Journal of Chromatography B, 877 (2009) 2224–2234

Contents lists available at ScienceDirect

Journal of Chromatography B

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

Key aspects of analytical method validation and linearity evaluation☆

Pedro Araujo*

National Institute of Nutrition and Seafood Research (NIFES), PO Box 2029, Nordnes, N-5817 Bergen, Norway

ARTICLE INFO

Article history:
Received 2 July 2008
Accepted 29 September 2008
Available online 2 October 2008

Keywords:
Validation
Terminology
Selectivity
Specificity
Accuracy
Precision
Linear function analysis
Range
Limit of detection
Limit of quantitation
Ruggedness
Robustness

ABSTRACT

Method validation may be regarded as one of the most well-known areas in analytical chemistry as is reflected in the substantial number of articles submitted and published in peer review journals every year. However, some of the relevant parameters recommended by regulatory bodies are often used interchangeably and incorrectly or are miscalculated, due to few references to evaluate some of the terms as well as wrong application of the mathematical and statistical approaches used in their estimation. These mistakes have led to misinterpretation and ambiguity in the terminology and in some instances to wrong scientific conclusions. In this article, the definitions of various relevant performance indicators such as selectivity, specificity, accuracy, precision, linearity, range, limit of detection, limit of quantitation, ruggedness, and robustness are critically discussed with a view to prevent their erroneous usage and ensure scientific correctness and consistency among publications.

1. Introduction

The word validation originates from the Latin validus meaning strong, and suggests that something has been proved to be true, useful and of an acceptable standard. The International Organization for Standardization defines validation as the confirmation by examination and provision of objective evidence that the particular requirements for a specified intended use are fulfilled [1]. This definition primarily implies that a detailed investigation has been carried out and gives evidence that an analytical method, when correctly applied, produces results that are fit for purpose as well as it confirms the effectiveness of the analytical method with a high degree of accuracy.

The importance of method validation has been emphasised since the late 40's when the American Chemical Society and Merck & Co., raised the issue of how mathematics and statistics are a necessary prerequisite to successful development and adaptation of new analytical methods [2,3]. By that time a survey of papers on development of analytical methods revealed that no comparisons were carried out with other or similar methodologies in order to check for accuracy in the reviewed articles [4]. In addition, it was pointed out that statistical data analysis was a subject neglected by chemists developing experimental methods [3]. In the early 70's a series of articles were published stressing the need of implementing a consistent set of definitions for determining the performance-characteristics of developed analytical methods and comparing unambiguously the advantages and disadvantages of the increasing volume of reported analytical methods [5–8]. This paved the way for the implementation of method validation in analytical laboratories since the late 70's and the current worldwide recognition that method validation is an important component in any laboratory engaged in the development and establishment of standard methods. Nowadays, there are several international renowned organisations offering guidelines on method validation and related topics. Basic references are the Association of Official Analytical Chemists (AOAC), the American Society for Testing and Material (ASTM), the Codex Committee on Methods of Analysis and Sampling (CCMAS), the European Committee for Normalization (CEN), the Cooperation on International Traceability in Analytical Chemistry (CITAC), the European Cooperation for Accreditation (EA), the Food and Agricultural Organization (FAO), the United States Food and Drug Administration (FDA), the International Conference on Harmonization (ICH), the International Laboratory Accreditation Cooperation (ILAC), The World

☆ This paper is part of a special issue entitled "Method Validation, Comparison and Transfer", guest edited by Serge Rudaz and Philippe Hubert.

* Tel.: +47 95285039; fax: +47 55905299.
E-mail address: pedro.araujo@nifes.no.

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.09.030

Health Organization (WHO), the International Organization for Standardization (ISO), the International Union of Pure and Applied Chemistry (IUPAC), the United States Pharmacopeia (USP), The analytical chemistry group EURACHEM, etc.

Articles are submitted for publication every year highlighting the determination of the parameters for method validation enforced by any of the above-mentioned regulatory bodies. However, in spite of the volume of articles submitted and published misinterpretation and miscalculation still persists due to some prevailing ambiguity in the definitions of some of the validation parameters, few existing guidelines to estimate some of them, and most importantly lack of attention to the mathematical and statistical tools involved in their calculation.

This article discusses some key aspects that should be considered when validating analytical methods, especially those concerning chromatography methods, in order to derive useful information from experimental data and to draw robust conclusions about the validity of the method.

Although the aspects described in this article apply to all types of analytical methods, in some instances they may not be applicable to certain analytical procedures. For instance, some animal models bioassays or some immunoassays have unique features which should be considered before submitting a validation report.

2. The general method validation steps

In a general context, method validation is a process that consists of at least five distinct steps, namely: system qualifications, sampling, sample preparation, analysis and data evaluation.

2.1. System qualifications

A general evaluation of system qualifications allows to verify that the instrument is suitable for the intended analysis, the materials (reagents, certified references, external and internal standards, etc) are suitable for use in analytical determinations, the analysts have the proper training and qualifications and previous documentation such as analytical procedures, proper approved protocol with pre-established acceptance criteria has been reviewed. If the general qualifications of a system are ignored and a problem arises, the source of the problem will be difficult to identify [9].

2.2. Sampling

The sampling step assists in the selection of a representative fraction of the material which is subsequently subjected to investigation. The choice of an appropriate sampling method is of great importance because it provides assurances that the sample selected is truly representative of the material as a whole for the purpose of meaningful statistical inferences. Within the statistical literature, there is a substantial body of work on sampling strategies, however the relative costs and time involved in each strategy should be evaluated in advance.

2.3. Sample preparation

Sample preparation is a key element to successful method validation. It has been pointed out that sample preparation represents 60–80% of the work activity and operating costs in an analytical laboratory [10]. The literature on sample preparation is ample and well documented. However, the analyst should remember that the selection of a specific preparation procedure depends upon the analytes, the analytical concentrations, the sample matrix, the sample size and the instrumental technique.

2.4. Analysis

The analysis is related to the instrument used to extract qualitative or quantitative information from the samples with an acceptable uncertainty level. The analysis could be visualised, in a broad sense, as a system possessing three interconnected basic elements, namely input → converter → output. In general, the input and output are designated by the letters x and y and they represent the concentration and the experimental response respectively. The choice of a particular analysis is based on many considerations, such as the chemical properties of the analytical species, the concentration of the analytes in the sample, the matrix of the sample, the speed and cost, etc.

2.5. Data evaluation

The main purpose of data evaluation is to summarise and gain insight into a particular data set by using mathematical and statistical approaches. Data evaluation allows extracting useful information and drawing conclusions about the inputs and outputs, and most importantly about the validation process in general.

3. Validation method parameters

In the early 80's, it was pointed out that the definition of the characteristic parameters for method validation and related topics were different between the existing organizations [11]. In 1990, the International Conference on Harmonization (ICH) was created as a unique project to bring together the regulatory authorities of Europe, Japan and the United States with the objective of achieving greater harmonization of parameters, requirements and, to some extent, also methodology for analytical method validation. The key criteria defined by the ICH and by other industrial committess and regulatory agencies around the world for evaluating analytical methods are: selectivity/specificity, accuracy, precision, linearity, range, limit of detection, limit of quantitation, ruggedness, and robustness.

3.1. Selectivity and specificity

The terms selectivity and specificity have been the subject of intensive critical comments essentially focusing on the ways in which they are often defined by analysts working in method validation [12–15]. By definition the selectivity refers to the extent to which a method can determine a particular analyte in a complex mixture without interference from other components in the mixture [16]. This definition is often wrongly used as equivalent to specificity, which is considered to be the ultimate in selectivity; it means that no interferences are supposed to occur [12]. Unfortunately, an inspection of the literature on method validation revealed that both terms are still used without distinction by some authors, even though by consulting the dictionary it is clear that these terms should not be used interchangeably. Selectivity should be connected with the word 'choose' while specificity with the word 'exact'. In this context, it is incorrectly to grade the term specificity (either you have it or you do not). An analyst involved in method validation should always remember that selectivity can be graded as low, high, bad, partial, good, etc., in order to choose the appropriate category for a particular purpose. The term specificity refers always to 100% selectivity [13,17,18] or, conversely, 0% interferences.

3.1.1. Experimental approaches to assess selectivity

During the last decade, some researchers have expressed concern about the lack of comprehensive recommendations from

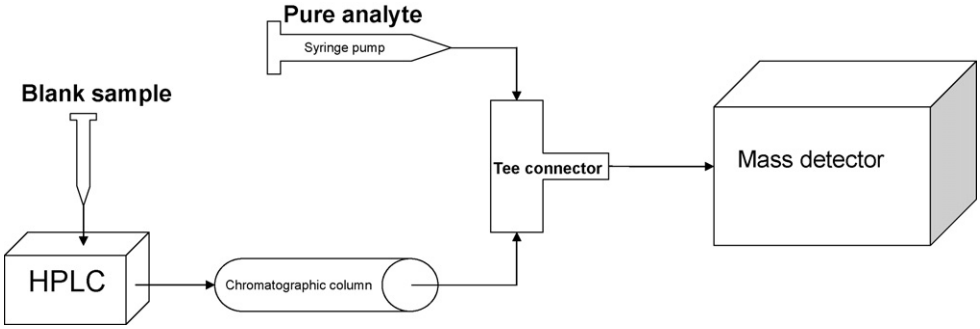


Fig. 1. Liquid chromatography mass spectrometry post-column infusion system.

accredited bodies, books and published articles on how to assess the selectivity of a method [14,19]. A non-exhaustive survey of papers on chromatographic method validation published in analytical journals during the last 10 years revealed that the three most popular approaches used to measure selectivity are:

1. Comparison of the chromatograms obtained after injection of blank samples with and without the analytes.
2. Comparison of the chromatographic response obtained after injection of analytical solutions with and without all the possible intereferents.
3. Analysis of certified reference materials.

Other approaches used in a lesser extend in the last 10 years are:

1. Calculation of the chromatographic response factor by discriminating the analytical species from closely related structures.
2. Comparison with certified methods.
3. Comparison of the slopes obtained by the standard addition and the external standard methods.

In addition to these previously well-known mentioned strategies, an interesting liquid chromatography mass spectrometry post-column infusion technique that enables detecting specific endogenous sample components that affect the target analytes has been proposed [20]. This approach uses a syringe pump and a HPLC system simultaneously coupled to a mass spectrometer through a tee connector (Fig. 1). The flow from the syringe pump delivers a constant amount of the analytes while the flow from the HPLC delivers a processed blank sample, in that way it is possible to study dynamically the effect of the matrix on the analytical responses over the entire chromatographic run, when different sample treatments, columns and mobile phases are used. An example from the literature is the evaluation of the influence of different extraction techniques on matrix effects and consequently the magnitude of these effects on the signal of sirolimus, an immunosuppressant that under specific chromatographic and mass spectrometry conditions elutes at approximately 6 min and presents a characteristics transition at m/z 931.6 → 864.6 [21]. The comparison of the various infusion chromatograms is showed in Fig. 2 which allows concluding that the observed signal suppression is mainly due to endogenous components in the whole blood and that the analytical signal is less prone to matrix interferences around the elution time when the solid phase extraction method is used.

Readers interested in analytical procedures to achieve selectivity are referred to comprehensive articles on the subject published elsewhere [14,17,22].

3.2. Accuracy

Accuracy is the degree of agreement between the experimental value, obtained by replicate measurements, and the accepted reference value. It has been pointed out that the accuracy is the most crucial aspect that any analytical method should address [23]. The determination of this parameter allows estimating the extent to which systematic errors affect a particular method. The preceding definition of accuracy is in accordance with several regulatory

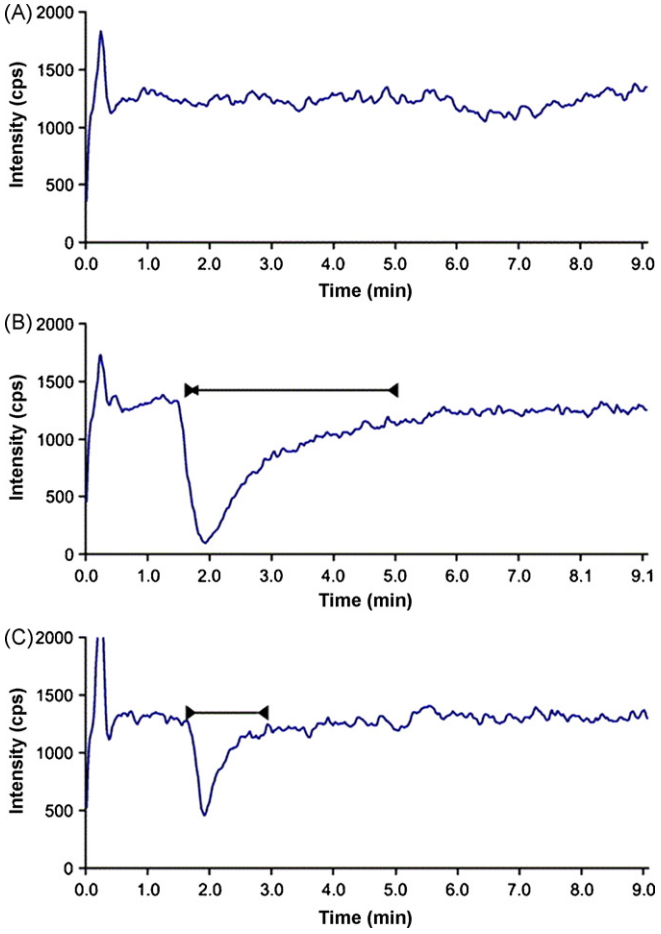


Fig. 2. Comparison of (A) mobile phase, (B) whole blood sample prepared by protein precipitation, and (C) a whole sample prepared by solid phase extraction by the post-column infusion method. The areas influenced by matrix effects are shown in B and C. The solid lines indicate the regions of altered ionization due to matrix effects. Reprinted from P.J. Taylor, Matrix Effects: The Achiles Heel of Quantitative HPLC-ESI-Tandem-MS. Clinical Biochemistry, 38 (4) (2005): 328–334 with permission from Elsevier.

Table 1
Requested conditions to measure the different precision components according to the ISO/DGuide 99999 [33].

Condition		Repeatability	Intermediate precision	Reproducibility*
Same	Analyst	yes	yes/no	yes/no
	Procedure	yes	yes	yes/no
	Instrument	yes	yes	yes/no
	Instrumental conditions	yes	yes	yes/no
	Laboratory	yes	yes	yes/no
	Material & reagents	yes	yes	yes/no
	Day	yes	no	yes/no

*One or more conditions should be changed.

bodies (ICH, FDA, and USP) and the IUPAC. However, it is important to mention that this definition is acknowledged by other accredited organisations as trueness or bias and the term accuracy as the combination of trueness and precision (ISO, EURACHEM, AMC). The concept of trueness as is stated in ISO-5725-1 has been invented in order to address some philosophical objections among medical and legal practitioners regarding the concept of statistical bias [24,25]. In this article, the term accuracy as defined by ICH, FAD, USP and IUPAC is used because it seems to be the preferred term in scientific journals. A ScienceDirect search using the keywords "validation trueness" and "validation accuracy" showed that only 55 articles used the former keywords while 3876 the latter. The difference becomes more dramatic when the search is performed without using the word "validation".

Several approaches have been suggested to evaluate the accuracy of a method. The main strategies currently used to assess the accuracy are:

1. Measuring the analyte in a particular reference material and comparing the result with the certified value.
2. Measuring the analyte in blank matrix samples spiked with known analytical concentrations and determining the percentage of recovery.
3. Comparing the results from the method under validation with those from a reference method.
4. Determining the analytical concentration in the sample by means of the standard addition technique.

The first strategy should be used as long as the reference material closely resembles the analytical concentration and the matrix of the sample under investigation.

The second strategy, despite its popularity, it has the disadvantage that the accuracy can be misestimated if the spiked analyte and the analyte contained in the sample behave differently due to dissimilarities in their chemical form and reactivity. For speciation purposes this approach is not recommended [26]. The IUPAC Harmonised Guidelines for In-House Validation of Methods of Analysis Technical Report recommends the use of the second approach only in cases where the method under validation is intended either for liquid samples or samples subjected to total destruction or dissolution [27].

A 'reference method' in the context of the third strategy refers to a nationally or an internationally fully validated method with different or similar measurement principles and sources of errors.

The fourth strategy is generally used in cases where blank samples are unavailable.

Another reported strategy to assess the accuracy of a method has been taking part in proficiency test schemes in order to compare the results of a particular method under validation with the consensus value obtained by the participating laboratories [28,29]. Unfortunately, the seconders of this proposal have not taken into account that proficiency testing is aimed at monitoring performance and competence of individual accredited laboratories [30] rather than assessing the accuracy of newly developed methods. It is irrational, and contrary to the general requirements for participating in proficiency testing, to take part in such schemes without a fully validated method.

The guidance for validation of analytical procedures issued by the ICH recommends checking the accuracy by performing a minimum of nine determinations over a minimum of three concentration levels (low, medium and high) corresponding to the whole analytical range investigated (3 levels × 3 replicates per level = 9 determinations) [31]. Although, this minimum requirement is suitable in general for chromatographic or spectroscopy methods, the analyst should follow the recommendations suitable for his/her particular method. For instance, the FDA guidance for validation of bioanalytical methods suggests evaluating the accuracy by measuring a minimum of three concentration levels prepared in pentaplicate in the range of expected concentrations (3 levels × 5 replicates per level = 15 determinations) [32].

3.3. Precision

The term precision is defined by the ISO International Vocabulary of Basic and General Terms in Metrology (ISO-VIM) as the closeness of agreement between quantity values obtained by replicate measurements of a quantity under specified conditions [33]. The determination of this parameter is one of the basic steps in the process of achieving repeatability and reproducibility in method validation. Assessing the precision implies expressing numerically the random error or the degree of dispersion of a set of individual measurements by means of the standard deviation, the variance or the coefficient of variation.

The regulatory bodies emphasize that the terms precision and accuracy should not be used as synonyms. A literature review of the last decade demonstrated that analysts engaged in validating methods are well aware of the difference between these terms. However, the review revealed that in many instances, the terms repeatability, reproducibility and intermediate precision are used interchangeably. The reader should keep in mind that the word repeatability in the present context, refers to obtaining the magnitude of a particular property of a sample more than once by keeping constant the

global factors (human, preparation, instrumental and geographical) over a short period of time; the term reproducibility refers to reproduce the magnitude of an already measured property by changing one or more of the global factors over a short or an extended period of time; and the term intermediate precision refers to obtaining the magnitude of a particular property of a sample more than once by using the same preparation, instrumental and geographical condition over an extended period of time. The differences between the three above mentioned precision components showed in Table 1 allow to derive the following conclusions:

- It is wrong to report a so-called "inter-day repeatability" term. Such a term should never be used in method validation.
- The term inter-day variation should be connected with intermediate precision or in some circumstances with reproducibility.
- The repeatability always ensures that the variability of the results remains constant under identical conditions and also that factors such as analyst, procedures, instrumental conditions, laboratories, and time have a negligible contribution to the variability of the results.
- The reproducibility always makes certain that the variability of the results remains constant under different conditions and that the variation of one or more factors does not contribute significantly to the variability of the results.

For validation of chromatographic methods, it is generally recommended checking the precision component by measuring a minimum of three concentration levels (low, medium and high) prepared in triplicate and covering the whole analytical range under study (3 levels × 3 replicates per level = 9 determinations) [30]. However, as was mentioned before, the readers must consider that minimum criteria vary according to the nature of the analytical procedures. For instance, the minimum number determinations to check the precision of a biological or a biotechnological method may be different from the minimum established in the present article.

3.3.1. Total error

Total error, a concept introduced 34 years ago, describes the overall error that may occur in a test result due to the simultaneous contribution of random and systematic error components of the measurement procedure [34]. In general terms, the total error (E_{total}) could be defined by the expression:

$$E_{total} = E_{systematic} + E_{random}$$

where the term E represents the error and the associated systematic and random subscripts define the individual error contribution. From the previous expression, it is evident that the mentioned differences between the regulatory bodies regarding the definition of the parameter that allows to estimate the systematic errors ($E_{systematic}$), termed for instance "accuracy" or "trueness", make mandatory the use of appropriate documentation of the particular definition used in the calculation of the total error. By documenting the specific definition used to express the systematic errors it is possible to derive alternative expressions for total error and consequently avoid ambiguities between definitions. For instance, the following alternative expressions can be found in the current literature on validation:

$E_{total} = Accuracy + Precision$ (using ICH terminology);

$E_{total} = Trueness + Precision$

or

$E_{total} = Accuracy$ (using ISO terminology).

The importance of estimating a total error in validation studies is that it provides a measure of quality that can be compared to the intended analytical quality of a test, which can be described in terms of an allowable total error [35] which sets a nominal limit for the total error tolerable in a single measurement or single test result. It is advisable, when determining the total error of a particular assay, to follow the recommended criteria. For instance, recent reports addressing the issue of best practices for ligand binding assays have established a nominal total error of 30% for this kind of assays based on their inherent variability [36,37]. In addition, it is compulsory to include all assay runs in the calculation of the total error. Rejection of particular assay runs in the calculation of the total error is allowed only in cases, where obvious and well-documented errors arise [36].

An important consideration is that total error is a quality parameter derived from two performance parameters (precision and accuracy) which contribute mutually to the quality of a test result. It is desirable to set goals for the allowable total error, rather than set individual goals for the allowable standard deviation and the allowable bias.

3.3.2. Uncertainty

The uncertainty of measurement is a parameter, associated with the result of a measurement that characterises the dispersion of the values that could reasonably be attributed to measurand (the quantity being measured) [33]. This parameter is estimated in order to judge the adequacy of a result for its intended purpose and to verify its consistency with other similar results.

Countless studies have reported the determination of the uncertainty of measurement results by using different approaches. However, the lack of consensus among the various approaches along with the absence of a worldwide-accepted criterion for the estimation and notation of measurement uncertainty were the main driving factors behind the development in 1980 of the general recommendation INC-1 for the expression of experimental uncertainties [33]. INC-1 states that the uncertainty in the results of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated.

- Uncertainty Type A is the evaluation of uncertainty by statistical methods. The components of this category are determined by means of the variances or standard deviations. The evaluation of this uncertainty applies to random and systematic errors.
- Uncertainty Type B is the evaluation of uncertainty by other means. The components of this category are usually based on scientific judgement using all the relevant information available. For instance, uncertainty assigned to reference data taken from handbooks, reference standards certificates, etc.
- Combined uncertainty is characterised by the summation of the Type A and Type B variances and expressed as standard deviations.

The task of determining the uncertainty of a particular measurement method requires the participation of experts familiar with the various aspects involved in the recommendations issued by accredited bodies. For instance, a laboratory interested in applying ISO 5725-3 recommendation [38] in the evaluation of the uncertainty of a chromatographic method will require the assistance of an expert familiar with the measurement method and its application and also an expert with experience in the implementation of the guidance recommended fully nested design and its statistical analysis.

It should be noted that even though the author of the present article has reported several studies on the determination of the uncertainty in linear calibration, central composite designs and

nested designs for proficiency testing, it is not his intention to present a detailed coverage of the subject. The interested reader is referred to the comprehensive work on uncertainty by Kimothi [39].

3.4. Linear function analysis

Linear function analysis is an area familiar to everyone involved in calibration experiments and perhaps this familiarity has led to the introduction of some fatal flaws when the linearity of a calibration is assessed. Nowadays, such erroneous procedures are described in oral presentations, laboratory reports and papers and accepted as correct.

Linearity, in the context of the previously described analysis system consisting of the basic elements input → converter → output, is the assumption that there is a straight line relationship between the input (x) and output (y) variables that can be written mathematically by the expression $y = f(x)$ if the straight line crosses through the origin or by the expression $y = f(x) + \delta$ if the straight line does not cross through the origin.

It is common practice to check the linearity of a calibration curve by inspection of the correlation coefficient r. A correlation coefficient close to unity ($r = 1$) is considered sufficient evidence to conclude that the experimenter has a perfect linear calibration. Although the Analytical Methods Committee and some articles on analytical validation discouraged using the correlation coefficient in the context of testing for linearity [40–43], many laboratories around the world base the linearity of their instrumental methods on a so-called (by the author of the present article) "r-test". Countless published papers reinforce the idea, perhaps indirectly, of using the r-test to check for linearity by reporting in their abstracts, discussions and conclusions statements such as

"This method has a linear calibration range of 1.00–1000 ng/ml with a correlation coefficient of 0.9999."
"The calibration graphs were linear with correlation coefficients greater than 0.999 for all compounds."
"It was clear that the calibration was linear as a result of a correlation coefficient close to 1 ($r_{experimental} = 0.9996$)."

The author of the present article is not stating that the previous quotations are incorrect, however it must be said that in some respect they are misleading in the context of linearity evaluation. In addition, they fail to indicate which statistical methods were used to evaluate their linear relationship.

The FDA guidance for validation of analytical procedures [31] which is based on the Text on Validation of Analytical Procedures issued by the ICH [44], recommends that the correlation coefficient should be submitted when evaluating a linear relationship and that the linearity should be evaluated by appropriate statistical methods. This guidance does not suggest that the numerical value of r can be interpreted in terms of degrees of deviation from linearity. Hence, it is extremely important to emphasise that an "r-test" to check for linearity does not exist. We cannot say that $r = 0.999$ is more linear than $r = 0.997$. It is surprising that despite the recommendations of the accredited bodies, few published articles have reported the use of statistical methods for linearity evaluation. This observation seems to indicate that the issue, pointed out over half a century ago, about statistical data analysis being a neglected subject by practitioners validating methods [3] is still a topic of contemporary relevance that needs an imperative attention by those currently engaged in validation of analytical methods and by those responsible for educating and training people to be embarked upon the various aspects of this important area of analytical chemistry.

3.4.1. Linear calibration function

The expression $y = f(x) + \delta$ can be rewritten as

$$\hat{y}_i = \phi x_i + \delta \quad (1)$$

where $\hat{y}_i$ and x_i represent the estimated experimental response and the analytical concentration respectively, both at a concentration level i. The coefficients ϕ and δ represent the sensitivity of the analysis and the intercept respectively.

The calibration function described by Eq. (1) must have an r close to 1 but the condition given by "if *linear*, then $r = 1$" is not equivalent to "if $r = 1$, then *linear*". In logical analysis terms, the best strategy to show that a condition is not equivalent to its converse is to provide a counterexample to it. To find a counterexample to "if *linear*, then $r = 1$" it suffices to show that there is something that has $r = 1$ but it is not linear. Appendix A provides a comprehensive counterexample based on statistical analysis of the various error sum squares and variance components from a linear calibration data set reported in the literature [45]. Before studying Appendix A, some important aspects such as replication, error sum squares and degrees of freedom should be discussed in advance.

3.4.1.1. Replication. Replication is an important aspect that must be considered when the experimenter wants to test if a particular experimental calibration model, for instance Eq. (1), is linear. The experimenter must have a reasonable number of standard solutions and instrumental replicates. It has been pointed out that the best calibration strategies are those with standard solution replicates higher than instrumental replicates [42,46]. A preparation error 26 times higher than the instrumental error has been reported in calibration experiments of triacylglycerols by LCMS [47]. This implies that the most serious problems are related to preparation and not to instrumental stability. Calibration experiments with only one standard per concentration level are a poor calibration strategy and must be avoided unless the standard solutions are effectively error-free. For the establishing of linearity the Analytical Methods Committee suggests preparing a minimum of six concentration levels in duplicates [42]. Even though duplication at each concentration level is considered an optimal design strategy by this accredited organisation, it is a poor approach that should not be followed. A study of the behaviour of the uncertainty as a function of the number of replicates for the model described by Eq. (1) demonstrated that performing between four and six replicates at each experimental level decreases the uncertainty significantly along the experimental range and produces a uniform confidence prediction region around the centre of the calibration graph which is an important feature for quantification experiments [48]. Based on the reported behaviour of the uncertainty as a function of the replication and considering that the minimum number of concentration levels proposed by various guidelines and articles on analytical validation varies between five and six, it is reasonable to measure the linearity of a calibration function by preparing a minimum of five concentration levels in triplicates [31,43].

3.4.1.2. Error sum squares. After selecting a sensible number of concentration levels (I) and replicating every concentration level J-times in a particular calibration experiment, the summation of three squared differences, namely the residual error sum of squares (SS_r) pure experimental error sum of squares (SS_ε) and lack-of-fit error sum of squares (SS_{lof}), must be calculated according to the following equations:

$$SS_r = \sum_{i=1}^{I} \sum_{j=1}^{J_i} (y_{ij} - \hat{y}_i)^2 \quad (2)$$

$$SS_\varepsilon = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(y_{ij} - \bar{y}_i)^2 \quad (3)$$

$$SS_{lof} = SS_r - SS_\varepsilon = \sum_{i=1}^{I}(\bar{y}_i - \hat{y}_i)^2 \quad (4)$$

The term y_{ij} represents the experimental response, $\hat{y}_i$ is the estimated response obtained by using Eq. (1), and $\bar{y}_i$ is the average response at every concentration level.

3.4.1.3. Degrees of freedom. The degrees of freedom (*DF*) associated to Eqs. (2)–(4) are respectively:

$$DF_r = (IJ - 2) \quad (5)$$

$$DF_\varepsilon = (IJ - I) \quad (6)$$

$$DF_{lof} = (I - 2) \quad (7)$$

The bracketed number 2 in Eqs. (5) and (7) and associated with Eqs. (2) and (4) respectively, represents the number of parameters described by Eq. (1) (the ϕ slope + the δ intercept = 2 parameters). If a model with a different number of parameters to those described by Eq. (1) were studied, for instance:

$$\hat{y}_i = \phi x_i + \varphi x_i^2 + \delta \quad (8)$$

The degrees of freedom associated with Eqs. (2) and (4) would be $(IJ-3)$ and $(I-3)$ respectively. The bracketed number 3 in this case, represents the three parameters $(\phi + \varphi + \delta)$ of Eq. (8).

By using Eq. (1) and the minimum criteria of five concentration levels ($I = 5$) in triplicates ($J = 3$) established in the previous section, it is possible to estimate 13, 10 and 3 degrees of freedom for SS_r, SS_ε and SS_{lof} respectively.

3.4.1.4. Acceptability of linearity data. A ScienceDirect search was performed using the keywords "linearity test" and revealed that several articles used these two words as a true measure of linearity in chromatography method validation. It has been reported repeatedly that an analyst engaged in any analytical validation should be aware that there is no test for linearity as such [42,43,49–51]. The Analytical Method Committee suggests using the *F*-test as a reliable approach to check the linearity of any calibration function. The procedure is as follows:

- The purely experimental variance and lack-of-fit variance designated by σ_ε^2 and σ_{lof}^2 are estimated by computing the quotients $SS_\varepsilon/(IJ - I)$ and $SS_{lof}/(I - 2)$ respectively.
- The calculated σ_ε^2 and σ_{lof}^2 variance terms are used to calculate the Fisher variance ratio or *F*-test by the expression:

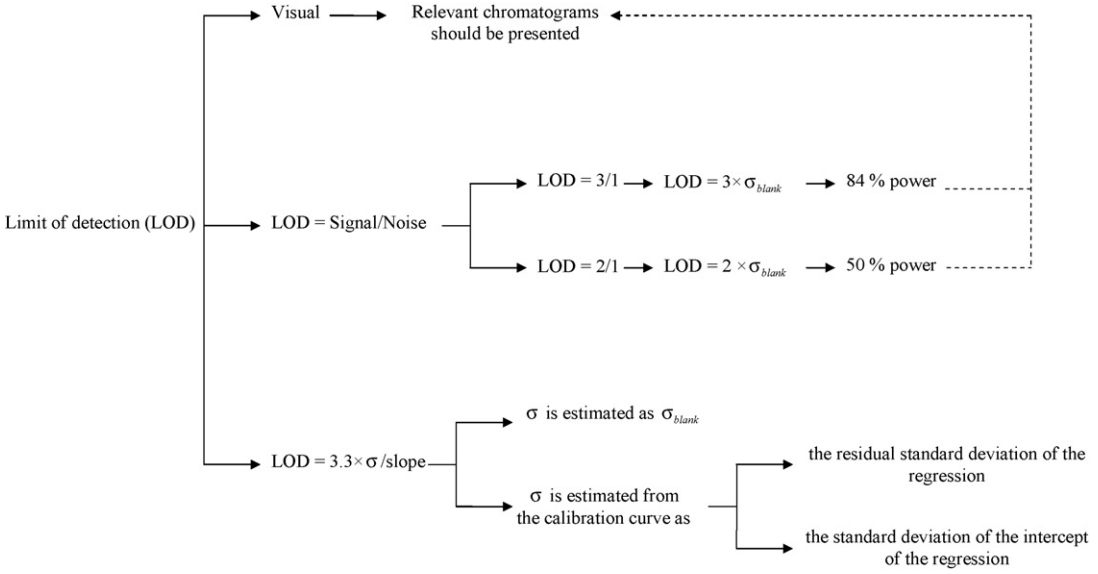
$$F_{(I-2)/(IJ-I)} = \frac{\sigma_{lof}^2}{\sigma_\varepsilon^2} \quad (9)$$

- The value of $F_{(I-2)/(IJ-I)}$ calculated experimentally is compared against the critical value of *F* found in statistical tables, generally at the 95% confidence level for $I - 2$ and $IJ - J$ degrees of freedom in the numerator and denominator respectively. If the experimental data set describes a genuine linear calibration of the form given by Eq. (1) then the condition $F_{tabulated} > F_{(I-2)/(IJ-I)}$ must be fulfilled. Otherwise there are grounds to suspect that a different model to the described by Eq. (1) must be proposed.

The estimation of the various error sum squares and the Fisher ratio for testing the acceptability of a linear model proposed in the literature is discussed in Appendix A.

3.5. Range

In general, the range of an analytical procedure can be defined as the interval between the upper and lower concentration of analyte for which suitable precision, accuracy and linearity have been demonstrated. The literature on method validation describes different ranges. For instance, linear range, analytical range, calibration range, dynamic range, working range. However, they can be summarised as working (or analytical) range and linear (or calibration, or dynamic) range. The former range which is wider than the latter, describes the interval between the lowest (limit of detection) and the highest concentration where the signal can be related to the concentration for the evaluation of random and systematic errors. The linear range corresponds to the valid interval of functional dependence of the signal on concentration or mass [52] which is usually determined by using the method of least squares, which



Scheme 1. Main approaches proposed in the literature for determining the limit of detection.

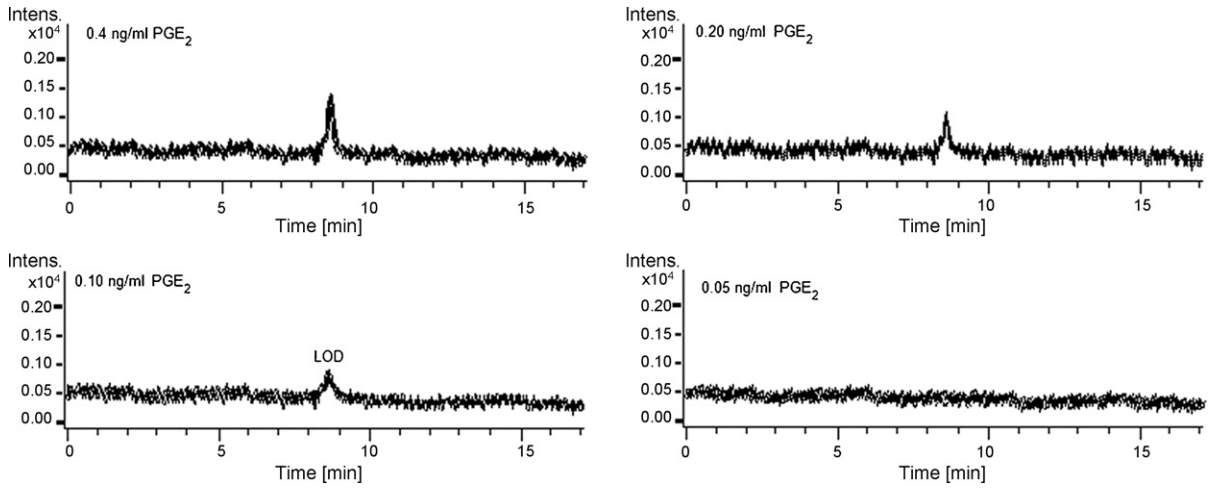


Fig. 3. Visual determination of the limit of detection for PGE_2 determined visually by diluting successively replicate solutions of PGE_2 at an initial concentration of 0.4 ng/ml.

in turns assumes homoscedasticy of the measurements over the linear range.

To demonstrate an acceptable linear range, it is generally suggested to prepare five different standard solutions from 50 to 150% of the target analytical concentration [43]. However, it is important to remember that the specified ranges often differ depending on the intended application of the procedure. For instance, the ICH recommends preparing five different standard solutions (plus a blank) from 80 to 120% for the assay of a drug or a finished product [31]. An analyst interested in validating a particular method should consult the guidelines to encompass the recommended minimum specified range of the intended method.

3.6. Limit of detection

The limit of detection (LOD) is commonly defined as the lowest amount of analyte in a sample that can be reliably detected but not necessarily quantitated by a particular analytical method. A decade ago, the International Organization for Standardization proposed to use the term "*minimum detectable net concentration*" [53] defined as the confidence with which it is possible to detect an amount of analyte in the sample larger than that in a blank sample with a statistical power of $(1-\beta)$.

Different criteria are used for evaluating the LOD. A summary of the main approaches proposed for determining the LOD in the literature on validation is presented in Scheme 1. Visual determination of the LOD is performed by preparing samples with known concentrations of the analyte and by establishing the level at which the analyte can be reliably detected. Fig. 3 shows the chromatography LOD for PGE_2 determined visually by diluting successively replicate solutions of PGE_2 at an initial concentration of 0.4 ng/ml. The replicate solutions are diluted up to a concentration level where the analyte is not longer detected visually. According to Fig. 3 the visual LOD corresponds to a concentration of 0.1 ng/ml PGE_2. The guidelines for validation of analytical procedures [31] recommend the presentation of relevant chromatograms when submitting a validation report where the LOD determination is based on a visual criteria. Another approach used for estimating the LOD is the calculation of the signal/noise relationship under the assumption that data normality, homoscedasticity and independency of residuals are met. An inspection of the literature on method validation revealed that this is the most popular approach among analysts performing validation studies. The signal to noise relationship is determined by comparing the analytical signals at known low concentrations with those of blank sample up to an analytical concentration that produces a signal equivalent to three or two times the standard deviation of the blank sample ($3\times\sigma_{blank}$ or $2\times\sigma_{blank}$).

The implications of using the previous criteria are illustrated in Fig. 4 where a limit of decision derived from the blank distribution is established at $+2\times\sigma_{blank}$. This limit of decision, at the 95% confidence level of the blank distribution is a probability that indicates whether or not a signal could be due to the sample ($>+2\times\sigma_{blank}$) or to the blank ($<+2\times\sigma_{blank}$). When $3\times\sigma_{blank}$ is used as a criterion for estimating the LOD (Fig. 4A), the probability β of obtaining a false negative is 16%. Conversely, it is possible to conclude that the concentration or amount of analyte in the analysed sample is larger than in a blank matrix with a statistical power $(1-\beta)$ of 84%. The statistical power measures the confidence with which it is possible to detect a particular difference if one exists [54]. There is not a conventional criterion to determine what is a suitable statistical power, however a value of 80% is generally considered the minimum desirable. When the criterion LOD = $2\times\sigma_{blank}$ is used (Fig. 4B) the probability β of reporting a false LOD ($\beta = 50\%$) is equal to the probability of finding a true LOD ($1-\beta = 50\%$). Although

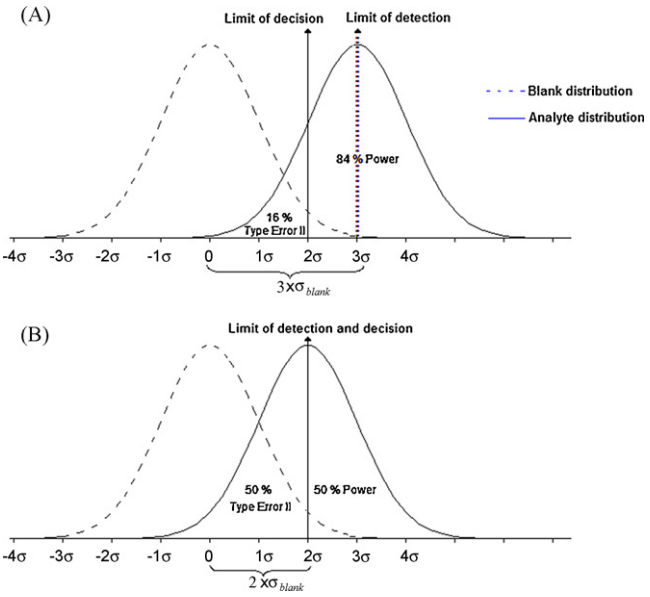


Fig. 4. Statistical power when (A) $3\times\sigma_{blank}$ and (B) $2\times\sigma_{blank}$ are used as criteria for estimating the limit of detection.

the minimum number of independent determinations required to establish the LOD is 10, it is advisable to increase the number of replicates when the LOD is defined as $2 \times \sigma_{blank}$ to avoid reaching wrong conclusions. Both LOD criteria ($3 \times \sigma_{blank}$ or $2 \times \sigma_{blank}$) should be justified by presenting the relevant chromatograms.

The final approach described in Scheme 1 is based on the quotient of two analytical parameters, namely the standard deviation (σ) and the slope of a regression curve (ϕ as in Eq. (1)). The former parameter could be expressed as the standard deviation of the blank (σ_{blank}), as the residual standard deviation of the calibration curve or as standard deviation of the intercept of the calibration curve. The guidelines on method validation do not express any particular preference for the approaches described in Scheme 1, however they recommend that when reporting the LOD the definition used in its evaluation should be stated.

3.7. Limit of quantitation

The limit of quantification (LOQ) is defined as the lowest concentration or amount of analyte that can be determined with an acceptable level of precision and accuracy [9]. Similarly to LOD, LOQ is evaluated by using different approaches, that is:

1. *Visual evaluation:* samples with known analytical concentration are prepared and the minimum level at which the analyte can be quantified with an acceptable level of uncertainty is established.
2. *Signal/noise ratio:* the signals of samples with known analytical concentrations are compared with those of blank samples up to an analytical concentration that produces a signal equivalent to 10 times the standard deviation of the blank sample ($10 \times \sigma_{blank}$).
3. *Standard-deviation/slope ratio (LOQ = 10 × σ/φ):* the parameters σ and ϕ are calculated in the same fashion as LOD.

The second approach, which is the most used in the literature, could be defined in a more general context as the lowest amount of analyte that can be reproducibly quantified above the LOD (LOQ = $n \times$ LOD).

Other definitions to express the LOQ different from those described above could be used, provided that the definition used in its evaluation is stated.

It is important to note that the discussed approaches for the evaluation of the LOQ do not demonstrate that at the LOQ concentration there is an adequate accuracy and precision. The different approaches proposed in the literature have been critically revised in an article which advocates using the accuracy profile approach in order to estimate an LOQ more in accordance with its contextual definition [55]. The reader interested in applying the accuracy profile approach is referred to the articles of Hubert et al. [56] and Boulanger et al. [57] who were the first to introduce this concept.

3.8. Ruggedness

This parameter evaluates the constancy of the results when external factors such as analyst, instruments, laboratories, reagents, days are varied deliberately. By considering these critical external factors and inspecting Table 1, it is evident that ruggedness is a measure of reproducibility of test results under normal, expected operational conditions from laboratory to laboratory and from analyst to analyst [58]. Ruggedness cannot be erroneously used as a synonymous of robustness as it is going to be explained in the next section.

3.9. Robustness

This parameter evaluates the constancy of the results when internal factors (no external factors as in ruggedness) such as flow rate, column temperature, injection volume, mobile phase composition or any other variable inherent to the method of analysis are varied deliberately. It is generally not considered in most validation guidelines [46].

Although robustness and ruggedness aim at testing the reproducibility of the test results regardless of internal or external influences respectively, the literature on method validation bears evidence that both terms are used interchangeably. The analyst performing a method validation should distinguish the similarities and differences between these validation parameters and avoid misconstruing ruggedness as robustness.

Classical and multivariate methodologies such as the one-factor-at-the time approach or a factorial design have been proposed to evaluate both ruggedness and robustness. However for more comprehensive studies on robustness and ruggedness evaluation the reader is referred to [59–62].

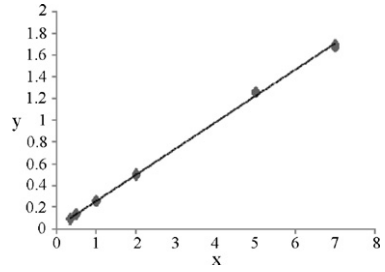
4. Final remarks

Method validation has evolved rapidly over the last half a century from being a neglected area of many scientific disciplines into a widely recognized process used to confirm that an analytical procedure employed for a specific test is appropriate for its intended use. However, despite this rapid evolution and the efforts of different regulatory bodies to reach greater harmonization of the relevant performance indicators commonly evaluated in method validation, some relentless inconsistencies in relation to their definitions, calculation and interpretation are repeatedly used. Practitioners should be vigilant over the herein described key aspects of method validation and bear in mind that their misapplication goes much further than a simple rejection of a submitted report or a waste of money, time and resources. The acceptance and application of erroneous terminology can have serious implications in results reliability, laboratory performance and institution credibility.

Appendix A. Demonstration that a correlation coefficient close to 1 is not a reliable indicator of linearity

The data printed in bold type have been reported elsewhere [45].

Comments	x	y			$\bar{y}$
The amount of analyte (ng), the chromatographic analyte/internal-standard peak area ratio and its average are designated as x, y and $\bar{y}$ respectively.	**0.35**	**0.0956**	**0.0948**	**0.0934**	**0.0946**
	0.50	**0.1356**	**0.1393**	**0.1361**	**0.1370**
	1.00	**0.2575**	**0.2551**	**0.2535**	**0.2554**
	2.00	**0.5028**	**0.4962**	**0.4940**	**0.4977**
	5.00	**1.2605**	**1.2534**	**1.2516**	**1.2552**
The calibration graph was obtained by plotting y vs x.	**7.00**	**1.6706**	**1.6950**	**1.6928**	**1.6861**
Proposed linear model by using the reported data.		$\hat{y} = 0.242x + 0.015$			
Reported squared correlation coefficient (r^2).		$\boldsymbol{r^2 = 0.9995}$			



Testing the acceptability of the above linear model by using the principles and equations described in Section 3.4.

x	$(y-\hat{y})^2$			$(y-\hat{y})^2$			$(y-\hat{y})^2$		
0.35	1.68×10^{-5}	2.40×10^{-5}	3.97×10^{-5}	1.00×10^{-6}	4.00×10^{-8}	1.44×10^{-6}	2.60×10^{-5}	2.60×10^{-5}	2.60×10^{-5}
0.50	1.60×10^{-7}	1.09×10^{-5}	1.00×10^{-8}	1.96×10^{-6}	5.29×10^{-6}	8.10×10^{-7}	1.00×10^{-6}	1.00×10^{-6}	1.00×10^{-6}
1.00	2.50×10^{-7}	3.61×10^{-6}	1.23×10^{-5}	4.55×10^{-6}	7.11×10^{-8}	3.48×10^{-6}	2.67×10^{-6}	2.67×10^{-6}	2.67×10^{-6}
2.00	1.44×10^{-5}	7.84×10^{-6}	2.50×10^{-5}	2.64×10^{-5}	2.15×10^{-6}	1.34×10^{-5}	1.78×10^{-6}	1.78×10^{-6}	1.78×10^{-6}
5.00	1.26×10^{-3}	8.07×10^{-4}	7.08×10^{-4}	2.84×10^{-5}	3.12×10^{-6}	1.27×10^{-5}	9.10×10^{-4}	9.10×10^{-4}	9.10×10^{-4}
7.00	1.47×10^{-3}	1.96×10^{-4}	2.62×10^{-4}	2.41×10^{-4}	7.86×10^{-5}	4.44×10^{-5}	5.23×10^{-4}	5.23×10^{-4}	5.23×10^{-4}
Residual error sum squares (Eq. (2))	$SS_r = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(y_{ij}-\hat{y}_i)^2$ (Eq. 2)	$SS_\varepsilon = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(y_{ij}-\bar{y}_i)^2$ (Eq. 3)	$SS_{lof} = \sum_{i=1}^{I}(\bar{y}_i-\hat{y}_i)^2$ (Eq. 4)						
Pure error sum squares (Eq. (3))									
Lack-of-fit error sum squares (Eq. (4))									
Results after applying Eqs. (2)–(4) →	$SS_r = 4.86 \times 10^{-3}$			$SS_\varepsilon = 4.69 \times 10^{-4}$			$SS_{lof} = 4.40 \times 10^{-3}$		
Degrees of freedom (DF)	$18-2=16$			$18-6=12$			$6-2=4$		
Associated variances (σ^2 = SS/DF)	$\sigma_r^2 = 3.04 \times 10^{-4}$			$\sigma_\varepsilon^2 = 3.91 \times 10^{-5}$			$\sigma_{lof}^2 = 1.10 \times 10^{-3}$		
Fisher ratio ($F = \sigma_{lof}^2/\sigma_\varepsilon^2$) calculated (if $F_{calculated} < F_{tabulated}$ then *Linear*)	28.151 (calculated) $\gg$ 3.259 (tabulated at the 95% with 4 and 12 degrees of freedom)								
Conclusions	Although $r^2 = 0.9995$ there is a significant non-linearity in the reported data. The proposed linear model $\hat{y} = 0.242x + 0.015$ is not adequate to describe the observed data due to some significant lack of fit.								

References

[1] International Organization for Standardization. Guide 8402: Quality Vocabulary. ISO, Geneva, 1994.
[2] W.J. Murphy, Anal. Chem. 19 (1947) 821.
[3] B.L. Clarke, Anal. Chem. 19 (1947) 943.
[4] F.C. Strong, Anal. Chem. 19 (1947) 968.
[5] A.L. Wilson, Talanta 17 (1970) 21.
[6] A.L. Wilson, Talanta 17 (1970) 31.
[7] A.L. Wilson, Talanta 20 (1973) 725.
[8] A.L. Wilson, Talanta 21 (1974) 1109.
[9] M.E. Swartz, I.S. Krull, Analytical Method Development and Validation, Marcel Dekker Inc., New York, 1997.
[10] http://fshn.ifas.ufl.edu/faculty/STTalcott/www/Food%20Analysis/Anthocyanin-SPE.doc.
[11] L. Huber, LC–GC Int. 11 (1998) 96, www.geocities.com/HotSprings/Spa/6896/methval.pdf.
[12] G.D. Boef, A. Hulanicki, Pure Appl. Chem. 55 (1983) 553.
[13] B. Persson, J. Vessman, Trends Anal. Chem. 17 (1998) 117.
[14] M. Valcárcel, A. Gómez-Hens, S. Rubio, Trends Anal. Chem. 20 (2001) 386.
[15] B. Persson, J. Vessman, Trends Anal. Chem. 20 (2001) 526.
[16] Western European Laboratory Accreditation Cooperation. Guidance Document No. WG D2. EURACHEM/WELAC Chemistry, Teddington, 1993.
[17] I. Taverniers, M. De Loose, E. Van, Bockstaele, Trends Anal. Chem. 23 (2004) 535.
[18] A.G. González, M.A. Herrador, Trends Anal. Chem. 26 (2007) 227.
[19] R.D. McDowall, Anal. Chim. Acta 391 (1999) 149.
[20] R. Bonfiglio, R.C. King, T.V. Olah, K. Merkle, Rapid Commun. Mass Spectrom. 13 (1999) 1175.
[21] P.J. Taylor, Clin. Biochem. 38 (2005) 328.
[22] D. Dadgar, P.E. Burnett, J. Pharm. Biomed. Anal. 14 (1995) 23.
[23] J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, Prentice-Hall, Harlow, England, 2000.
[24] International Organization for Standardization. Guide 5725-1: Accuracy (Trueness and Precision) of Measurement Methods and Results, Part 1: General Principles and Definitions. ISO, Geneva, 1994.
[25] W.W. Hauck, W. Koch, D. Abernethy, R.L. Williams, Pharmacol. Forum 34 (2008) 838.
[26] L. Ebdon, L. Pitts, R. Corneils, H. Crews, O.F.X. Donard, P. Quevauviller (Eds.), Trace Elements Speciation for Environment Food and Health, Royal Society of Chemistry, Cambridge, 2001.
[27] International Union of Pure Applied Chemistry. Harmonised Guidelines for In-House Validation of Methods of Analysis (Technical Report). IUPAC, Budapest, 1999. <http://old.iupac.org/divisions/V/501/draftoct19.pdf>.
[28] R. Boqué, A. Maroto, J. Riu, F.X. Rius, Grasas y Aceites 53 (2002) 128.
[29] P. Van Zoonen, R. Hoogerbrugge, S.M. Gort, H.J. Van de Wiel, H.A. Van't Klooster, Trends Anal. Chem. 18 (1999) 584.
[30] International Organization for Standardization/International Electrotechnical Commission. R103: General Requirements: Proficiency Testing for ISO/IEC 17025 Laboratories. ISO/IEC, 2008. <http://a2la.org/requirements/A2LA_General_Requirements_for_Proficiency_Testing.pdf>.
[31] International Conference on Harmonization. Guidance for Industry. Q2B: Validation of Analytical Procedures: Methodology. ICH, 1996. <http://www.fda.gov/Cder/Guidance/1320fnl.pdf>.
[32] US Food and Drug Administration. Guidance for Industry. Bioanalytical method Validation. FDA, 2001. <http://www.fda.gov/CDER/GUIDANCE/4252fnl.pdf>.
[33] International Organization for Standarization. DGuide 99999: International Vocabulary of basic and General Terms in Metrology (VIM). ISO, 2004. <http://www.ntmdt.ru/download/vim.pdf>.
[34] J.O. Westgard, R.N. Carey, S. Wold, Clin. Chem. 20 (1974) 825.
[35] J.O. Westgard, R.W. Burnett, Clin. Chem. 36 (1990) 1629.
[36] 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, AAPS J. 9 (2007) E30.
[37] R. Bischoff, G. Hopfgartner, H.T. Karnes, D.K. Lloyd, T.M. Phillips, D. Tsikas, G. Xu, J. Chromatogr. B 860 (2007) 1.

[38] International Organization for Standardization. Guide 5725-3: Accuracy (Trueness and Precision) of Measurement Methods and Results, Part 3: Intermediate Measures of the Precision of a Standard Measurement Method. ISO, Geneva, 1994.
[39] S.K. Kimothi, The Uncertainty of Measurements, Physical and Chemical Impact and Analysis, ASQ Quality Press Publication, Milwaukee, USA, 2002.
[40] Analytical Methods Committee, Analyst 113 (1988) 1469.
[41] M. Thompson, Anal. Proc. 27 (1990) 142.
[42] Analytical Methods Committee, Analyst 119 (1994) 2363.
[43] J.M. Green, Anal. Chem. 68 (1996) 305A.
[44] International Conference on Harmonization. Guideline for Industry. Q2A: Text on Validation of Analytical Procedures. ICH, 1995. <http://www.fda.gov/cder/Guidance/ichq2a.pdf>.
[45] S. Umeda, G.W. Stagliano, M.R. Borenstein, R.B. Raffa, J. Pharmacol. Toxicol. Methods 51 (2005) 73.
[46] C. Hartmann, J. Smeyers-Verbeke, D.L. Massart, R.D. McDowall, J. Pharm. Biomed. Anal. 17 (1998) 193.
[47] P. Araujo, T. Nguyen, L. Frøyland, J. Am. Soc. Mass Spectrom. 16 (2005) 388.
[48] P. Araujo, L. Frøyland, Accred. Qual. Assur. 10 (2005) 185.
[49] R. Cassidy, M. Janoski, LC–GC 10 (1992) 692.
[50] C.A. Dorschel, J.L. Ekmanis, J.E. Oberholtzer, F.V. Warren, B.A. Bidlingmeyer, Anal. Chem. 61 (1989) 951A.
[51] Analytical Methods Committee. AMC Technical Brief 3, 2000, 1. <http://www.rsc.org/images/brief3_tcm18-25924.pdf>.
[52] O. Mathias, Chemometrics: Statistics and Computer Application in Analytical Chemistry, Wiley–VCH, Weinheim, 1999.
[53] International Organization for Standarization. ISO 11843-1: Capability of Detection. Part 1: Terms and Definitions. ISO, Geneva, 1997.
[54] P. Araujo, L. Frøyland, J. Chromatogr. B 847 (2007) 305.
[55] E. Rozet, A. Ceccato, C. Hubert, E. Ziemons, R. Oprean, S. Rudaz, B. Boulanger, Ph. Hubert, J. Chromatogr. A 1158 (2007) 111.
[56] Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, Anal. Chim. Acta 391 (1999) 135.
[57] B. Boulanger, P. Chiap, W. Dewé, J. Crommen, Ph. Hubert, J. Pharm. Biomed. Anal. 32 (2003) 753.
[58] http://www.labcompliance.com/tutorial/methods/default.aspx.
[59] M. Mulholland, Trends Anal. Chem. 7 (1988) 383.
[60] E. Aguilera, R. Lucena, S. Cárdenas, M. Valcárcel, E. Trullols, I. Ruisánchez, Trends Anal. Chem. 25 (2006) 621.
[61] J.J.B. Nevado, M.J.V. Llerena, C.G. Cabanillas, V.R. Robledo, J. Chromatogr. A 1123 (2006) 130.
[62] B. Dejaegher, Y.V. Heyden, J. Chromatogr. A 1158 (2007) 138.