used as well as a detailed description of the method and define personnel and responsibilities of the validation team.

### 4.2. Validation process

#### 4.2.1. Validation of identity tests

Identity testing is the most common application of TLC. The validation parameter to be checked for acceptance is the specificity of the method. Substances expected to interfere and cause disturbances of spot shape and Rf value like matrix components, are developed simultaneously side-by-side to the substance to be tested on the same chromatographic plate. No other spots or peaks should be detected on similar Rf value of the substance investigated, nor should the peak be distorted. This approach may be used for any

**Table 2**
Types of analytical procedures and required validation characteristics (as requested by ICH Guideline Q2(R1) [2]).

| Type of analytical procedure | Identification | Testing for impurities | | Assay, dissolution (measurement only) content/potency |
|---|---|---|---|---|
| Characteristics | | quantitat. | limit | |
| 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 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.

type of detection—$UV_{254}$ or $UV_{366}$ or any visualisation by pre- or post-chromatographic derivatisation.

The identity of botanicals is proven by not only by Rf-values of the components, but also by similarity of fingerprints which are indicative for the composition and the ratio of the different components, too [17]. In these cases beside the prevalidation and the specificity test, the precision of the method at three concentration levels should also be tested. The observed standard deviation of the Rf-values of the separated zones is determined and limited.

*4.2.2. Validation of quantitative TLC procedures*

Although real quantitative TLC may only be achieved by using objective, instrumental detection and evaluation of the chromatograms by scanning densitometry or videodensitometry, visual evaluation based on comparison of spot sizes and intensities may also provide a quantitative estimation of the analyte. There are two types of visual evaluation in TLC: dichotomic limit tests (yes/no decisions) and the semi-quantitative tests allowing an estimation of the concentration of the analyte.

*4.2.2.1. Limit-test.* The objective of these tests is to allow the decision whether the substance under investigation is present in the sample under or above a given specification limit. Nearly all the pharmacopeial TLC purity tests still contained in Ph. Eur. are limit tests of this kind. In addition, limit tests are widely used as in-process controls and release tests especially in chemical synthesis like API manufacturing [19,25]. The validation of a limit test consists of a specificity test and a check of the detection limit (DL), the smallest quantity of the analyte that can be detected. As visual estimation is quite subjective, at least three analysts should perform this evaluation and the mean value is regarded as DL. The Eurachem Guide [1] recommends 10 determinations. In addition, it should be checked using spots of concentrations of the substance under investigation that exceed or fall below the given specification by app. 20 or 50% whether this difference can be judged reliable.

*4.2.2.2. Semiquantitative estimation.* In many cases, however, the dichotomic "yes/no" decision is not sufficient and more detailed information is needed instantly. In that case a semiquantitative estimation can be performed by visually comparing the spot of the impurity to spots of either the main component or one particular, known and isolated impurity applied in appropriate concentration. In pharmaceutical analysis for example, USP general Chapter <466> "ordinary impurities" [26] describes this test in case the individual monograph of a substance introduces a limit for the sum of all impurities. This sum of all impurities is then calculated by addition of the quantities of the individual impurities to be considered, based on visual comparison with spots of known concentration of the main, monographed component. A "calibration sequence" is constructed using corresponding low concentrations of the main component. If appropriate, instead of the main component a more suited, characterised impurity may be used as reference substance. The quantities of the individual impurities are than estimated by visually comparing the sizes and intensities of the spots to that of calibration spots. For validation, a calibration sequence consisting of at least four spots, the twofold of DL, 50%, 100%, and 150% of the proposed limit has to be constructed. These calibration spots corresponding to different quantities of the analyte should be clearly discriminated form each other by the analysts. Usually, this has to be proven by at least three analysts. The precision of a semiquantitative estimation can be considerably increased by applying more calibration spots of different concentrations.

The current appropriate way to document validations of limit tests and semiquantitative estimations is using state-of-the-art techniques like colour photographs or video documentation.

*4.2.2.3. Quantitative determination.* Quantitative in situ evaluation in TLC is performed by scanning the developed chromatogram using spectrodensitometers, video-scanning or as latest development using MS-detection coupled techniques. Validation of quantitative TLC or HPTLC is basically identical or at least very similar to that of HPLC procedures taking into account the special features of the open TLC technique. Although the ICH Guideline [2] is mandatory only for pharmaceutical applications with focus on approved and licensed active pharmaceutical ingredients (APIs, bulk drug substances) or finished products, its *general* principles may also be applied to all other quantitative determinations based on TLC. However, it must emphasised again, that the very stringent and inflexible ways of interpretations of the requirements of the ICH Guideline as practiced in pharmaceutical industry should not be transferred without criticism to other applications that need more flexible approaches.

According to ICH Guideline Q2 (R1), the validation characteristics to be determined in case of a quantitative impurity testing are as follows: specificity, detection and quantitation limit (DL/QL), accuracy, precision, linearity and range. For an assay procedure in pharmaceutical application with an expected label claim of the analyte, all the parameters except DL/QL are determined.

*4.3. Testing specificity*

For an impurity test the procedure must be shown to be specific and clearly separate all known (characterised) and unknown (not characterised) but expected impurities. If only one specific impurity is to be quantitatively determined, separation of this compound from all other impurities must be shown. Pursuant, in case of an assay procedure the procedure must be shown to be specific for the component(s) to be determined only. These compounds may be active components or APIs of a drug product, a crude natural material, and preservatives or colouring substances in a foodstuff.

The specificity test is nearly the same in all cases: all substances available and expected to interfere or distort should be applied either as pure reference substance or as partially purified material or prepared in situ together with the tested substance on the same chromatographic plate, and Rf and Rs values are determined. This includes by-products of the synthesis or biosynthesis as well as degradation products or components from matrix, excipients or other materials used.

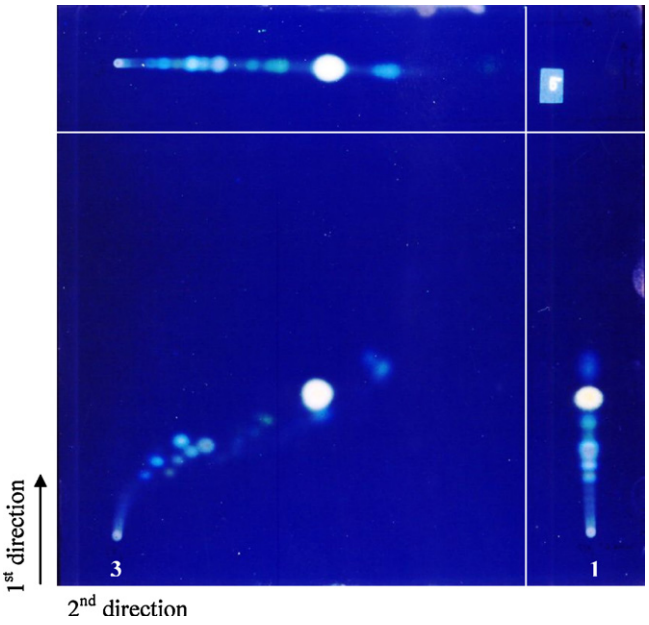


**Fig. 2.** Two-dimensional separation for testing spot-purity of thermally stressed nandrolone decanoate, using different eluents in the two directions. Eluents: 1st direction, heptane–acetone (7 + 3), 2nd direction, chloroform–acetone (95 + 5). Applications 1–3: see in Fig. 1.

If the quantitative TLC procedure is intended to be used in stability studies of a hitherto unknown substance or a not yet tested composition or matrix, prove of specificity must include potential unknown and possible degradation products. In case of new chemical entities this is usually achieved by performing stress tests and testing the corresponding degradation mixtures in parallel as required for *new* pharmaceutical active pharmaceutical ingredients (APIs) by ICH Guidelines Q1A (R2) [27] and Q3A(R) [28]. Degradation products induced by the effect of accelerated temperatures, humidity, acidic, alkaline, oxidative and reductive conditions or light stress are chromatographed beside the original, unstressed sample on the same chromatographic plate. However, there are two common misconceptions when it comes to stress testing:

Unrealistic stress conditions are used not mimicking the conditions to be expected in reality, thus destroying (or "killing") the analyte rather than inducing partial degradation. The authors consider stress causing 5–10% degradation the ideal conditions.

In addition, there are numerous manuscripts describing stress degradation of already known compounds with well characterised degradation products and pathways. ICH Guidelines Q1A (R2) [27] and Q3A(R) [28] clearly refer to *new* entities.

The efficacy of the separation of the degradation products expressed as peak purity of the main component can be checked by two-dimensional separation using eluents of different selectivity in the two directions (Fig. 2). The second, increasingly popular test to assure peak purity test is based on comparative UV-spectra recorded at the start, maximum and end of the peak of the main component in question. This method is suitable, however, only if the UV-spectra are sufficiently different.

### 4.4. Linearity and range

Wherever possible, the concentration range in which the analyte is expected should be taken into consideration. This is relatively easy when it comes to testing of approved or registered pharmaceutical products or any corresponding API with given specifications for assay and impurities. It becomes more demanding for tests in early development and R&D or in case of wide expectation ranges like components in foodstuff or active principles in natural products like botanicals of rather unknown composition and even more challenging for rather undefined compositions like environmental samples.

In TLC the calibration is an extremely critical parameter. The simplest mode would be using linear calibration, but in TLC methods the linear or "pseudo-linear" range of calibration and consequently the working range is significantly narrower than, e.g. in HPLC. This is caused by the fact that in contrast to HPLC, in TLC/HPTLC the Lambert–Beer's law is not effective, as in TLC most of the measurements are performed in diffuse reflectance mode whereas HPLC predominantly uses spectrophotometric detection in the liquid phase. As a consequence, analysts have to find the quasi-linear working range for linear regression [19]. Once the linearity of the working range has been proven in the course of the validation, a three-point calibration may be used when performing routine analysis as proposed by the corresponding monograph on quantitative TLC in the European Pharmacopoeia [29]. The widespread habit of using one point calibration in routine TLC analysis is definitively not encouraged by the authors.

For proving linearity the calibration function should be based on not less than 6 different concentrations of analyte applied, preferably eight or more. Solutions should be made by direct independent weighing and not by diluting one common stock solution, an unfortunately wide spread approach. Application should be performed at least in duplicate, equidistant with more spots at highest and lowest concentration or accumulated at either end of the concentration range; spots should be randomly applied. Calibration spots are quantitatively evaluated and the regression line is constructed by plotting the peak areas or peak heights as appropriate against the applied quantities.

The most widespread malpractice – not only in TLC – besides preparation of the calibration concentrations by diluting one stock solution is the "prove" of linearity based solely on the calculation of the correlation coefficient. Although this approach was declared incorrect and unacceptable numerous times [30–33] this is the still predominant way linearity is considered to be demonstrated. Fig. 3 presents a wide range calibration which apparently is non-linear, but would be accepted to be linear according to the correlation coefficient.

The most convenient way of testing linearity beside a visual assessment is plotting residuals [6,31,34], i.e. the distances of the experimentally determined points from the regression line against the quantities of the analyte applied. If the calibration graph is linear in the tested range, the residuals are randomly distributed around the regression function and do not show any tendency (Fig. 4).

If the calibration is non-linear, the residuals show clear tendency, i.e. a bow shaped trend (see in Fig. 3).

Statistical tests may also be used for proving linearity, like Mandel's test for comparison errors of residuals of quadratic and linear regression by an *F-test* at a chosen significance level or other statistical test like lack of fit test by ANOVA or testing homoscedasticity, i.e. the homogeneity of variations of residuals [8,10,34].

In case of a linear calibration, as a next step the confidence interval of the calibration function should be calculated. The confidence interval of intercept of a linear regression without bias includes the origin ("zero"). If not, results mainly in the lower range of the calibration near to zero may be biased. As a consequence it is extremely important to properly select the correct calibration and working range. If the calibration is definitively non linear, mathematical models of non linear calibrations like quadratic or Michaelis–Menten regression should be used. Mathematical transformation to linear calibrations will lead to severe bias and must be avoided [27]. In this case at least four calibration points should be used in daily routine testing.

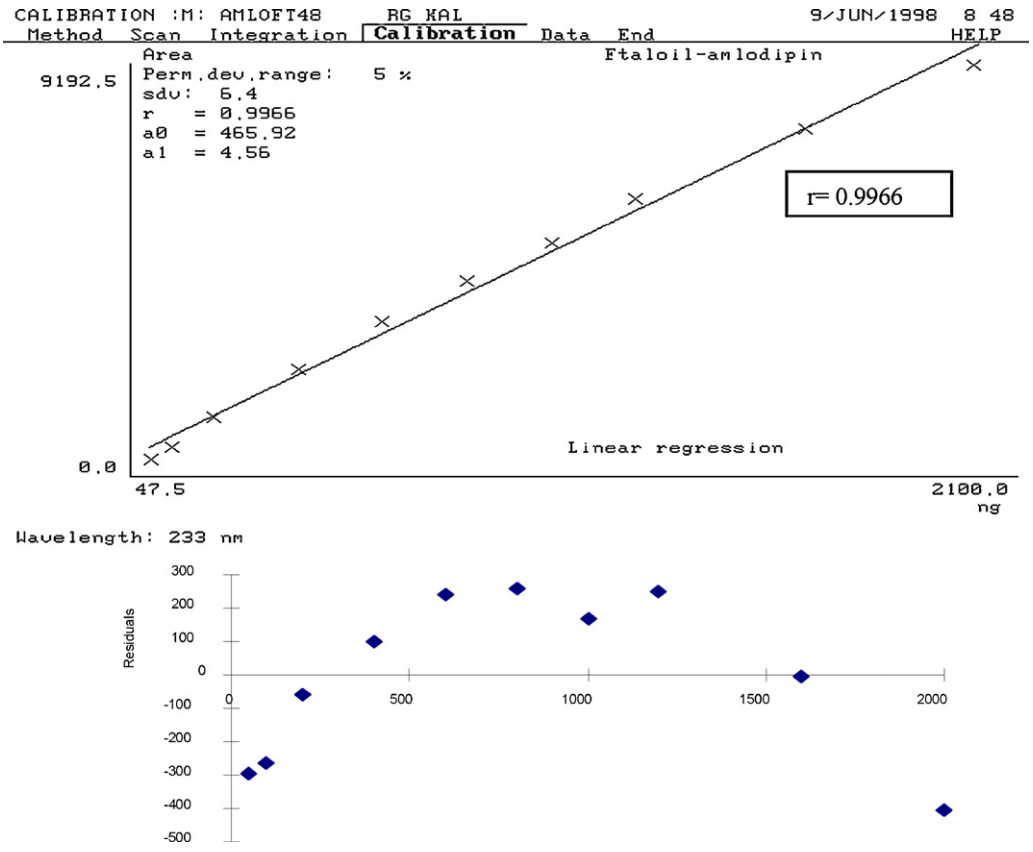


**Fig. 3.** Non-linear calibration, calculated with linear regression. Residuals prove the non-linearity.

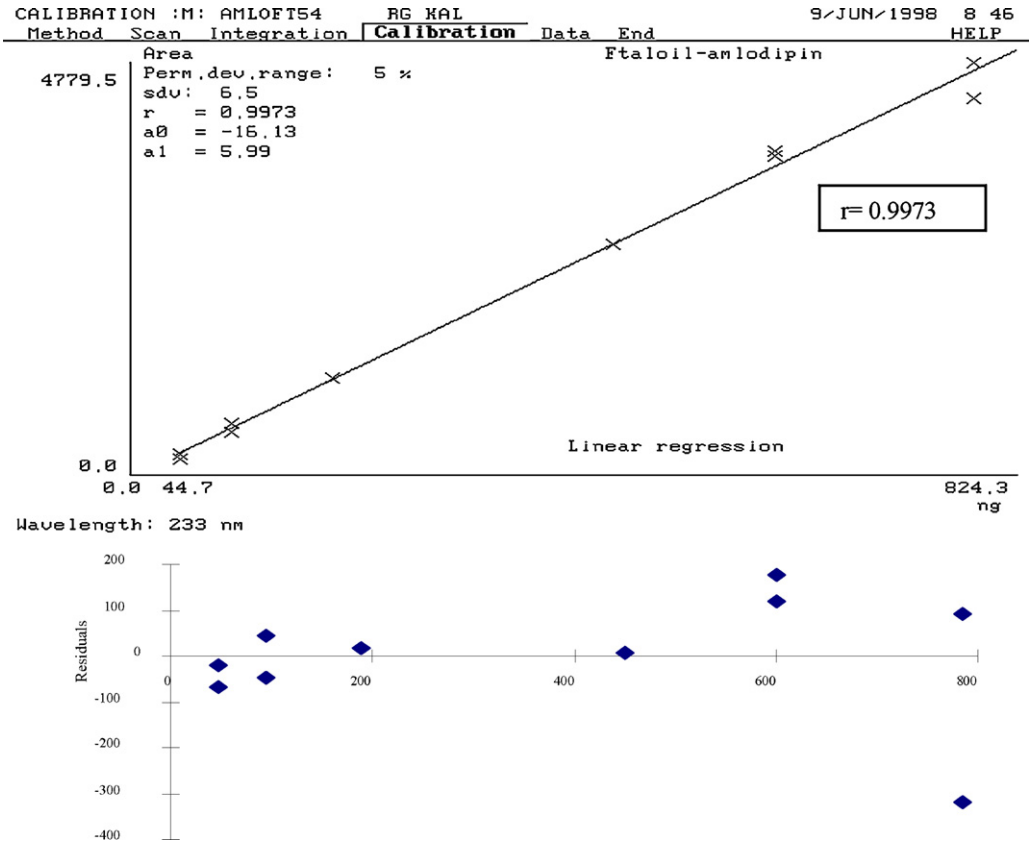


**Fig. 4.** Linear calibration, calculated with linear regression. Residuals prove the linearity.

In pharmaceutical applications like assay of a drug substance or a finished product the working range must cover at least 80–120% of the expected analyte concentration, for determination of an impurity 50% (QL) to 120% of the acceptance criterion must be covered.

### 4.5. Precision of the method

Any analytical result is affected by two different kinds of experimental or operational error—systematic and random error. Systematic error or bias are associated with the accuracy of the procedure and are defined as the difference of the measured mean value from the "true" value.

Random errors are associated with the precision of the procedure. The precision (or imprecision) of the procedure is a combination of the errors of its individual steps: sampling, weighing, sample-preparation, dilution, sample application, development, detection, evaluation and calculation. As a consequence, precision must considered the ". . .degree of agreement among individual test results when the procedure is applied repeatedly to multiple samplings of a homogeneous sample" [5].

According to ICH Guideline Q2A (R2) precision can be determined at three levels [2–5]:

- Repeatability expresses the precision obtained with the same procedure on identical test material under the same operating conditions over a short interval of time. Repeatability is also termed intra-assay precision.
- Intermediate precision expresses within-laboratories variations: precision obtained with the same procedure on identical test material but on different days, different analysts, different equipment, etc. Sometimes termed intra-laboratory precision.
- Reproducibility expresses the precision obtained with the same procedure on identical or standardized test material between laboratories as determined by collaborative studies. Usually applied to standardization of methodology and not part of "routine" validation. Sometimes termed inter-laboratory precision.

#### 4.5.1. Repeatability

The easiest to be examined parameter is a procedure's repeatability. This is performed by assaying not less than 6 independent subsamples that have been processed through the complete analytical procedure from sample preparation to sample application, subsequent chromatography and following evaluation to final test result. In case of a linear calibration not less than 3 concentrations of reference standard, in case of non-linear calibration not less than 4 concentrations of reference standard are applied onto the same chromatographic plate.

The ICH Guideline recommend that repeatability should be assessed using a minimum of nine determinations covering the specified range for the procedure (i.e. 3 concentrations and 3 replicates of each concentration or using a minimum of 6 determinations at 100% of the test concentration).

#### 4.5.2. Intermediate precision

For the second level, the determination of intermediate precision, no clear expectations are published in any guidance document. It has become widely accepted that these experiments are performed on different days by different analysts in the same laboratory using different, but equivalent equipment and devices if available and considered to be used in routine. In this case results are affected by *unintentionally* varied different environmental conditions.

#### 4.5.3. Reproducibility

The third level – reproducibility – is usually not a part of analytical validation, but may be very important in case of analytical transfers.

Repeatability of sample application and/or repeatability of scanning of spots are part of equipment qualification (see above) and must not be misinterpreted as performance parameter describing a procedure's precision [19].

#### 4.5.4. Acceptance criteria

Acceptance criteria for a procedure's repeatability or intermediate precision must be based on the intended use of the analytical method. A repeatability showing a random standard deviation (RSD) of app. 1.5% may be acceptable for assaying a drug substance in a drug product with a given specification range of label claim ±5%. It is, however, disqualifying for use of the procedure to assay the drug substance itself with a given specification range from 98.5% to 101.5%. Acceptable precision data for impurity testing may vary between 5% and 25%, depending on the concentration of the impurity.

For non pharmaceutical applications as assaying active principles in herbs or impurities in food or environment far higher RSDs may be acceptable.

Acceptance criteria for intermediate precision are usually ranging from 1.3 to 1.7 × RSD of repeatability.

### 4.6. Accuracy/trueness of the method

Accuracy provides information about the difference between the mean measured and the "true" value. It defines the bias of the method. The accuracy of the method can be determined by adding known quantities of the analyte to the sample (drug, matrix), i.e. "to spike" it and to calculate accuracy as the percentage of recovery of the known added amount of analyte in the sample by the assay. A recovery of 100% would be the theoretical or added amount.

The sample should be spiked with known quantities of the component to be determined at three different concentrations. As sample preparation and pre-treatment is a vital part of any analytical method the spiking must *always* be performed prior to the essential parts of the sample preparation like extraction, filtering, pre-concentration by heating or by solid phase extraction. Simply adding reference substance to solutions ready to be spotted on the plate or even worse to "overspot" the chromatogram is unacceptable. The original and the spiked samples once processed and transferred to the plate are then separated and evaluated side-by-side on the same chromatographic plate referring to the same calibration samples.

For pharmaceutical applications, the ICH Guideline recommends using a minimum of 9 determinations over a minimum of 3 concentration levels, covering the specified range; i.e. 3 concentrations and 3 replicates of each concentration.

Again, acceptance criteria are not easy to define, but should not be set too ambiguously. For impurity testing, recoveries ranging from 80% to 120% for impurity levels ≤0.5% with an RSD of app. 10% are acceptable.

For non pharmaceutical applications especially when analysing complex matrices like herbal extracts or herbs, food or environmental samples far lower recoveries may be deemed as adequate.

A second possibility for testing accuracy of a proposed method is the comparison of the results of the procedure with those of a second, well characterised and validated analytical procedure with defined or known accuracy. It may be an existing HPLC in case of a quantitative TLC/HPTLC impurity test or assay. The comparison of the results can be performed by *t*-test.

It is an often overlooked requirement in pharmaceutical analysis that replacing a compendial, pharmacopoeial or submitted procedure by an alternative procedure is only allowed and accepted by authorities, if the alternative procedure is at least as reliable and shows performance characteristics as good as the original one. Sim-

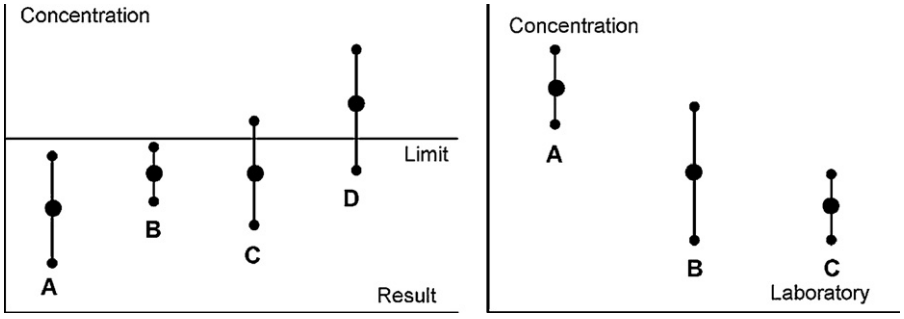


**Fig. 5.** Left: impurity results reported with analytical uncertainty. Results A and B are well below the allowed level, even when their measurement uncertainty is taken into consideration. The mean of result C is below the limit, but there is certain possibility that the maximum allowed concentration is exceeded because of its uncertainty. In food industry, this sample has to be rejected. Result D has the same consequences, although there might be some possibility that the limit is not exceeded. Right: an identical sample is analysed by three different laboratories using the same method. Laboratories A and B will accept the results of each other, and so will laboratories B and C. However, laboratories A and C will run into arguments, as mean values including analytical uncertainty do not overlap. (Taken from [39]).

ply replacing an existing HPLC method by an alternative TLC/HPTLC method only because TLC/HPTLC will offer economical benefits is not appropriate in case the performance of the original method is not met!

### 4.7. Uncertainty of the method

In parallel to the concept of analytical validation the concept of analytical uncertainty was developed in the early 1990s of the last century [35,36]. Initially developed to trace physical measurements back to defined standards or SI units, it soon was adapted to chemical analytical measurements in the Guide "Quantifying uncertainty in analytical measurement" [37]. As a consequence, ISO norm 17025, "General requirements for the competence of testing and calibration laboratories", now demands the determination of the measurement uncertainty of any reported analytical results in its Paragraph 5.4.6 [38]. The concept of measurement uncertainty is an extension of the well known "error analysis and propagation", but takes into consideration not only the random errors, but also sources of bias to be included in this probabilistic approach.

Measurement uncertainty is determined by addition of the variances of the individual steps of an analytical method or by an approach which starts with one of the above-mentioned precision data. The Guide provides explicite directions for the use of validation and related data in the construction of uncertainty estimates. As a result, analytical uncertainties of reported results are usually wider than intermediate precision data, as bias from recovery and other sources is included in the estimate.

The rationale behind reporting the measurement uncertainty is to assure that analytical results independently from the testing and reporting laboratory can be used to decide on acceptance or rejection of a given material or to judge whether legally defined limits are controlled or exceeded. In addition, equivalence of the performance of testing laboratories can be checked and assessed. Results of different laboratories can be shared and accepted in case an interlaboratory test has analytical result including its uncertainty will show some overlap (see in Fig. 5) [39].

Although widely accepted in other fields of application, the concept of measurement uncertainty was never adapted by the pharmaceutical industry. This neglectance was mainly driven by the attitude of industry and regulators, that the public would feel uncomfortable when noticing that analytical results in a sensitive field like health care and drugs may be "corrupted" with an inheritent uncertainty.

Attempts to advocate a more open minded approach and to realise that the reliability of analytical procedures and method used in pharmaceutical industry are widely overestimated [40] showed only limited success.

### 4.8. Detection limit (DL) and quantitation limit (QL)

The detection limit (DL) is the lowest quantity which can be detected with a ß-error (type II or false negative error) of 50%, or in other words the probability that the presence of the analyte cannot be proven is 50%. The quantitation limit (QL) is the lowest quantity which can be determined with a suitable precision [1–5]. Their determination is only compulsory in case of purity tests but may be useful in case of an assay intended for determination of active principles or components in natural products, like botanicals, or pharmacological samples, where the content to be expected is highly varying.

There are two possibilities for determination of these operational characteristics:

(a) A regression line is constructed by spotting decreasing quantities of the substance to be investigated randomly in triplicate, together with a blank (i.e. the pure solvent) application. The standard deviation ($s$) is than calculated based on not less than 20 blank peak heights. Both, DL and QL may than be calculated on the basis of the signal-to-noise ratio using the following equations:

$$\mathrm{DL} = \frac{3s}{S} \quad \text{and} \quad \mathrm{QL} = \frac{10s}{S},$$

where $s$ is the standard deviation of blank peaks' heights and $S$ is the slope of the regression line

(b) The analyte (with matrix!) is repeatedly ($n \geq 5$) spotted in decreasing quantities and the RSD is calculated based on the peak areas or heights for each concentration. The lowest concentration for which the RSD can be accepted (e.g. RSD 10% or RSD 20% depending on use of the procedure) is the QL (Eurachem approach).

Usually, the QL should not be higher than 50% of the specification limit to be controlled using this procedure.

### 4.9. Permanent validation: checking the method in the daily use

After a successful method transfer the validated method is introduced into the routine analytical work, which may be a challenge to any new procedure. Because TLC is an "open system", methods are highly affected by experimental and environmental conditions;

therefore one has to check the goodness of the chromatographic investigation on each chromatoplate.

#### 4.9.1. Checking the selectivity and the sensitivity

For any HPLC or GC method a system-suitability test (SST) has to be performed prior to starting any daily testing of real samples. In TLC the SST is performed simultaneously with testing samples and reference samples on the same chromatographic plates. Results of the SST and results of the samples tested are available simultaneously. The result of the chromatographic test can be accepted if the SST complies to its specifications.

The selectivity of the actual separation can be checked by testing two substances showing closely Rf-values. The system's selectivity can be accepted if these spots or peaks are separated properly. As an example, the TLC purity test of Levonorgestrel of the monograph of the Ph. Eur. [41] uses ethinylestradiole for this purpose.

Controlling sensitivity in routine applications is very important especially in case of methods using post-chromatographic derivatisation, or when the eluent contains low-volatile or strongly adsorbed components like diethyl amine or pyridine. In that case a small quantity of the calibration standard substance, – e.g. the twofold of the DL – is applied and the sensitivity of the chromatogram can be accepted, if the control spot or peak can be detected after development.

#### 4.9.2. Checking the precision and accuracy: monitoring the methods (control charts)

Unintentional or accidental variation of the experimental conditions may have significant effect on the test, leading to results out of normal distribution or expectation. Interfering or disturbing influences can be checked by simultaneous testing of a defined control sample in parallel to the analytical sample on the same chromatographic plate. Control samples are samples of well-known and representative quality, which should have been tested several times, preferably characterised by more than twenty independent results. Based on these data *control charts* can be constructed [8,9,18,42]. If the actual results of the control sample are well within the action limits of the control charts, the result obtained for the sample can be accepted. The accuracy of the actual determination can be checked using an *average chart*, and the precision can be controlled by using a *range chart*.

The use of the control samples allows excellent trending by collecting numerous measurement data over longer periods of time, and RSDs calculated from these data give the real information about the intermediate precision and/or method uncertainty and allow a prediction of the reproducibility of the method.

## 5. Conclusion

Thin Layer Chromatography (TLC) or its high performance variation HPTLC may be used for widely varying purposes and applied to very different matrices. As a result, validation requirements cannot follow a "one size fits all" approach, but must be tailored to address the particular application's needs and characteristics.

Although the pharmaceutical industry was trend setting when defining structured and formalised ways how to perform analytical validation and what performance parameters to cover, this stringent approach might not be suited for TLC/HPTLC applications in other fields like foodstuff, environmental, or clinical analysis. In addition acceptance criteria should be based on the method's intended use and performance rather than on the ambition to demonstrate "how good the laboratory is". An estimation of the method's real uncertainty according to the known guidance may help to avoid unrealistic expectations on reliability and validity of results.

As validation is only a snapshot in a method's life cycle, ongoing trending and monitoring of analytical performance via control samples or control charts is highly recommended.

## References


[1] EURACHEM: the fitness for purposes of analytical methods. A laboratory guide to method validation and related topics. http://www.eurachem.org/guides/pdf/valid.pdf, 1998.
[2] ICH harmonised tripartite guideline: validation of analytical procedures: text and methodology, Q2 (R1), Geneva. http://www.ich.org/LOB/media/MEDIA417.pdf, 2005 (Parent Guideline dated 27 October 1994, Complementary Guideline on Methodology dated 6 November 1996 incorporated in November 2005).
[3] Draft guidance analytical procedures and method validation. US Food and Drug Administration, Centre for Drugs and Biologics, Department of Health and Human Services. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm122858.pdf, August 2000.
[4] CDER reviewer guidance: validation of chromatographic methods, US Food and Drug Administration, Centre for Drugs and Biologics, Department of Health and Human Services. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM134409.pdf, November 1994.
[5] United States Pharmacopeia 33, First Supplement to the USP-NF Reissue, General Chapter <1225>, Validation of Compendial Methods, US Pharmacopoeial Convention, Rockville, MD, April, 2010, R-660.
[6] J. Ermer, J.H.McB. Miller, Method Validation in Pharmaceutical Analysis. A Guide to Best Practice, Wiley-VCH Verlag GmbH & Co, KGaA, Weinheim, 2000.
[7] G.A. Shabir, W.J. Lough, A.A. Shafique, T.K. Bradshaw, J. Liq. Chromatogr. Related Technol. 30 (2007) 311.
[8] A. Shabir, J. Chromatogr. A 987 (2003) 57.
[9] W. Funk, V. Damman, C. Vonderheid, G. Oehlmann, Statistische Methoden in der Wasseranalytik, VCH, Weinheim, 1985.
[10] W. Funk, V. Damman, G. Donnevert, Qualitätssicherung in der Analytischen Chemie, VCH, Weinheim, 1992.
[11] W. Funk, V. Damman, G. Donnevert, Quality Assurance in Analytical Chemistry, VCH, Weinheim, 1995.
[12] K. Ferenczi-Fodor, Z. Végh, Zs. Pap-Sziklay, J. Planar Chromatogr. 6 (1993) 198.
[13] G. Szepesi, J. Planar Chromatogr. 6 (1993) 187.
[14] B. Renger, H. Jehle, M. Fischer, W. Funk, Pharm. Ind. 56 (1994) 993.
[15] B. Renger, H. Jehle, M. Fischer, W. Funk, J. Planar Chromatogr. 8 (1995) 269.
[16] G. Indrayanto, M. Yuwono, in: J. Cazes (Ed.), Encyclopedia of Chromatography, 2nd ed., Frances & Taylor LLC, 2005, p. 1759.
[17] E. Reich, A. DeBarr, J. AOAC Int. 91 (2009) 13.
[18] K. Ferenczi-Fodor, Z. Végh, A. Nagy-Turák, B. Renger, M. Zeller, J. AOAC Int. 84 (2001) 1265.
[19] K. Ferenczi-Fodor, B. Renger, Z. Végh, J. Planar Chromatogr. 23 (2010) 173.
[20] Monograph 2.01.03, Ultraviolet Ray Lamps for Analytical Purposes, European Pharmacopoeia (suppl. 6.8), 6th ed., European Pharmacopoeia Commission, Strasbourg, France, 2010, p. 21.
[21] A. Nagy-Turák, Z. Végh, K. Ferenczi-Fodor, J. Planar Chromatogr. 8 (1995) 188.
[22] United States Pharmacopeia 30, General Chapter <1225>, Validation of Compendial Methods, US Pharmacopoeial Convention, Rockville, MD, 2007, p. 680.
[23] B. Dejaegher, Y. Vander Heyden, J. Chromatogr. A 1158 (2007) 138.
[24] B. Dejaegher, Y. Vander Heyden, LC–GC Europe (July) (2006) 418.
[25] K. Ferenczi-Fodor, Z. Végh, B. Renger, J. Chromatogr. A 1218 (2011) 2722.
[26] United States Pharmacopeia 32, General Chapter <466>, Ordinary Impurities, US Pharmacopoeial Convention, Rockville, MD, 2009, p. 162.
[27] ICH harmonised tripartite guideline: stability testing of new drug substances and products Q1A (R2), Geneva. http://www.ich.org/LOB/media/MEDIA419.pdf, 2003 (Parent Guideline dated 30 March 1995).
[28] ICH harmonised tripartite guideline: impurities in new drug substances, Q3A (R1), Geneva. http://www.ich.org/LOB/media/MEDIA422.pdf, 2006 (Parent Guideline dated 16 September 1992).
[29] Monograph 2.02.27, Thin-Layer Chromatography, European Pharmacopoeia (suppl. 6.8), 6th ed., European Pharmacopoeia Commission, Strasbourg, France, 2010, p. 43.
[30] S. Ebel, Analytical Chemistry Progress, Springer, Berlin-Heidelberg, 1984, p. 71.
[31] Analytical methods committee: is my calibration linear? AMC Technical Brief No. 3. http://www.rsc.org/amc, December 2000.
[32] J. Ermer, in: J. Ermer, J.H.McB. Miller (Eds.), Method Validation in Pharmaceutical Analysis. A Guide to Best Practice, Wiley-VCH Verlag GmbH & Co, KGaA, Weinheim, 2005, p. 88.
[33] J.V. Loco, M. Elskens, C. Croux, H. Beernaert, Accredit. Qual. Assur. 7 (2002) 281.
[34] J.W. Einax, M. Reichenbächer, Anal. Bioanal. Chem. 384 (2006) 14.
[35] Guide to the Expression of Uncertainty in Measurement, ISO, Geneva, 1993. ISBN 92-67-10188-9 (Reprinted 1995).
[36] Joint committee for guides in metrology: evaluation and measurement data—guide to the expression of uncertainty in measurement, vol. 100. http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf, 2008.

[37] S.L.R. Ellison, M. Rösslein, A. Williams (Eds.), Eurachem/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, 2nd ed., 2000. http://www.eurachem.org/guides/pdf/QUAM2000-1.pdf.
[38] EN ISO/IEC 17025, General Requirements for the Competence of Testing and Calibration Laboratories, EN Management Centre, Brussels, 2005.
[39] V.R. Meyer, J. Chromatogr. A 1158 (2007) 15.
[40] B. Renger, J. Chromatogr. B 745 (2000) 167.
[41] Monograph Levonorgestrel, European Pharmacopoeia (suppl. 6.8), 6th ed., European Pharmacopoeia Commission, Strasbourg, France, 2010, p. 2266.
[42] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.