We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Analytical Calibrations: Schemes, Manuals, and Metrological Deliberations

Marwa S. El-Azazy

Additional information is available at the end of the chapter

Abstract

Chemical measurement processes (CMPs) must be performed in a setup of controlled statistical conditions. Thus, validation of such a measurement process and assessment of its ability to accurately measure the analyte is important. Analytical calibration is the most crucial step in any analytical procedure targeting the estimation of analyte concentration. As a key component of any validation procedure, calibration must be properly conducted. To achieve that, firm knowledge with the realms of the calibration process must exist. Several jurisdictions help to build up this acquaintance, including the terminology and definitions, the international guidelines and how they differ, schemes and manuals to be used to build a calibration model, metrological considerations, and assessment procedures. Careful thinking prior to any of the previous calibration aspects is necessary and helps to improve the product of the calibration process. Throughout this chapter, aspects of the calibration assembly will be thoroughly discussed. Different types of calibration will be revealed with a focus on analytical calibration for a CMP. Steps for a successful calibration will be described. The reader will be able to use information given throughout this chapter as a guide for an effective calibration process.

Keywords: analytical calibration, terminology, regulatory agencies, multi-, one- and two-standard calibrations, calibration methodologies

1. Introduction

Millions of analytical investigations are instigated every day. Despite the massive progress and advancements implemented to the developed techniques and instrumentations, calibration stays as the most critical stage in every analytical practice leading to the estimation of the target analyte.

An analytical measurement process is a setup with a demarcated configuration that has been carried to be statistically controlled under the designated experimental conditions. To substantiate the efficacy of analytical processes and subsequently the applicability in routine analysis, the ability of such a method to "quantify" must be assessed. Thus, and to fetch such a status of statistical management, key elements including *validation*, and hence its metrological frontier, *calibration*, must be clearly comprehended [1].

In the latest definition released by the Joint Committee for Guides in Metrology (JCGM) in their 3rd edition of the "International Vocabulary of Metrology, VIM", calibration is: "operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication" [2, 3]. Validation, in the same edition, was defined as "verification, where the specified requirements are adequate for an intended use."

Though *validation* as an idiom is already well-known, the protocols of its application are not clear for many of the analytical chemists. No need to say that validation of an already developed analytical process must be performed following a clearly written protocol and through a series of laboratory experiments. Moreover, different regulatory bodies (e.g., IUPAC, ICHQ2R1) do have different nomenclature for such a term (as well as its components) and hence dissimilar manuals, an issue that in turn leads to different performance and approval criterions [4, 5].

As a component of the validation process, calibration is also a subject of controversy in terms of vocabularies, the perception of the calibration procedure starting with method development to fitting of results obtained, implementation of the appropriate linearity testing, and hence the assessment of goodness-of-fit and deviation from linearity.

It is very important to recognize that though the existence of intrinsic discrepancies between chemical (CMPs) and physical (PMPs) measurement processes in terms of uncertainty associated with the results and the availability of reference materials; both are still treated with the same metrological approach. Yet, an imperative difference between both processes must be carefully considered which is *calibration* [6–8].

The purpose of this treatise is to shed light on the "appropriate" definition of calibration as a process that encompasses metrological/statistical as well as procedural evaluation of the analytical measurement. The different types of calibration will be revealed. Analytical calibration, across the different guidelines and with respect to definitions and terminologies, schemes, metrologies, and methodologies will be discussed.

Though in some sections of this piece complicated terminologies would be used, a reader of this chapter, even if not from the scientific community, would be able to understand information given with the help of definitions revealed in almost every section.

2. Calibration in analytical sciences: fundamentals

Several definitions exist in literature for calibration. In addition to the previously mentioned definition given by the VIM [2, 3], IUPAC definition of calibration can be viewed as a "general" description where it is given as "an operation that relates an output quantity to an input quantity" [9]. Unfortunately, these definitions instead of giving a clear-cut understanding of the term and the corresponding process have created a kind of confusion where it is common to find the wrong term being given to the wrong process or similar names given to different types of processes, etc.

However, it is noteworthy to mention that the additional "notes" given by the JCGM [3] on the definition of calibration would clarify this misunderstanding where: "A calibration may be expressed by a statement, calibration function, calibration diagram, calibration curve, or calibration table. In some cases, it may consist of an additive or multiplicative correction of the indication with associated measurement uncertainty" and "Calibration should not be confused with adjustment of a measuring system, often mistakenly called 'self-calibration', nor with verification of calibration." Furthermore, and according to JCGM, "Often, the first step alone in the above definition is perceived as being calibration."

Yet, and as per these definitions, it is important to distinguish between the different types of calibration and whether it is designed for a qualitative or a quantitative purpose. As a relation between an input quantity and another output quantity, quantitative calibration can be performed *directly* (where the measurement and the reference values are being compared
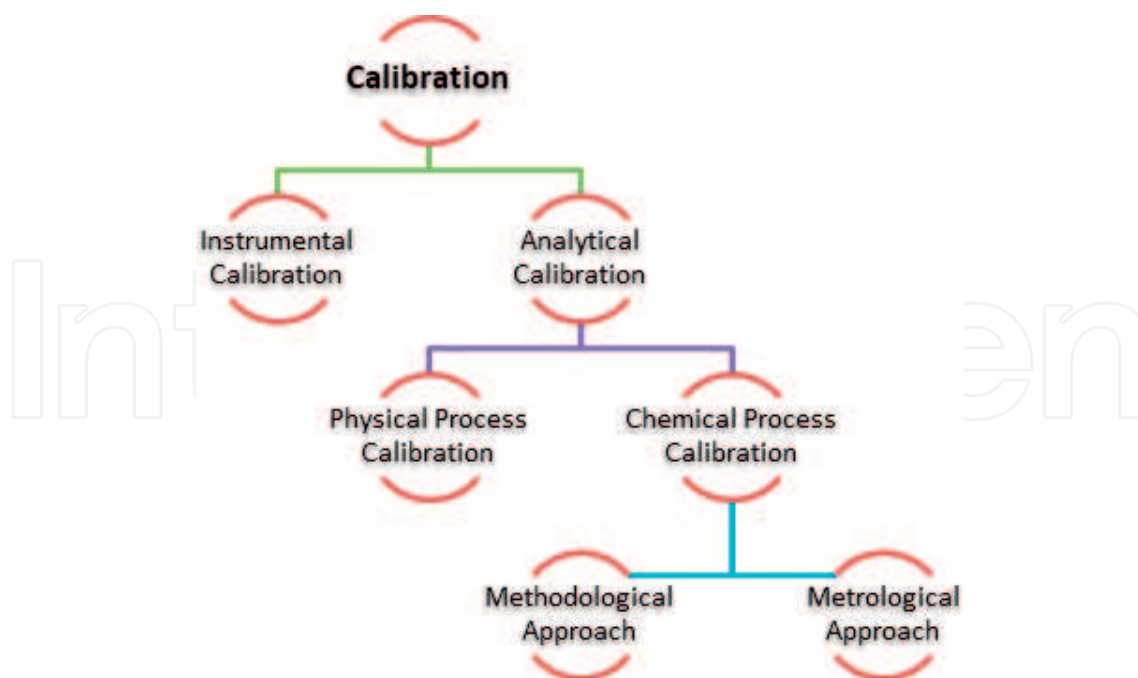

Figure 1. Schematic representation of the different approaches for calibration.

employing the same units) or *indirectly* (where the measured response is being decoded into the corresponding quantity to be determined, i.e., *analytical calibration*). Both *direct* and *indirect* calibrations can target the equipment as well as the process itself [10]. More details on these subdivisions will be given under the relevant section.

Calibration then can be tackled using different standpoints depending on its implication. In other words, is the calibration targeting the system of measurement and its quality so it is *metrological calibration* or it is an *analytical calibration* that merely describes the relation between the analyte and the corresponding response? Distinction of direct from indirect calibration and then process and instrument can be performed using the metrological maneuver. Another approach to see the calibration process would be in terms of methodologies and schemes followed to achieve such a status. **Figure 1** shows a schematic representation of the calibration process with the different approaches commonly found in literature. The following subsections will be dealing mainly with analytical calibration of a chemical measurement processes in terms of *steps and guidelines, schemes, manuals and methodologies,* and *metrological considerations*.

3. Analytical calibration

3.1. Steps and guidelines

As previously mentioned, the term *analytical calibration* is used when the calibration process cannot be performed *directly*. In general, the objective of doing calibration is to establish an experiential liaison between the instrument response signal "*y*-variable" and the reaction factors "*x*-variable." The purpose of establishing such a liaison is to be able to assess the influence of these variables on the response and hence quantify the analyte.

Surveying the literature shows that different validation strategies proposed by the different regulatory institutions usually involve quite different guidelines for analytical calibration. In addition to the differences in terminologies used to define analytical calibration and hence associated terms, other major differences can be found as follows.

3.1.1. Proposing a strategy for a calibration study

Planning is the preliminary step in conducting the calibration study. The conventional scheme for performing calibration is to prepare a set of standards (plus a blank) followed by quantifying the response signal for such a set [11–13]. Common several "How" questions usually evolve as the analyst is getting ready to conduct this study:

- How many standards will be used?

- How the target of calibration will affect the composition of calibration standards?

- How the selected number of standards will be patterned and disseminated on the studied concentration range?

- How to select the concentrations that will be measured?

- How the measurement procedure would be like?

- How many times the analysis should be repeated (replications)?

- How the calibration mode will be set? (details will be discussed later)

The elements of *calibration hierarchy* according to JCGM [3] *are one or more measurement standards and measuring systems operated according to the measurement procedure*. Typically, a minimum of 5–6 calibration standards is used for this purpose. Yet, the number of standards used might vary according to the performed analytical process as well as the guidelines proposed by the supervisory body followed. The calibration standard might be *matrix-free* if the purpose is to calibrate solvents, for example, or *matrix-matched* (*MMC*) if it is expected that the presence of matrix would affect the response signal and hence the calibration outcomes. In this case, a blank sample (analyte free) should be used.

Careful distribution of the selected concentration levels over the working range is necessary for appropriate calibration. In this concern, discrimination between narrow and wide calibration ranges is essential. Attention should be paid for the case where a wide concentration range is calibrated where keeping the selected levels at very wide distances, a common approach in literature, might deteriorate the detecting system of the instrument, an issue that produces erroneous readings. The best approach is to keep the data points consistently dispersed across the selected range. Moreover, selected concentrations should be independently prepared (no serial dilutions) to avoid augmentation of error.

Selection of the concentration range to be covered should be based on the expected content of the real samples taking in consideration the matrix and the intended application of the proposed procedure [14]. According to ICH guidelines, for example, if the calibration is performed on an active ingredient or a final product, the range is usually 80–120% of the analyte concentration [4]. In case of using MMC, the blank sample (zero concentration, solvent) should be considered.

The appropriate protocol for a measurement will be the one that simulates the actual circumstances. In this itinerary, it is recommended that calibration samples are to be unevenly analyzed instead of being measured in an increasing concentration sequence. Moreover, insertion of calibration standards randomly in between the unknown samples within the measurement stream is commended.

Every experiment is associated with an error! Diminishing the random error (measurement uncertainty) and hence improving the precision is usually one of the goals when analytical calibration is performed. Replicate analysis is usually the approach. The number of recommended replicates differs according to the implemented guideline. While EMA, FDA, and AOAC indorse five replicates, ICH recommends three replicates or six replicates at a single concentration level compared by replication for 2–3 times at 6–10 concentrations evenly spaced across the linear range by Eurachem [4, 5, 14–17]. However, and due to economic considerations, triplicate analysis is the common approach.

Some guidelines impose more regulations than those previously mentioned. For example, FDA for bioanalytical method validation [18] necessitates that at least four concentrations (lower limit of quantification LLOQ, low, medium, and high) measured in six runs in duplicate/run.

3.1.2. Assembling and modeling of experimental data

Following the fulfillment of the previous checklist of "How questions," the subsequent movement will be to corroborate the liaison between the measured concentration and the equipment response. This liaison is usually established via *regression analysis* and hence *calibration graphs (commonly described as curves)*. According to JCGM [3], *calibration curve* is "expression of the relation between indication and corresponding measured quantity value", and "a calibration curve expresses a one-to-one relation that does not supply a measurement result as it bears no information about the measurement uncertainty."

3.1.2.1. Construction of calibration curve

The calibration curve is generally constructed by plotting the response values (*y*-axis, *dependent variable*) against the known standard concentration values (on *x*-axis, *independent variable*, *predictor*) either manually or by operating popular software like Excel®, for example. Performing regression analysis and drawing a regression line require a cautious decision on a bundle of three main components: *model, mode, and fitting technique*.

Typically, the number of predictors and so the type of response variable differ between various measurements. Accordingly, the regression pattern would be different. A common regression model is the *linear regression* where a *best-fit* straight line is drawn between *x* and *y* variables. Other types of regressions include *logistic, polynomial, stepwise (forward selection and backward removal), and ridge* regression.

In the simple linear regression, one independent variable is involved compared to more than one in case of multiple linear regression. The best-fit line is usually obtained employing the method of least squares (the most popular technique). This regression line is usually presented by the equation: $y = ax + b$, where *a* and *b* are the slope and the intercept, respectively. In this method, the line is calculated by minimizing the sum of squares of the residuals for each data point.

Regression analysis based on *principle component analysis* (*PCA*) is known as *principle component regression* (*PCR*), in which the response is regressed against a set of variables and using the PCA to find the regression coefficients. Other regression methods such as *partial least-squares (PLS)* establish a linear regression model by protruding *x* and *y* variables to a new space. This technique is mainly used when the number of data points is less than the number of variables [19, 20].

The last step after deciding upon the method and the model used is the selection of fitting technique. Adopting the case of a linear regression model being generated using the method of least squares, two approaches are commonly followed to find the best-fit line: *ordinary*

(*linear*) *least squares* (*OLS*) and *weighted least squares (WLS)* [21, 22]. As the name implies, OLS is the least-squares regression approach used when errors have a constant variance across the working range, *homoscedasticity*. That is of course in addition to the general assumptions of the OLS; errors are not correlated, conditional mean of errors is zero, and regressors are not linearly dependent (no multicollinearity). In contrary, WLS should be only used when variances are different, *heteroscedasticity,* and the working range is wide.

As an example of how to construct a calibration graph, spectrophotometric determination of tioconazole (antifungal, electron donor) using chloranilic acid (electron acceptor) via charge transfer reaction, and other calculated parameters needed to establish the regression relationship between [drug] and absorbance are shown in **Table 1**. Equations used to calculate essential regression parameters, *r (correlation coefficient)* and hence the *coefficient of determination* (R^2), *slope (a)* and *intercept* (*b*), are shown in **Figure 2**, which is the calibration graph plotted from data shown in **Table 1**.

3.1.2.2. Assessment of performance: model metrics

Evaluation of a linear relationship between concentration and response is usually performed by assessing the regression statistics, calibration graphs, and residual plots of the proposed model. Inspection of linearity is usually made *visually* by observing the calibration plot. Again, different guidelines do use different terminologies to describe the linearity and range, FDA for example uses the term calibration (standard) curve, compared to ICH guidelines which clearly defines linearity and Eurachem which uses the term working range [4, 14, 16, 18].

x_i	y_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \hat{y}$	$(y_i - \hat{y})^2$	$(x_i - \bar{x})(y_i - \hat{y})$
40	0.073	−180	32400	−0.4537	0.205844	81.666
80	0.175	−140	19600	−0.3517	0.123693	49.238
120	0.276	−100	10000	−0.2507	0.06285	25.07
160	0.372	−60	3600	−0.1547	0.023932	9.282
200	0.473	−20	400	−0.0537	0.002884	1.074
240	0.582	20	400	0.0553	0.003058	1.106
280	0.677	60	3600	0.1503	0.02259	9.018
320	0.781	100	10000	0.2543	0.064668	25.43
360	0.878	140	19600	0.3513	0.123412	49.182
400	0.98	180	32400	0.4533	0.205481	81.594
Σ =						
2200	5.267	0	132000	0	0.838412	332.66
$\bar{x}$ = 220	$\hat{y}$ = 0.5267					

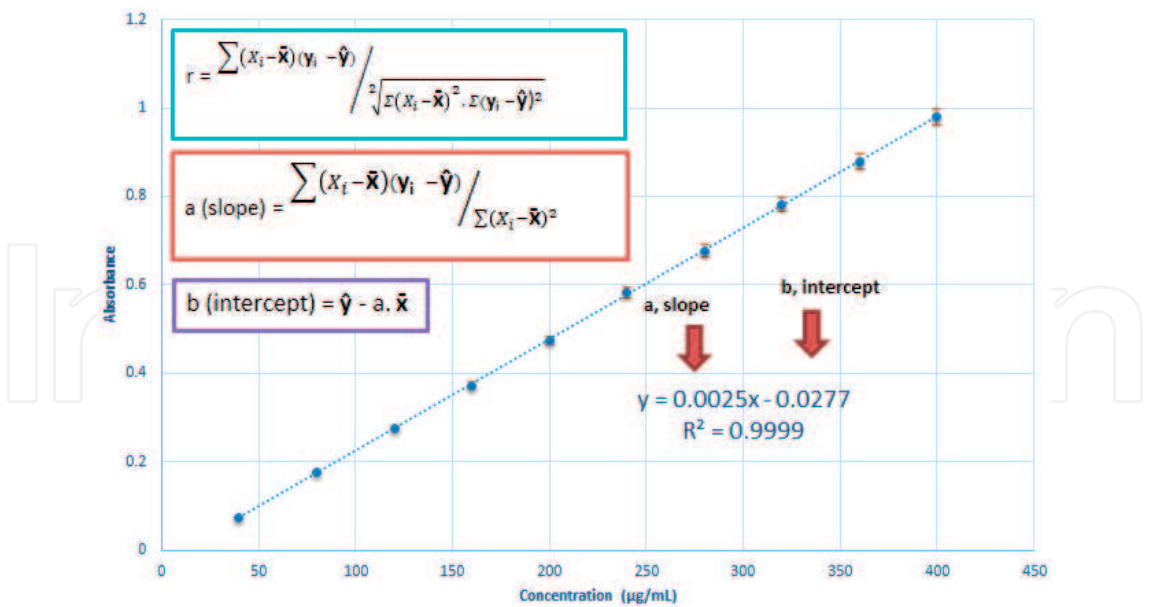**Table 1.** Parameters needed for the calibration graph (**Figure 1**).

Figure 2. Calibration graph plotted from data presented in **Table 1**.

Figure 3 shows three commonly used terms to describe the range: *analytical (dynamic range)*, *working (calibration) range*, and *linear range*.

The analytical or the dynamic range is the range in which the equipment is showing a response to the tested concentration, and this response is changing as the concentration varies. This relationship might be linear or nonlinear. The calibration range, in which the liaison between response and analyte concentration has an adequate uncertainty, usually starts with the limit
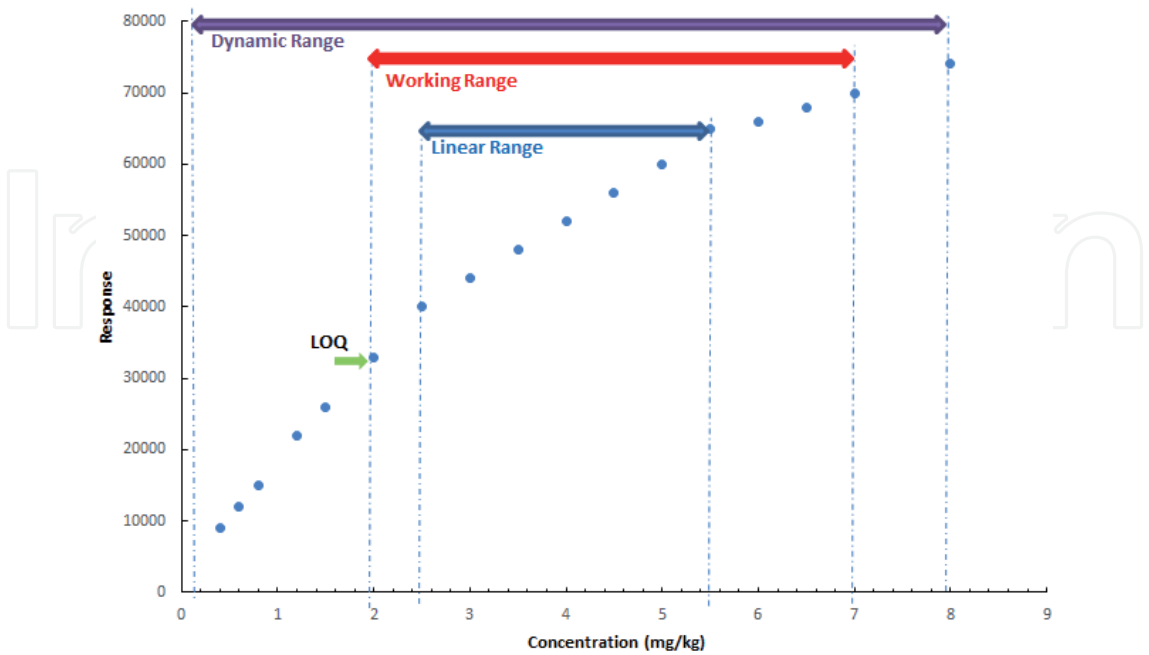

Figure 3. A demonstration of different ranges: analytical, working, and linear ranges.

of quantitation (LOQ) and ends where there is an obvious deviation from linearity. Working range is usually wider compared to the linear range. Thus, the latter can be defined as the range where there is a direct proportionality between concentration and response [14, 23, 24].

Though not being a component of the validation process, sensitivity is mentioned in variety of guidelines with the purpose of method evaluation. As a parameter, sensitivity can be easily estimated from the linear calibration graph as the function gradient. As per FDA guidelines [16], sensitivity is defined as "the lowest analyte concentration that can be measured with acceptable accuracy and precision (i.e., LLOQ)." In this itinerary, parameters such as *limit of detection* (*LoD*) and *limit of quantitation* (*LoQ*) need to be distinguished [23].

Once the status of "linearity" is established, statistical analysis is needed. Model metrics such as the correlation coefficient, slope of the regression line, and the intercept should be included (**Figure 2**). A comparison between the linearity assessment practices as per the different guidelines will be revealed in the following subsections. **Table 2** shows a comparison

Assessment approach	Recommended by	Pros	Cons	Ref.
Graphical inspection				
Residuals plot	IUPAC, NATA, INAB	Helpful together with the visual inspection in detecting linearity	Not a powerful tool in confirming linearity and needs a former experience with the different residual patterns	[14, 25–27]
Visual inspection (nongraphical)	–	Easy and useful in clear-cut situations	Subjective and cannot be used alone to indicate linearity	[16, 18]
Statistical analysis				
Analysis of variance (ANOVA)	IUPAC	$F_{calculated}$ value can be easily calculated	Not decisive	[9]
Lack-of-Fit (LOF)	IUPAC, INAB	Easy to be implemented in many software spreadsheets	Greatly dependent on the method precision, and usually several replicates are needed	[25, 27]
Mandel's fitting test	IUPAC	Easy to calculate and is mainly used when variances of two calibration functions are similar	Needs more samples compared to regular fitting tests and needs an estimation of the nonlinear model	[9]
Numerical evaluation				
Coefficients of correlation (r) and determination (R^2)	ICH, Eurachem, IUPAC, INAB, NATA	Widely used and implemented in almost all software	Sometimes deceptive and is monotonously getting higher as the number of variables increases	[4, 14, 25–27]
Residual standard deviation (RESSD)	NATA	Easy to understand and calculate	Depends on the measurement tool and different from one equipment to another	[26]

Table 2. A comparison between different linearity assessment approaches.

between the nongraphical, graphical, statistical, and numerical evaluation approaches for linearity evaluation. Contrast is shown in terms of the pros and cons of each approach as well as the recommending guideline(s).

i. *Graphical inspection:* this approach is recommended by most of validation guidelines. The preliminary step is to construct a plot between concentration and response on the *x* and *y* axes, respectively. The second step involves examining the plot visually. Majority of guidelines support using the plot of residuals as a tool to inspect linearity. Residuals can be defined as the difference between an observed value for a dependent measurement (*y*) and the estimated value of this measurement. As an approach, plot of residuals is a plot where calculated residuals are shown on the *y*-axis and the independent variable is shown on the *x*-axis. Linearity is confirmed when points are randomly scattered around the horizontal *x*-axis. Some data are not suitable candidates for plotting residuals; e.g., heteroscedastic data and outliers [14, 25–28].

ii. *Nongraphical approach*: visual inspection of data without plotting the graph or using statistical tools cannot be used as a linearity assessment tool by itself [16–18].

iii. *Statistical assessment:* statistical evaluation of data is a vital tool to confirm linearity when visual and residual plots cannot confirm a status of linearity. Generally, tests of significance are the methods used to infer whether stated claims about a sample of data extracted from a certain population are in favor or against the stipulated evidence. In other words, the significance tests are testing whether the *null hypothesis* (H_0) is being verified or not. Examples for significance tests include the student *t*-test and the *F*-test. Significance tests reported in literature to test linearity can be summarized as follows:

- *Analysis of variance* (*ANOVA*): this test depends on calculation of combined variances (S^2) between or within a group of data replicates assembled together in a certain way. This test is only recommended by IUPAC [9]. As a significance test, $F_{calculated}$ is compared with $F_{tabulated}$. The calculated *F*-values is found using the following formula: $F_{calculated} = (S_{y/x}/S_y)^2$, $S_{y/x}$ is the standard error for residuals and S_y is the pure error.

- *Lack-of-fit (LOF) test:* this test is a part of IUPAC validation guidelines [25, 27]. The calculated *F*-value is a ratio of mean sum of squares of random error (MSS_{error}) as a measure for divergence of points from the regression line being caused by the haphazard distribution of the points following replicate measurements to the mean sum of squares due to the lack-of-fit (MSS_{LOF}) as a measure for deviation of points caused by incongruity of the calibration paradigm. A comparison between the calculated and the tabulated value is then performed. Another approach to perform LOF test is to find the probability, *p*-value. Having a *p*-value higher than 0.05 means that the lack of fit is not significant [29, 30].

- *Mandel's fitting test*: this test is used to compare between two models (one is linear and the other is nonlinear) in terms of linearity when the variances are similar. The first step is to calculate the residual standard deviation for both models [9]. Again, if $F_{calculated}$ is greater than $F_{tabulated}$, the linear model cannot be accepted.

iv. *Numerical assessment*: numerical fitting parameters are used as a measure of goodness-of-fit (GOF) in regression analysis. The following parameters are commonly used:

- *Correlation coefficient (r)* and *coefficient of determination (R^2)*: these two parameters are commonly used to express the GOF of a model. In general, R^2 is now more applicable compared to r, where the former measures the proportion of variance of the dependent variable being diminished by prediction of the independent variable, while the latter is just a measure for the correlation between the two variables. In general, a value of r/R^2 close to 1 is an indication for linearity [31].

- *Residual standard deviation (RESSD)*: the smaller the value of RESSD, the better the obtained fit. RESSD measures the digression of data away from a fitted regression line.

3.2. Schemes

As previously mentioned under steps and guidelines for a successful calibration, the first step is to decide on how many standards will be used for calibration? Usually, the most common approach is the use of more than one standard "multi-standard calibration." It is noteworthy to mention that the term standard can be described as "realization of the definition of a given quantity, with stated quantity value and associated measurement uncertainty, used as a reference" and in *NOTE 1A* "realization of the definition of a given quantity can be provided by a measuring system, a material measure, or a reference material" and in *NOTE 9* "The term 'measurement standard' is sometimes used to denote other metrological tools, e.g., 'software measurement standard'" [3]. Another term is usually used then to describe the measurement standard, which is *reference materials (RM).*

As per JCGM [3], RM is "material, sufficiently homogeneous and stable with reference to specified properties, which has been established to be fit for its intended use in measurement or in examination of nominal properties." The composition of RM would vary depending on the application. For example, substance RM has an individual pure component in solvent of use, compared to matrix RM, which consists of analytes prepared in a matched matrix. When RM is "accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceabilities, using valid procedures", it will be known as certified RM, CRM [3].

Several schemes are usually available to perform calibration depending on the number of used standards.

3.2.1. Multi-standard calibration

This is the most popular approach for calibration where a minimum of three standards is usually used. Different guidelines do have different specifications in this concern and in terms of replicates and the measurement levels (please see Section 3.1.1).

3.2.2. Two-standard calibration

This approach is usually used for investigations performed at a narrow concentration range and after the linearity of the employed function has been confirmed, probably as a continued calibration. It can be also used when the applied procedure has a background. As a condition, the analyte concentration needs to be within the range covered by the two standards.

The real [analyte] can be calculated using the formula: $[anal] = [std_{1 \text{ or } 2}] + k (y_{unknown} - y_{std \text{ 1or2}})$, where the brackets express concentrations, k is the reciprocal of slope (sensitivity), and y is the response for the unknown and the standard, respectively [11, 12, 32]. Examples for this calibration are the pH meter and temperature sensor calibrations. A special scheme of a two-point calibration is known as bracketing calibration. In this approach, the [anal] is bracketed between the two standards assuming that a linear arithmetical interpolation can be proposed based on the knowledge of $[std_{1 \text{ and } 2}]$. The uncertainty associated with this approach is thus small if compared to the overall uncertainty [33, 34].

3.2.3. Single-standard calibration

As a direct calibration technique, this approach is applicable only when the calibration function linearity is established (especially in the region covering the [anal] and between the selected [std] and the origin) and if the graph intercept is zero [11, 12]. In this case, [anal] can be calculated using the calibration factor CF (which is the ratio between [std] and the average analytical response for the standard), where the unknown $[anal] = CF*y_{unknown}$. This simple calibration is generally used to test the drift from the response.

Multi-standard application then seems to be the most feasible and accurate scheme for calibration. However, this is not the case when, for example, the detector response varies with the time. In this case, the one-standard calibration is advantageous assuming that the unknown signal is within ±10–50% of the standard signal depending on whether the maximum analyte concentration limit has been surpassed or not [11, 12]. Depending on the analyte, availability of the standard, nature of the process, presence of concomitant analytes/interferences, and matrix effect, the procedure of calibration significantly varies and any of the previously reported schemes can be chosen.

3.3. Methodologies and manuals

While external and internal calibrations are the major themes, standard addition method (AC) and matrix-matched calibration (MMC) are also employed when required. Therefore, different methodologies for calibration can be proposed depending on how the RM will be applied within the course of calibration process. Through this section, emphasis will be basically on the CMPs, and the common methodologies usually followed to calibrate such a process.

3.3.1. External calibration (EC)

This approach is commonly known as "solvent/ standard calibration." As the name implies, EC is performed externally applied, i.e., the known standard solution, which is a substance RM prepared in the working solvent, is prepared and then analyzed distinctly from target samples. This approach can be applied using any of the previously mentioned schemes for calibration. The analysis protocol involves comparing the response for the unknown sample to the response for the target in the standard solution. One of the drawbacks of this methodology is the postulation that the impact of the difference between the matrices (standard and sample) can be ignored, an issue that leads to incorporation and propagation of a matrix systematic error. Nonetheless, this approach can be used when there is a minor or no contribution from the matrix effect and the instrumental drift can be ignored [11].

3.3.2. Matrix-matched calibration (MMC)

In contrast to the EC, MMC is used when the matrix has an impact on the response to the analyte. Both matrix RMs or substance RMs (together with an analyte-free matrix) can be employed for this approach. Attention should be paid that the matrix should be carefully matched. Again, the presence of analytes other than the target in the matrix could produce a matrix effect [11, 35].

3.3.3. Standard addition calibration (AC)

In this approach, known amounts of the analyte are added to aliquots of the test solution. Measurement is then performed by extrapolation of the calibration line to the zero response (no analyte). This approach can explain only certain types of the matrix effect; however; it cannot account for the effect of instrumental drift. Before the implementation of this method, the linearity of the calibration line should be confirmed over the whole concentration range. Moreover, the added concentration should be at least five times as high as the [anal] but within the linearity limits.

The actual [anal] is calculated using the equation: [anal] = CF $((y_{unknown}/y_{spiked} - y_{unknown}))$, where y_{spiked} and $y_{unknown}$ are the responses for the spiked and the unknown sample, respectively [11, 36].

3.3.4. Internal standard calibration (IC)

This approach is used to correct for both matrix effect and the drift over time. This technique is not the opposite to the previously mentioned EC; however, they can be used together. The matrix RM or as commonly known, internal standard (IS), which is structurally analogous to the analyte, is added for both unknown samples as well as the standards. The IS is selected in such a way that it can be distinguishably measured from the analyte. Moreover, there should be no interference between the IS and the analyte from one hand, and between the IS and the matrix of the unknown from the other hand. In addition to saving time and effort, the presence of the IS serves to compensate for sample loss during the preparation process [11]. The only limitation of this procedure is the availability of the ideal IS that can satisfy the previous conditions and emulate the matrix effect and the instrumental drift.

3.4. Metrological considerations

The product of the calibration scheme is usually portrayed as a mathematical model after performing the appropriate regression. Assessment of the proposed model depends on estimation of the experimental error which in turn affects linearity. Moreover, an important feature of the validation process which can be viewed as a direct calibration is the recovery studies. The concentration in the coming subsections will be on the metrological features of calibration in terms of error associated with the measurement and the recovery studies.

3.4.1. Uncertainty

As previously mentioned, the product of calibration is an experiential formula that relates the instrumental response to the analyte concentration. Thus, in other words, the actual value of

a measurement is equated with the experimental value. As a result, the uncertainty associated with the measurement needs to be determined. Principally talking about the systematic error of a measurement, and according to JCGM [3], it can be defined as "component of measurement error that in replicate measurements remains constant or varies in a predictable manner." NOTE 1 "A reference quantity value for a systematic measurement error is a true quantity value, or a measured quantity value of a measurement standard of negligible measurement uncertainty, or a conventional quantity value." As per NOTE 2 "Systematic measurement error, and its causes, can be known or unknown. A correction can be applied to compensate for a known systematic measurement error" and NOTE 3 "Systematic measurement error equals measurement error minus random measurement error."

For a linear calibration function generated from a multi-standard calibration approach using any of the methodologies of EC or IC, the linear regression line can be described by the equation: $y = ax + b$. This straight-line equation can be used to find an unknown concentration assuming that the response for this concentration is known. As the location of the regression line varies with the uncertainties associated with the regression parameters, a and b, the predicted concentration of the unknown would also be associated to uncertainty [37]. Metrologically, uncertainty of calibration is estimated using the following formula:

$$u(x_0) = \frac{S_{y/x}}{b} \sqrt{\frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_0 - \bar{x})} + \frac{1}{m} + \frac{1}{n}} \tag{1}$$

where $u(x_0)$ is the uncertainty associated with the unknown measurement, $S_{y/x}$ is the residual standard deviation, m is the number of replicates, n is the number of calibration points, $\bar{x}$ is the mean of x data points. It is noteworthy to mention that uncertainty associated with a measurement would be also sourced from the random error.

The *accuracy* and *trueness* are the terms used by majority of guidelines [1–5, 15–18]. However, there is a metrological difference between both terms. The term accuracy expresses how close an individual measurement to the real value of this measurement; however, trueness measures how close the mean of large number of values to the true value [3]. Thus, method trueness is measured as *absolute bias* or *relative bias*, which is expressed as % error and % relative error (%RE), respectively. Random error, however, affects the precision, which is calculated from the formula of standard deviation and in turn it affects the method accuracy [38]. Thus, uncertainty is affected by both bias as well as standard deviation.

3.4.2. Recovery

Generally, recovery investigations performed within the course of validation and following the calibration process can be treated as direct calibration of the proposed methods. Simply, the recovery is equal to = [found]/[actual]. It is important to declare that recovery outcome would differ per data point investigated and that the recovery value obtained at a certain value cannot be extrapolated to find the recovery at another data point.

For a linear function, the relation between recovered and actual analyte can be given as: [actual] = a [found]+ b, where a and b are the slope and the intercept, or the proportional and the additive errors, respectively [39].

4. Conclusion

Thousands of analyses and so validations are being performed every day. Calibration is a fundamental module of any analytical validation procedure. Different regulatory bodies propose different idioms and hence procedures for putting calibration in effect. Existence of a well-defined terminology for calibration and therefore a harmonized procedure would significantly improve the outcome of the analytical measurement. Appropriate selection of the calibration scheme and the subsequent methodology are the key factors for the success of analytical calibration. This chapter has outlined the process of analytical calibration in terms of appropriate designation (and considering the different releases by different documentary agencies), schemes (multi-, one-, and two-standard calibrations), and the operating manuals. Moreover, the metrological aspects of the calibration process have been revealed throughout the discussion with a focus on the recovery and uncertainties associated with analytical measurement.

Author details

Marwa S. El-Azazy

Address all correspondence to: marwasaid@qu.edu.qa

Department of Chemistry and Earth Sciences, College of Arts and Sciences, Qatar University, Doha, Qatar

References

[1] Currie LA. Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995). Pure and Applied Chemistry. 1995;**67**:1699-1723

[2] Bièvre PD. The 2012 International Vocabulary of Metrology: "VIM". Accreditation and Quality Assurance. 2012;**17**:231-232

[3] "BIPM—International Vocabulary of Metrology (VIM)," International Vocabulary of Metrology (VIM). Basic and General Concepts and Associated Terms, 3rd ed. Joint Committee for Guides in Metrology (JCGM), 200: 2012 (JCGM 200:2008 with minor corrections)

[4] ICH Harmonized Tripartite Guideline. Validation of analytical procedures: Text and methodology, Q2(R1), Current Step 4th version, Parent guidelines on Methodology Dated November 6, 1996, Incorporated in November 2005

[5] Association of Official Analytical Chemists (AOAC). Official Methods of Analysis. Vol. 1. 15th ed. Arlington, VA: AOAC; 1990. p. 673

[6] Muijlwijk R. Is metrology in chemistry really special? Accreditation and Quality Assurance. 1999;**4**:477-478

[7]   Valcárcel M, Ríos A. A metrological hierarchy for analytical chemistry. TrAc—Trends in Analytical Chemistry. 1999;**18**:68-75

[8]   Valcárcel M. Metrology in chemistry. In: Neidhart B, Wegscheider W, editors. Quality in Chemical Measurements. Berlin, Heidelberg: Springer; 2001. pp. 89-107

[9]   Danzer K, Currie LA. Guidelines for calibration in analytical chemistry. Part 1: Fundamentals and single component calibration. Pure and Applied Chemistry. 1998;**70**: 993-1014

[10] Rodriguez LC, Gracia LG, López EMA, Sánchez JL. Calibration in chemical measurement processes: I. A metrological approach. Trends in Analytical Chemistry. 2001;**20**:195-206

[11] Rodriguez LC, Gracia LG, López EMA, Sendra JMB. Calibration in chemical measurement processes. II. A methodological approach. Trends in Analytical Chemistry. 2003;**20**:620-636

[12] Rodríguez LC, González MGB, Viñas MS, Casado AG, Sáez AMG. Principles of analytical calibration/quantification for the separation sciences. Journal of Chromatography A. 2007;**1158**:33-46

[13] Raposo F. Evaluation of analytical calibration based on least-squares linear regression for instrumental techniques: A tutorial review. Trends in Analytical Chemistry. 2016; **77**:167-185

[14] Magnusson B, Örnemark U. editors. Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics. 2nd ed. Eurachem; 2014. ISBN 978-91-87461-59-0

[15] European Medicines Agency (EMA). Committee for Medicinal Products for Human Use (CHMP). Guideline on bioanalytical method validation; 2011

[16] US FDA. United States Department of Health and Human Services Food and Drug Administration. Guidance for industry. Bioanalytical method validation, no. May 2001

[17] AOAC peer-verified methods, program manual on policies and procedures; 1993

[18] US FDA. United States Department of Health and Human Services Food and Drug Administration, Centre for Drug Evaluation and Research (CDER) and Centre for Veterinary Medicine (CVM). Guidance for industry. Bioanalytical method validation (Biopharmaceutics), no. September 2013

[19] Amemiya T. Advanced Econometrics. Cambridge, Massachusetts: Harvard University Press; 1985. pp. 57-60. ISBN 0-674-00560-0

[20] Wold S, Sjöström M, Eriksson L. PLS-regression: A basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 2011;**58**:109-130. DOI: 10.1016/S0169-7439 (01)00155-1

[21] Hayashi F. Econometrics. Princeton, New Jersey, USA: Princeton University Press; 2000. pp. 3-58. ISBN 0-691-01018-8

[22] Nascimento RS, Froes RES, eSilva NOC, Naveira RLP, Mendes DBC, Neto WB, Silva JBB. Comparison between ordinary least squares regression and weighted least squares regression in the calibration of metals present in human milk determined by ICP-OES. Talanta. 2010;**80**:1102-1109

[23] IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Oxford: Blackwell Scientific Publications; 1997. XML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook. Last update: 2014-02-24; version: 2.3.3

[24] Guideline on bioanalytical method validation. European Medicines Agency; 2011

[25] Thompson RWM, Ellison SL. IUPAC technical report. Harmonized guidelines for single-laboratory validation of methods of analysis. Pure and Applied Chemistry. 2002;**74**: 835-855

[26] National Association of Testing Authorities, Australia (NATA). Technical note 17, guidelines for the validation and verification of quantitative and qualitative test methods; 2013

[27] Irish National Accreditation Board (INAB). Guide to method validation for quantitative analysis in chemical testing laboratories PS 15. 2012

[28] Montgomery DC, Peck EA, Vining GG, Introduction to Linear Regression Analysis, 4th ed. USA; John Wiley & Sons; USA, 2006

[29] Elazazy MS. Determination of midodrine hydrochloride via Hantzsch condensation reaction: A factorial design based spectrophotometric approach. RSC Advances. 2015; **5**:48474-48483

[30] Elazazy MS, Ganesh K, Sivakumar V, Huessein YHA. Interaction of p-synephrine with p-chloranil: Experimental design and multiple response optimization. RSC Advances. 2016;**6**:64967-64976

[31] Draper NR, Smith H. Applied Regression Analysis. New York, USA: Wiley-Interscience; USA. ISBN 0-471-17082-8

[32] Cardone MJ, Palermo PJ, Sybrandt LE. Potential error in single-point-ratio calculations based on linear calibration curves with a significant intercept. Analytical Chemistry. 1980;**52**:1187-1191

[33] Thompson M, Ellison SLR. A review of interference effects and their correction in chemical analysis with special reference to uncertainty. Accreditation and Quality Assurance. 2005;**10**:82-97

[34] ISO11095: 1996. Linear Calibration using Reference Materials. Vol. 1996. Geneva, Switzerland: ISO;

[35] Steger HF. Uses of matrix reference materials. In: Fagelj A, Parkany M, editors. The Use of Matrix Reference Materials in Environmental Analytical Processes. Cambridge: The Royal Society of Chemistry; 1999. p. 128

[36] AMC technical briefs. In: Thompson M, editor. Analytical Methods Committee AMCTB No 37 March 2009, Standard additions: Myth and reality, pp. 1-2

[37] Kadis R. Evaluation of the measurement uncertainty: Some common mistakes with a focus on the uncertainty from linear calibration. Journal of Chromatography A. 2017;**1499**:226-229

[38] Kruve A, Rebane R, Kipper K, Oldekop M-L, Evard H, Herodes K, Ravio P, Leito I. Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part II. Analytica Chimica Acta. 2015;**870**:8-28

[39] Harmonised Guidelines for the Use of Recovery Information in Analytical Measurement, IUPAC Technical Report. Pure and Applied Chemistry. 1999;**1**:337