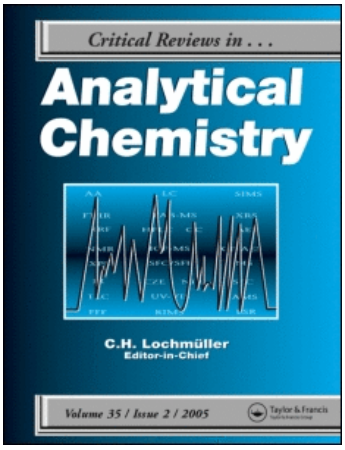

Critical Reviews in Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713400837

The Role of the Robustness/Ruggedness and Inertia Studies in Research and Development of Analytical Processes

L. Cuadros-Rodríguez[a]; R. Romero[a]; J. M. Bosque-Sendra[a]
[a] School of Qualimetrics, Department of Analytical Chemistry, University of Granada, Granada, Spain

To cite this Article Cuadros-Rodríguez, L. , Romero, R. and Bosque-Sendra, J. M.(2005) 'The Role of the Robustness/Ruggedness and Inertia Studies in Research and Development of Analytical Processes', Critical Reviews in Analytical Chemistry, 35: 1, 57 — 69
To link to this Article: DOI: 10.1080/10408340590947934
URL: http://dx.doi.org/10.1080/10408340590947934

Critical Reviews in Analytical Chemistry, 35:57–69, 2005

ISSN: 1040-8347 print
DOI: 10.1080/10408340590947934

The Role of the Robustness/Ruggedness and Inertia Studies in Research and Development of Analytical Processes

L. Cuadros-Rodríguez, R. Romero, and J. M. Bosque-Sendra
School of Qualimetrics, Department of Analytical Chemistry, University of Granada, Granada, Spain

The knowledge of the robustness/ruggedness of analytical processes is an essential feature in the analytical validation. The definitions and the different aims of such studies are given. A dual general procedure for checking robustness/ruggedness in both intrinsic and extrinsic validation is presented and the different steps are discussed. The denomination of "inertia study" is introduced to designate this methodology in extrinsic validation and precision and trueness tests are proposed inside the inertia study. Several experimental designs, depending on the goal of robustness/ruggedness study and the type of variables (quantitative or qualitative) are considered.

Keywords inertia study, intrinsic and extrinsic validation, precision test, robustness and ruggedness study, trueness test, variable effect

INTRODUCTION

Over the last few years, there has been much development in analytical processes. With this in mind, aside from technique selection, it was necessary to select the procedural conditions for this study, so that they could be established using optimization strategies. Finally, the method was applied to real samples, ensuring the quality and correct interpretation of the analytical results. Lately, there has been an emphasis on the principles and approaches to the validation of analytical methods to ensure the quality and integrity measurements and results.

In order to validate the analytical process, it is necessary to know if the analytical properties are appropriate to characterize apply the process. There are different approaches related to validation in analytical chemistry, and the each situation should be clarified. In a recent paper Valcárcel et al. (1) offer a basic approach of validation, taking into account its meaning in generic and quality domain, to establish the convergent and divergent facets with other relevant keywords used in analytical quality, such as qualification, verification or inspection.

In the technical domain, validation of either processes or objects is synonymous with its demonstration of suitability for intended use, and it is generally defined as "the confirmation via examination and provision of objective evidence that the particular requisites for a specific utilization have been fulfilled" (2).

In this sense, the feature of analytical process validation can be applied with two aims (1): On the one hand, the validation process allows characterizing the elements of the analytical process and the analytical process itself, and implies to establish its performance characteristics, fully related to analytical quality properties; in this case, the validation is termed "intrinsic." On the other hand, it can be considered as the process to ensure that the performance characteristics obtained are "fit-for-purpose" to the intended use required by routine acceptance criteria, legislation limits and/or client requirements, and, now, the validation is designated "extrinsic." "Integral" validation is the sequential combination of both alternatives (intrinsic and extrinsic).

The validation of an analytical method has been defined in a more restrictive way as "the process by it is established, by laboratory study, that the performance characteristics of the method meet the requirements for the intended analytical applications" (3) or the "checks to ensure that the performance characteristics of the method are understood and demonstrate that the method is scientifically sound under the conditions which it is to be applied" (4). In this particular approach, the validation of an analytical process consists in the verification, by means of experimental work, of the two fundamental analytical properties (5), representativeness (related to activities outside the laboratory) and accuracy (mainly related to intralaboratory work).

Received 27 March 2001; accepted 5 November 2001; updated 19 January 2005.

Address correspondence to L. Cuadros-Rodríguez, School of Qualimetrics, Department of Analytical Chemistry, University of Granada, E-18071, Granada, Spain. E-mail: lcuadros@ugr.es

Representativeness is achieved through correct sampling, sample conditioning, and sample storage; it can be defined as "the consistency between the analytical results and the composition of the bulk sample studied" (6). This analytical property is not usually involved in the set of attributes of a validated method.

Accuracy is a qualitative concept concerned to analytical results produced by the method to be validated and it can be characterized in terms of trueness and precision of the analytical process (7). Although, for single laboratories, the evaluation of the precision and a verification of the trueness should be enough, a complete validation also implies the estimation of the quality parameters associated with the performance characteristics derived of such basic analytical properties (8).

Furthermore, another performance characteristic of the analytical validation termed ruggedness and/or robustness must be checked (8). Although ruggedness and/or robustness are frequently used interchangeably, a distinction between these two terms must be made (9). Ruggedness is defined as "the degree of reproducibility when the procedure is subjected to changes in external conditions such as different laboratories, analysts, instruments" (3): robustness is defined as "the capacity of an analytical procedure to remain unaffected by small but deliberately introduced variations in method parameter and provides an indication of its reliability during normal usage" (10).

In our opinion, the difference between both terms should be discussed because different information from analytical methods can be obtained. Whereas ruggedness would be reserved to check the analytical method when a noninherent aspect of the analytical process is changed (i.e., column, supplies, brands and/or chemicals, equipment and/or instrument, operators, laboratories, experimental periods, etc.) (11, 12), robustness would characterize the behavior of the analytical method when experimental variables inherent to the analytical process are slightly modified. Here, we should take into account the definition given by Green for chromatographic analytical methods (13): "its ability to remain unaffected by small changes in parameters such as percent organic content and pH of the mobile phase, buffer concentration, temperature, and injection volume." Therefore, ruggedness would inform on the interlaboratory method transferability and robustness would advise on the method practicability and stability related to experimental physico-chemical variables.

Another term, inertia, has been proposed by Cuadros-Rodríguez et al. (8). It is defined as "the capacity of the method to generate inaccurate results when the nominal experimental and/or habitual working conditions are slightly modified." This definition, although more general, is in agreement with the definition given by the French SFSTP for robustness of an analytical methods (14), which states that, it "is its capacity to yield accurate results in the presence of small changes of experimental conditions such as might occur during the utilization of these procedures," where small changes in experimental conditions are meant as "any deviation of a parameter of the procedure compared to its nominal value as described in the procedure of analysis." (English translation from Vander Heyden (15)).

Consequently, for intrinsic validation, robustness/ruggedness identifies the procedural variables, which have a significant effect on the outcome of the analytical process, providing an indication of its reliability during normal usage. Nevertheless, when extrinsic validation is applied, the inertia study is more able to determine if trueness and precision are still satisfactory for a particular analytical problem; in other words, if these analytical properties change significantly when there are small variations of the procedural conditions.

That is why not only the robustness/ruggedness conclusion must be used to indicate the factors that can significantly influence on the outcome of an analytical method, but the inertia study results might be used to know if the trueness and precision change when there is a slightly variation of the variable from the optimal (nominal) values.

Robustness/ruggedness can be partly assured by good system suitability specifications, that is, a series of system suitability parameters should be established to ensure that the validity of the analytical procedure is maintained whenever used (11, 16, 17). System suitability testing is an integral part of many analytical procedures. The test is based on the concept that the equipment, electronics, analytical operations, and samples to be analyzed constitute an integral system that can be evaluated as such. System suitability test parameters to be established for a particular procedure depend on the type of procedure being validated (18).

The robustness/ruggedness tests were originally performed at the end of the method validation process (3), however this involved the risk that when a method is found not to be robust, it should be reoptimized; thus, it is recommended to perform robustness much earlier in the life cycle of the method development (18), by implementing a prevalidation process (19), because the development of an analytical process are not independent of the validation process (1, 20). However, the inertia study must be performed when the analytical method has been definitively optimized, as the authors have expressed in a previous article (8). The flow diagram of the integral validation of an analytical process, according to García-Campaña et al. (8), is showed in Figure 1.

From a metrological point of view, the robustness/ruggedness measures the ability of an optimum to produce truthful and precise results, thus, this study might be suitable to determine the experimental conditions for the recommended procedure (21). Indeed, these performance characteristics are geometrically related with the profile of the response surface (where the real functional relationship between the response and the factors is drawn) in the neighborhood of the studied optimum. In this sense, if the optimum is located in a plateau (Figure 2a), the analytical process will be more robust than if it is in a sharp curve (Figure 2b), where the variation on the response are bigger when the variables change. In this case, the analytical process will not be robust. In the same way, if there are two optimums with same instrumental response value, and one is very pointed (nonrobust) (Figure 2c) and the second one has a "plateau" sufficiently broad (Figure 2d) for comfortable experimental work (very robust), the

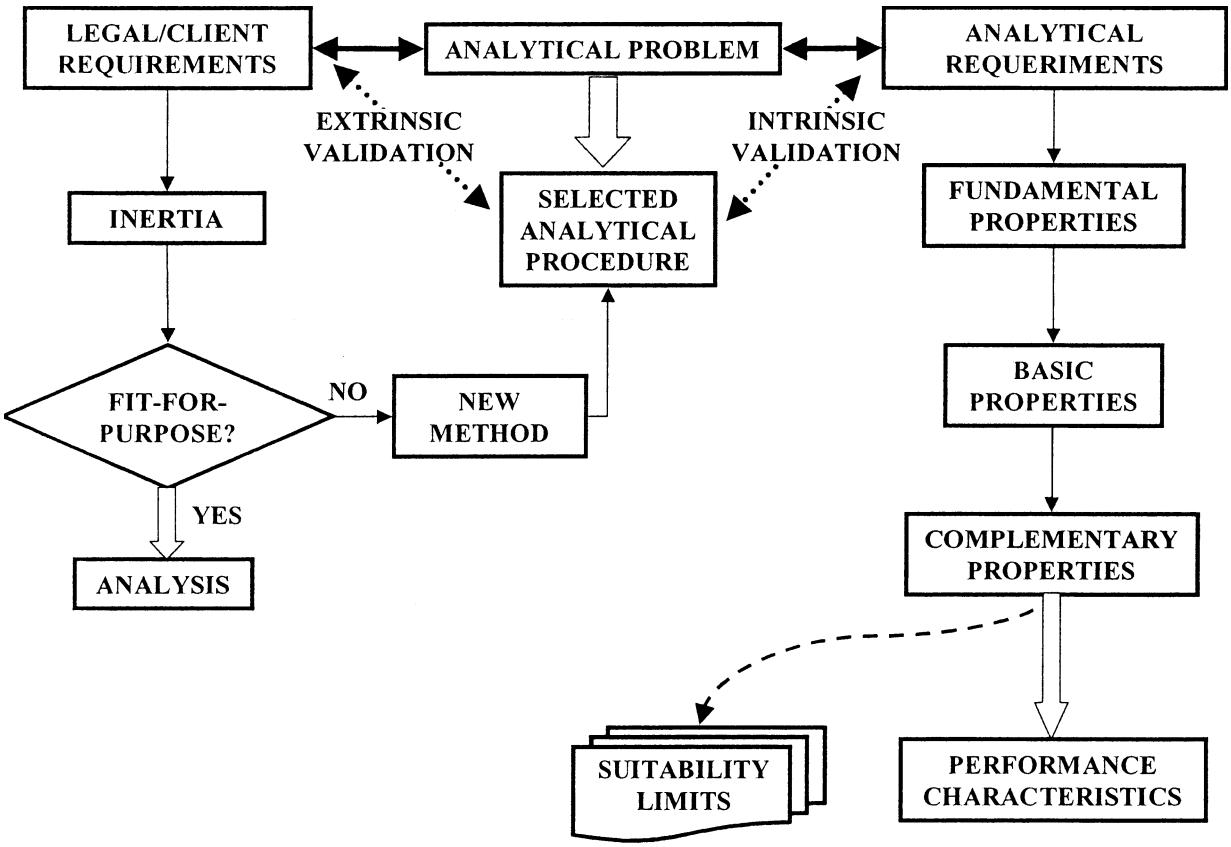

FIG. 1. Flow diagram of the integral validation of analytical processes.

experimental conditions for the second one must be selected as procedural conditions. In this sense, a method is robust toward those variables whose response curves present flatness in the region close to the nominal value.

Different criteria to obtain a robust/rugged optimum in mixture design and HPLC optimization have been proposed, some of them use a Multi Criteria Decision Making (MCDM) technique (22–25) and Taguchi methods (26–28) to select the optimum. However, other criteria, where no MCDM is needed, have been developed (29). In addition, several applications of robustness/ruggedness test for different analytical techniques: capillary electrophoresis (30–33), GCMS (34), size-exclusion chromatography (35), ion-exclusion chromatography (36), HPLC (37–39), HPLC-ICP-MS (40), flow injection analysis (41–43), molecular spectrofluorimetry (44, 45), solid-phase spectrophotometry (46), adsortive stripping voltammetric (47), polarography with multivariate calibration (48) and ^{1}H-NMR (49) has been proposed. On the other hand, other robustness/ruggedness approaches based on the stability of the lineal range (36) or of the calibration curves have also been recently published (50).

A general methodology to study robustness/ruggedness has been established in few papers. The first one was proposed by Youden (51), in which a Plackett-Burman design and the appropriate statistical treatment was used. The following proposed methodologies, which are based on the Youden procedure, use different experimental designs (52–54) or different statistical interpretations (55, 56). A review covering the use of robustness and ruggedness in Analytical Chemistry has been edited (57) and a guidance for Robustness/Ruggedness Tests in Method Validation has been published (58). Finally, expert systems for the study of ruggedness in HPLC method validation were also developed (59–61).

In this article, the robustness/ruggedness is examined, giving the essential features of a general methodology inside the validation process, increasing the amount of information available regarding the quality of the analytical process, which allows us to test the presence of main total effects, curvature, and main sided effects of the different variables. Also, the inertia study is implemented in the validation framework and statistical tests to check globally both precision and trueness of the analytical process are suggested.

THEORETICAL ASPECTS

The robustness/ruggedness test is usually carried out using Experimental Design Methodology. In this sense, an experimentation plan is defined as the experimental strategy constituted by the experiences in relation to the selected variables (factors) so the realization of each experience implies that each variable

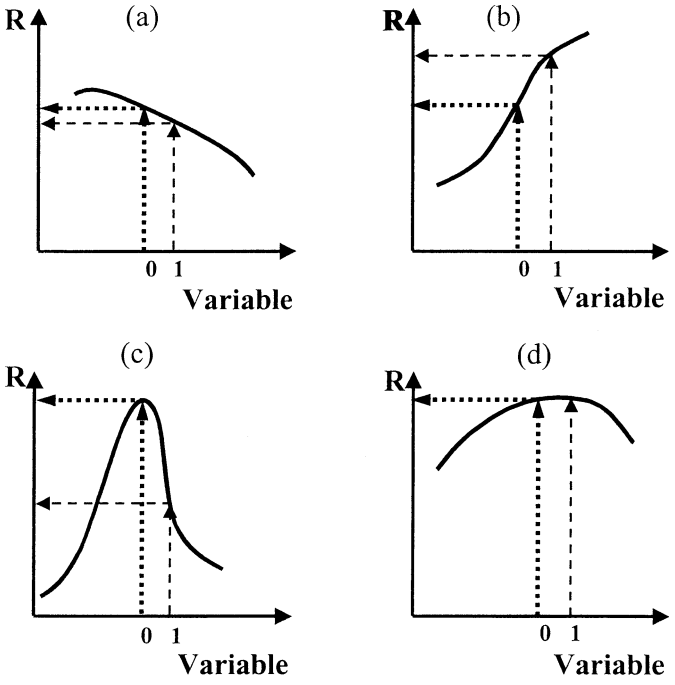

FIG. 2. Geometrical relationship between robustness and the response surface profile: (a) optimum located in a plateau, (b) optimum located in a sharp curve, (c) nonrobust maximum, and (d) robust maximum.

takes a certain value (level), which is defined by the structure of the design, and depends on the type of design. Taking into account that the matrix of experiences is the table that gathers the characteristics of each experience in relation to the different levels of the factor, the experimentation plan is the application of the formal structure of the matrix to a particular problem.

Two types of variables can be distinguished in the experimental design context: quantitative and qualitative. The values of quantitative variables are associated with points on a numeric scale. As well, they are divided in continuous (associated with any value on that scale) or discrete (associated with certain values on the scale). Qualitative variables cannot be characterized by an order of magnitude. Also, they can be distinguished in binary (yes/no, lower/higher) and manifold (more than two possibilities, such as type of column used for a chromatographic separation).

The different levels of the variables, used in the experimentation plan, are expressed as coded values, such as $0, +1, -1$, in the matrix of experiences. The nominal level is usually 0, whereas the positive $(+1)$ and negative (-1) values are deviations from this nominal level in the system under test.

The experimental strategy is used to establish a relation between the responses and the variables, and the target can be to explore an experimental field (exploratory) or to elaborate the best possible descriptive models of the studied phenomenon. If the mathematical model cannot be theoretical (it does not come from a physical law), polynomial functions are usually used to fit the experimental data, applying the regression by ordinary least squares.

To be able to fit the model, the number of experiences, N, must be equal or greater than the number of terms, p, of the polynomial function. The order of the polynomial function selected to fit the experimental data, depends on the number of levels, l, studied by each variable, being k the number of variables. For two level designs, only a linear model can be fit, whereas if a three level design is used, quadratic models can be fit.

Main Total and Main Sided Effects of the Variables

For the robustness test, a mathematical model of additives effects is generally used, expressed as:

$$R^p = R_{(0)} + E_1 + E_2 + \cdots\cdots + E_k + \Delta, \qquad [1]$$

where R^p is the predicted response by the model; $R_{(0)}$ is the expected response of the analytical system when a experience is performed under nominal conditions (i.e., nominal response); E_j indicates the main effect of the j variable and it represents the variation introduced in the response when the value of the variable j changes in relation to the nominal value; and Δ is the total error of prediction ($\Delta = R^{real} - R^p$), which can be seen as the sum of two contributions, random or experimental error, which are unidentified and normally distributed and systematic error or bias, understood as a persistent positive or negative deviation from the accepted reference value.

The effect of each variable can be positive or negative according to the response increases or decreases with regard to the nominal response. These main effects could be estimated from (see Figure 2):

$$E_j = R_{(1)j} - R_{(0)}, \qquad [2]$$

where $R_{(1)j}$ indicates the found response when the experience is carried out for the modified value of the j variable.

New effects, denominated interaction terms, could be taken into account when the effect of a variable on the response depends on the value of the others. In these cases, the model is:

$$R^p = R_{(0)} + E_1 + E_2 + \cdots\cdots + E_k + E_{1,2} + E_{1,3} + \cdots\cdots + E_{k-1,k} + \Delta, \qquad [3]$$

where, for instance, $E_{1,2}$ indicates the effect of the interaction of the variable 1 with the variable 2. In general, the robustness study is mainly concerned about the main effects of variables, since the range of the factor levels is so small that the interaction terms are negligible.

On the other hand, when the value of the variables is decreased (-1) and increased $(+1)$ respect to the nominal value (0) in the same experiment, it can be possible to distinguish between the main effects. Main total effects, E_j, indicates the total variation that is introduced in the response, when the factor changes its value from lower (-1) until upper $(+1)$, whereas main sided effects, $E_{(+)j}$ and $E_{(-)j}$, are denominated when the variable is

modified from the nominal value (0) to the upper or lower, where:

$$E_j = R_{(+1)j} - R_{(-1)j}; \quad E_{(-1)j} = R_{(-1)j} - R_{(0)}; \quad E_{(+)j} = R_{(+1)j} - R_{(0)} \qquad [4]$$

$$E_j = |E_{(-)j}| + |E_{(+)j}| \qquad [5]$$

These two sided main effects usually have opposite signs, and they can be equals or different in magnitude (Figure 3). If the effect of a factor is linearly related to their levels (linear model), and if the nominal level is situated in the middle of the interval, $E_{(-)j}$ and $E_{(+)j}$ will be uniform (Figure 3a), and they will only differ due to experimental error. On the other hand, if the effect of a factor is not linearly related to their levels (nonlinear model), the two main sided effects will be nonuniform, being $E_{(-)j}$ and $E_{(+)j}$ (Figure 3b). In this case, the difference between these two main sided effects will be higher than experimental error.

When these main sided effects are included in the mathematical model, this one is given by:

$$R^p = R_{(0)} + [E_{(-)1} \text{ or } E_{(+)1}] + \cdots\cdots + [E_{(-)k} \text{ or } E_{(+)k}] + \Delta, \qquad [6]$$

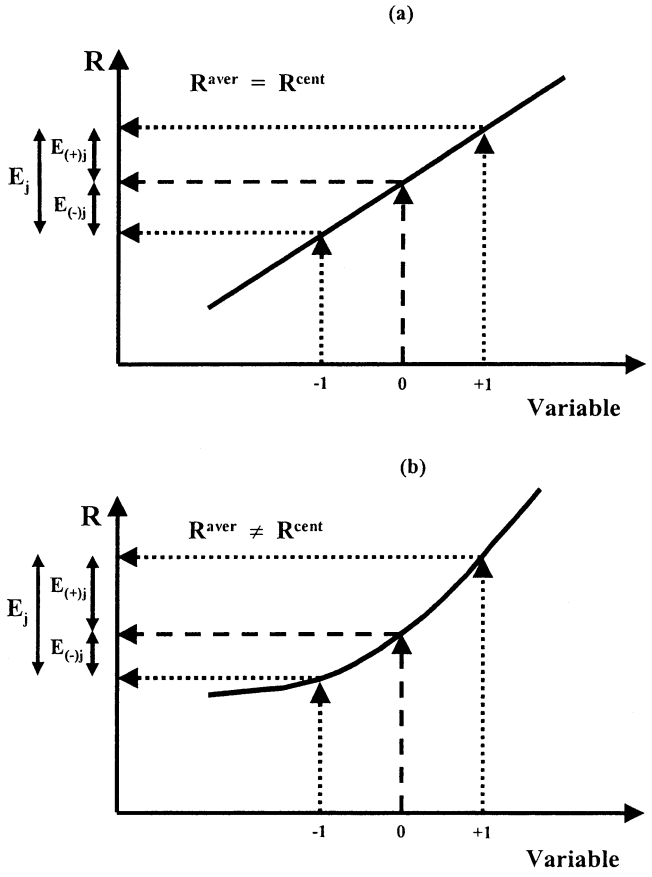

FIG. 3. Difference among main sided effects in: (a) linear model, (b) nonlinear model.

where each j main sided effect must be selected according to the j variable has been increased or decreased.

To calculate these main effects, two level experimental designs are efficient enough to estimate just the main total effects (44, 55), but for the simultaneous estimation of main sided effects, it requires a three level design (44, 46, 54).

In general, it would be more coherent to interpret the analytical effect as the deviation that takes place in the response from the modified value (-1 or $+1$) in relation to the nominal value (0). In this case, and when the main sided effects are uniforms, the analytical effect and the main sided effect are equals and these ones coincide with half of the main total effect.

Regression Coefficients and Effects

It is also possible to apply Multiple Linear Regression (MLR) to estimate each effect (30), and therefore, the mathematical model are given by:

$$R^p = b_0 + b_1X_1 + b_2X_2 + \cdots\cdots + b_kX_k + \Delta, \qquad [7]$$

where b_j represents the regression coefficients and the X_j terms are the variables. Each term can be identified with its homologous one in the effects model, thus, b_0 is the nominal response (all X_j terms are zero), and the b_jX_j terms are the main effects.

The regression coefficients provide information on the behavior of the analytical system close to the nominal value, since they measure the sensitivity of the response in relation to the considered variables. In fact, the estimated regression coefficients obtained, represent the change that takes place in the response for each variable unit that it is modified (62). The regression coefficients are always identified with the analytical effects, and their numerical value agree with the effect when the experimentation plan has been carried out from a two levels matrix, in which the variables are codified with 0 and $+1$ (or 0 and -1), whereas the value of the main total effect is the double that the corresponding regression coefficient, due to the own definition of analytical effect:

$$E_j = 2b_j. \qquad [8]$$

If the main sided effects are nonuniforms, a quadratic model must be considered:

$$R^p = b_0 + b_1X_1 + b_{11}X_1^2 + \cdots\cdots + b_kX_k + b_{kk}X_k^2 + \Delta, \qquad [9]$$

which can be also expressed as:

$$R^p = b_0 + (b_1 + b_{11}X_1) \cdot X_1 + \cdots\cdots + (b_k + b_{kk}X_k) \cdot X_k + \Delta, \qquad [10]$$

where it can be appreciated that the main effect, given by the coefficient of each variable in brackets, depends on the value of the own variable and therefore it is different to lower or upper values than the nominal value.

As it has been quoted above, other nonlinear terms, which are different from the quadratics, can take place when interactions are not negligible:

$$R^p = b_0 + b_1X_1 + b_2X_2 + \cdots + b_kX_k + b_{12}X_1X_2 + \cdots + b_{k-1,k}X_{k-1}X_k + \Delta \quad [11]$$

and the same as before:

$$R^p = b_0 + (b_1 + b_{12}X_2)\cdot X_1 + \cdots + (b_{k-1} + b_{k-1,k}X_k)\cdot X_{k-1} + b_kX_k + \Delta, \quad [12]$$

where it can be observed that the effect of a variable (in brackets) depends on the value of another variable.

Curvature of the Model

When main-sided effects are nonuniforms, the mathematical model must be completed with nonlinear terms, and a response surface with curvature appears. Consequently, if curvature is demonstrated, nonuniform effects can be expected, taking into account that the interactions are usually negligible in a robustness study.

If both main sided effects are equals in magnitude, its sign is opposite, and therefore, the differences of their absolute values are equal to zero. As a result, when experiences are carried out at both modified values (upper and lower) of the variables, the average response (R^{aver}) is the same than the response at nominal conditions (response in the central point (R^{cent}) (see Figure 3). When curvature exits, the main sided effects are different to each other and, now, the difference of its absolute values is different to zero. Likewise, the average response is not equal to the response at central point, and this fact will be taken advantage of to study the existence of curvature in the model (44, 46)].

As example to explain it, let us to consider one-variable model (Table 1 indicates the calculations and shows the found results). Assuming a linear relationship between the response and the factor, $R = b_0 + b_1X_1$, the main total effect is $2b_1$ and the main sided effects are b_1 and $-b_1$, respectively, and both responses, R^{aver} and R^{cent}, are similar to b_0.

When curvature exists, a second order model must be applied, $R = b_0 + b_1X_1 + b_{11}X_1^2$. The main total effect is the same than the prior case but the main sided effects are now nonuniforms (they are $b_1 + b_{11}$ and $-b_1 + b_{11}$, respectively). Consequently, the average response and the central response are now different and their values are b_0+b_{11} and b_0, respectively (if in fact, they differ in the value of the coefficient of the quadratic term).

General Model with Qualitative Variables

The qualitative factors can be varied only in discrete steps and the levels must be specified, disappearing the distinction between upper and lower. Whereas for quantitative factors, one total effect can be calculated, for qualitative factors, more than one effect must be calculated. For example, if a qualitative factor is defined at three levels (L_1, L_2, L_3), three "effects" (56) may be present: the effect caused by the change from L_1 to L_2, from L_1 to L_3 or from L_2 to L_3. The general effect model can be defined for one qualitative factor L as:

$$R^p = R_{(0)} + E_{L_1,L_2} + E_{L_1,L_3} + E_{L_2,L_3} + \Delta, \quad [13]$$

where each $E_{Li,Lj}$ is calculated according to:

$$E_{L_i,L_j} = \frac{\Sigma R(L_i)}{n_{L_i}} - \frac{\Sigma R(L_j)}{n_{L_j}}, \quad [14]$$

and $\Sigma R(L_i)$ and $\Sigma R(L_j)$ are the sums of the responses where the factor L is at levels i and j respectively, and n_{Li} and n_{Lj} are the number of runs with factor L at level i and j, respectively.

If qualitative and quantitative variables are used in the robustness study, when interaction and quadratic terms have been considered negligible, the model is performed with both types of variables as:

$$R_p = R_{(0)} + E_1 + \cdots + E_k + E_{L_1,L_2} + E_{L_1,L_3} + E_{L_2,L_3} + \Delta. \quad [15]$$

TABLE 1
Estimation of main total effect, main-sided effects, average response, and central response from a one-variable mathematical model using regression coefficients

Model	Feature	Calculation	Result
Linear	E_j	$R_{(+1)} - R_{(-1)} = [b_0 + b_1\cdot(+1)] - [b_0 + b_1\cdot(-1)]$	$2b_1$
($R_p = b_0 + b_1X_1$)	$E_{(+)j}$	$R_{(+1)} - R_{(0)} = [b_0 + b_1\cdot(+1)] - [b_0 + b_1\cdot(0)]$	b_1
	$E_{(-)j}$	$R_{(-1)} - R_{(0)} = [b_0 + b_1\cdot(-1)] - [b_0 + b_1\cdot(0)]$	$-b_1$
	R^{aver}	$1/2\ \{R_{(+1)} - R_{(-1)}\} = 1/2\{[b_0 + b_1\cdot(+1)] + [b_0 + b_1\cdot(-1)]\}$	b_0
	R^{cent}	$R_{(0)} = b_0 + b_1\cdot(0)$	b_0
Quadratic	E_j	$R_{(+1)} - R_{(-1)} = [b_0 + b_1\cdot(+1) + b_{11}\cdot(+1)^2] - [b_0 + b_1\cdot(-1) + b_{11}\cdot(-1)^2]$	$2b_1$
($R_p = b_0 + b_1X + b_{11}X_1^2$)	$E_{(+)j}$	$R_{(+1)} - R_{(0)} = [b_0 + b_1\cdot(+1) + b_{11}\cdot(+1)^2] - [b_0 + b_1\cdot(0) + b_{11}\cdot(0)^2]$	$b_1 + b_{11}$
	$E_{(-)j}$	$R_{(-1)} - R_{(0)} = [b_0 + b_1\cdot(-1) + b_{11}\cdot(-1)^2] - [b_0 + b_1\cdot(0) + b_{11}\cdot(0)^2]$	$-b_1 + b_{11}$
	R^{aver}	$1/2\ \{R_{(+1)} - R_{(-1)}\} = 1/2\{[b_0 + b_1\cdot(+1) + b_1\cdot(+1)^2] + [b_0 + b_1\cdot(-1) + b_1\cdot(-1)^2]\}$	$b_0 + b_{11}$
	R^{cent}	$R_{(0)} = b_0 + b_1\cdot(0) + b_1\cdot(0)^2$	b_0

When qualitative factors are used, only conclusions respect to the examined levels selected can be made and it does not allow drawing any conclusion about the total population. For example, if the type of column is examined using three different types of column, only conclusions about the selected columns can be made.

METHODOLOGY

The methodology of the robustness/ruggedness test is different according to the type of validation (intrinsic or extrinsic) used. In this sense, if intrinsic validation is used, the main goal is to identify the experimental variables, which affect significantly to response variables, and to anticipate the problems that may arise during its application, so an effect signification test must be applied in order to know the significant experimental variables which require a strict control.

On the other hand, if extrinsic validation is applied, the goal is to verify if the performance characteristics (trueness and precision) are changed or not, so the global precision and trueness test (explained below) must be used. If at least one of these tests is significant, an intrinsic validation can be applied in order to identify the significant variables. In this sense, depending on the type of validation, there are different possibilities to apply a robustness study, as can be observed in Figure 4, but both of them can be made at the same time if the analyst needs it.

Intrinsic Validation: Robustness and Ruggedness Study

The first step is the selection of variables and their levels. The variables (quantitative or qualitative) must be related to the analytical procedure and are selected for the description of the analytical method (63). These factors are related to the sample preparation, separation, and/or detection steps. The number of factors selected is usually limited for practical reasons. In general, when the number of factors is higher than eight, it is recommended to split the set of factors into two or three groups (64), provided that interactions between variables of different groups are not suspected.

For robustness study, quantitative variables are mainly considered. The factor levels are usually defined symmetrically around the nominal level prescribed in the operating procedure although asymmetrical intervals around the nominal level can be also examined (65). The selection of the levels can be based in the precision or the uncertainty (66), or in a percentage previously selected (similar or different for each variables) (44, 46). The number of levels tested for each factor could be three (low, high, and nominal levels) but, for a economy of work, is better to performance a design with two levels (low and high) and at least three central points (nominal level).

The responses, defined by the analyst, can describe a quantity such as the content of a substance, peak area in chromatographic methods, absorbance in spectrophotometric methods, or it should consider quality parameters such as resolution and migration times.

After the selection of factors, their levels and response variables, a design must be selected. These designs are usually saturated designs, such as two-level saturated fractional designs (46) or Plackett-Burmann designs (55, 67). In the saturated designs, the number of experiments required, without counting the central points, is equal to the number of variables plus one ($N = k + 1$), and all main effects are not confounded with each other. For large number of variables, supersaturated designs are used (68), in which the number of factors examined exceeds the number of experiments. These designs do not allow estimating the effects of the individual factors because of confounding between the main effects, but the total variance of the responses in them could be used as a measure for the robustness of the method.

After the main total effects are calculated (55, 64), it is necessary to identify the significant main total effects (see below), using an analysis of variance (44, 69), or a t-test (46, 64, 70) (see the next section). In this last one, it must be verified the existence of a linear relationship between the factor and the response. When there is curvature, the main-sided effects needs to be estimated (44, 46) using three-level designs, such as reflected two-level designs [15], three-level Plackett-Burmann designs (54, 67), three-level factorial designs (31, 44), Box-Behnken designs (71), or D-optimal desings (47).

On the other hand, qualitative and quantitative variables are habitually included in the ruggedness study and different types of designs can be used. One choice is the use of nested designs for qualitative variables (15, 72), whereas quantitative variables can be studied by factorial designs. However both types of variables can be studied in the same design using asymmetrical factorial designs (70), which contain factors examined at different levels. It is clear that the number of experiments increases when qualitative factors are included.

In all the cases, the different experiments must be carried out in a random order, especially if central points are used, to take into account uncontrolled factors. Sometimes, if drift occurs, it is necessary to correct the response results to obtain a real estimation of factors (73).

The conclusions obtained from the statistical analysis of the experimental data set can be distorted by the presence of outliers, so a previous detection and removing process of them is needed (46, 74). The presence of outliers requires a new realization of experimental design, or at least, the runs that give outliers must be repeated.

Finally, when the factors that cause significant determination in performance are identified, the system suitability limits can be established (75, 76). In this way, Vander Heyden et al. (69) proposed the use of the worst-case situation to define the system suitability limits. In this case, only main total or linear effects are considered in the prediction of the worst-case situation, assuming a linear relationship between the variables and response. This approach is only valid when the curvature test is not significant because, on the contrary, the significant main sided effects might be used to obtain the system suitability limits.

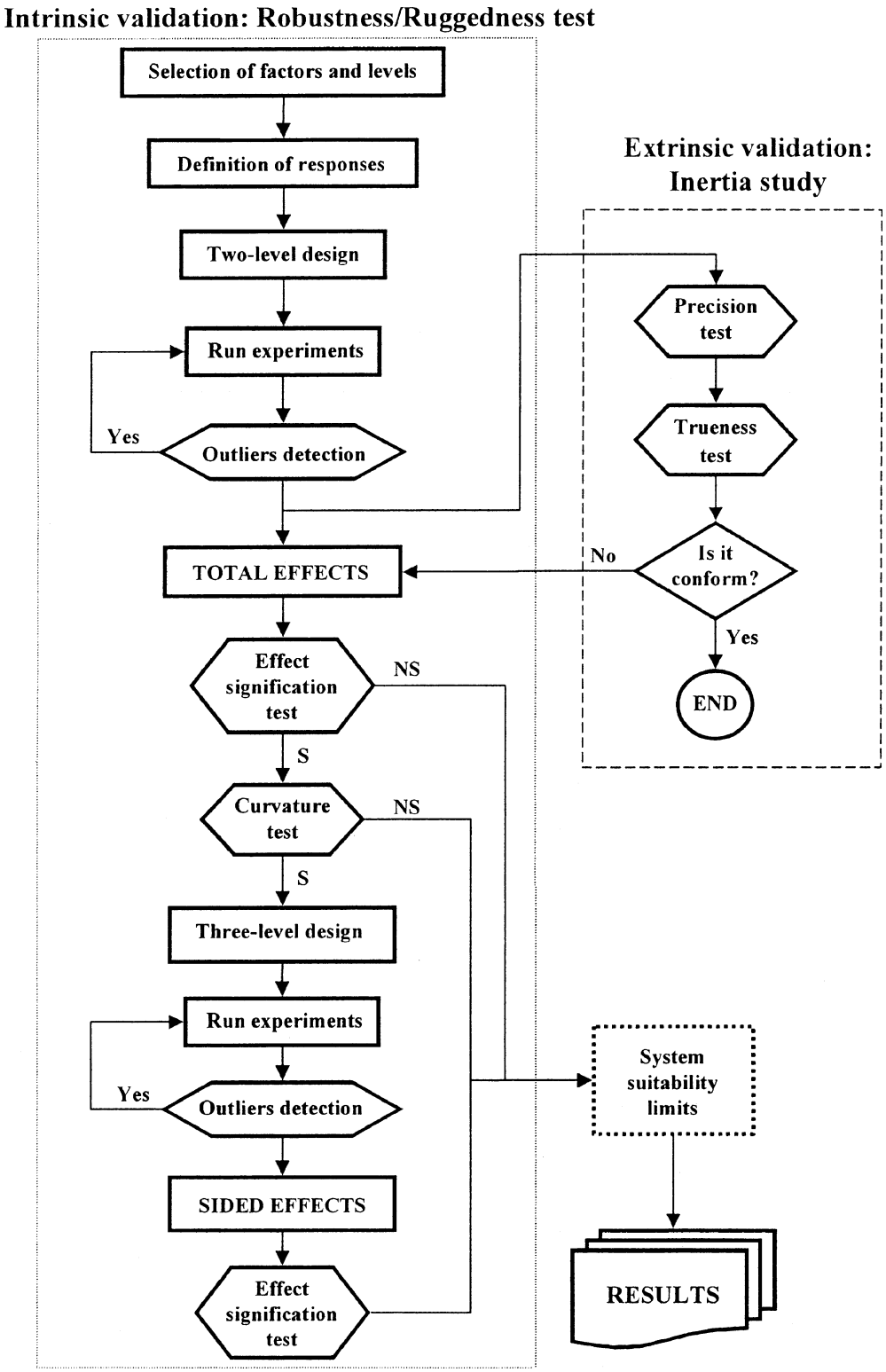

FIG. 4. Flow chart of the general methodology of the robustness/ruggedness and inertia study to validate analytical processes. (S = significant conclusion; NS = non significant conclusion).

Effect Signification Test

In order to identify statistically significant effects, a t-test is usually used. The statistic is $|E_x|/(SE)_E$, where $(SE)_E$ is the standard error associated to the effect, and this value is compared with the critical two-tailed Student-t value, $t_{crit}(\alpha, f)$, for a confidence level normally 95 percent ($\alpha = 0.05$) with f degrees of freedom. A variable is considered to be significant if the calculated t-test value is lower than the tabulated one. This test determines whether the modification of the instrumental response due to the effect of the variation of each experimental variable is less or equal than the random error of the procedure, that is, if a bias takes place, in which case it is considered not significant, and if it is greater, the effect is significant. In this sense, the obtained P-value indicates the probability that the bias found for the effect is only because of random experimental errors, and thus a high value for P-value corroborates the null hypothesis, while values of less 0.05 might indicate the effect is significant.

The conclusions of this test are different, depending on the way to estimate $(SE)_E$, which represents the experimental variability within the design, and can be estimated:

(1) from the residual standard deviation obtained for all the data set (42, 55);
(2) from the effects of the two-factor interactions (55, 77);
(3) using dummy factor chosen in such way they are confounded with the most probable chemically consistent interactions, and having no influence on the design structure (45, 55, 67, 78);
(4) from the standard deviation on predicted response estimated from calibration data set (44–46);
(5) from replicates of any experimental run of the design (45, 62);
(6) from randomized replicate measurements at the nominal level of the procedure-related variables (67);
(7) from the distribution of effects (algorithm of Dong) (56). The number of freedom degrees depends on the way to estimated $(SE)_E$.

In this sense, $(SE)_E$ obtained from replicate measurements at nominal level or from replicates at any experimental run of the design, underestimates the experimental error, since it can be considered as a measure of variability determined under repeatability conditions. On the other hand, if the standard deviation from all the data set is used, $(SE)_E$ is an overestimation of the experimental error (55). However, the best choices to calculate $(SE)_E$ are dummy factors, randomized replicates at nominal level in the design, or using the effects of the two-factor interactions (67).

Curvature Test

When a-two level design is chosen for the robustness study, it assumes that a linear relationship between the factors and the responses exists. This assumption must be verified using a curvature test and it must be only applied if there are significant main total effects, since if there are not significant effects, the linear relationship can be assumed, and therefore, the main sided effects are similar.

This test is based on determining whether or not the mean of the observed responses, R^{aver}, belongs to the confidence interval for the response at nominal condition, R^{cent}:

$$R^{aver} \in R^{cent} \pm t \cdot s_d. \quad [16]$$

In this equation, t is the tabulated two-tailed Student t-value with $r - 1$ freedom degrees and $\alpha = 0.05$, where r is the number of replicate measurements at nominal level. It can be observed that this test is similar to the trueness one (following), where it will be explained.

A significant conclusion implies curvature in the global mathematical model, and not necessarily in a particular effect, and for this, a subsequent individual study of the sided effects should be carried out. For this study, three-level designs must be used (44, 54), examining the factors that have significant total effects.

Extrinsic Validation: Inertia Study

This type of validation requires the same experimental work than before, but the goal is not the same. That is why, after the detection of outliers, the methodology is different (Figure 4), and now, the inertia study is a robustness/ruggedness study just minimized to ensure that precision and trueness (performance characteristics) are still satisfactory, so a precision and trueness test must be applied. These two tests have not been described in bibliography previously and they will be considered below with a deeper rigor.

Precision Test

The purpose of this test is to establish if slight modifications in the factors, when an inertia study is carried out, causes a lost in the precision of the analytical process in nominal conditions. The uncertainty associated with the robustness experiments, expressed as standard deviation (s_R^{exp}), is given by:

$$s_R^{exp} = \sqrt{\frac{\Sigma E_j^2}{k}}, \quad [17]$$

where E_j is the effect of each factor and k is the number of main estimated effects. The associated freedom degrees are k.

On the other hand, the uncertainty associated with the nominal procedure is needed, and it can be obtained from a repeatability test (using central points); the standard deviation is calculated by:

$$s_R^{cent} = \sqrt{\frac{\Sigma \left(R_i^{cent} - \bar{R}^{cent}\right)^2}{r - 1}}, \quad [18]$$

where R_i^{cent} is each of the responses obtained at nominal level, $\bar{R}^{cent}$ is the mean of these values and r is the number of replicate measurements at nominal level. The associated freedom degrees are $r - 1$.

If it is possible, it could be better to calculate this uncertainty as standard deviation of the predicted response in the central point from the linear regression analysis on the calibration data set of the analytical method (s_R^{reg}) (44, 46):

$$s_R^{reg} = \sqrt{\frac{s_{resid}^2}{n} + \frac{s_b^2}{b^2} \cdot d_R^2} = \sqrt{\frac{s_{resid}^2}{n} + SRD_b^2 \cdot d_R^2}, \quad [19]$$

where s_{resid} is the standard deviation of the residuals; n, the total number of pairs of points used to calculate the regression line; b, the slope of calibration (or sensitivity); s_b, the standard deviation of the slope; SDR_b, the relative standard deviation of the slope ($SDR_b = s_b/b$); and d_R, the distance (or difference), in response units, between the predicted response and the mean response of the calibration data set. In this case, the freedom degrees are $n - 2$.

The analyst can also establish this uncertainty based on the knowledge of the analytical procedure. In this case, the freedom degrees are infinity.

The variances are then compared in a F-test. The statistic with $(k, r - 1)$ freedom degrees, if variance from a repeatability measurements is used, is calculated to the following expression:

$$F_{calc} = \frac{\left(s_R^{exp}\right)^2}{\left(s_R^{cent}\right)^2} \quad [20]$$

or, if variance is estimated from a regression analysis, from:

$$F_{calc} = \frac{\left(s_R^{exp}\right)^2}{\left(s_R^{reg}\right)^2} \quad [21]$$

with $(k, n - 2)$ freedom degrees.

If the test is significant, it means that at least one factor is significant and, when its nominal level is changed, the uncertainty increases. If the test is not significant, the precision of the analytical procedure does not vary although the studied factors change slightly.

Trueness Test

This test evaluates the probability that bias is not zero (whereby systematic error is present) when the inertia study is applied. The trueness test requires to determine whether or not the average response of the N experiments of the design used in the inertia study, R^{aver}, Eq. (22), belong to the confidence interval for the response at nominal conditions or predicted by the calibration curve, Eq. (23), where R_i is the response of each experiment.

$$R^{aver} = \frac{\Sigma R_i}{N} \quad [22]$$

$$R^{aver} \in \bar{R}^{cent} \pm t \cdot s_d \quad [23]$$

The calculation of s_d depends on the results of the previous precision test:

1) If the result of the precision test confirms that the precision does not change, a pooled standard deviation is calculated:

$$s_d = \sqrt{\left(\frac{k \cdot (s^{exp})^2 + (r-1)(s^{cent})^2}{k + r - 1}\right)\left(\frac{1}{k} + \frac{1}{r-1}\right)} \quad [24]$$

In this case, the freedom degrees are $k + r - 1$.

2) If the variances are different, s_d is calculated by Welch-Satterthwaite approximation (79):

$$s_d = \sqrt{\frac{(s^{exp})^2}{k} + \frac{(s^{cent})^2}{r-1}} \quad [25]$$

The freedom degrees are now obtained using Eq. [26].

$$\nu_d = \frac{s_d^4}{\frac{((s^{exp})^2/N)^2}{k} + \frac{((s^{cent})^2/r)^2}{r-1}} \quad [26]$$

If the result of this test is not significant, with a significance level of 95 percent ($\alpha = 0.05$), it means that the difference between R^{aver} and R^{cent} is due to random error, and the trueness does not change when the slight variation of the studied factors are produced. If the test is significant, it means that, as the precision test, at least, one factor is significant, and it can be detected, applying the previous methodology.

CONCLUSIONS

The robustness/ruggedness study is an essential subject in the validation of the analytical processes. The purpose of these studies in intrinsic and extrinsic validation is different. The robustness/ruggedness study allows us to estimate the main total effects, to check the curvature of the model around the procedural conditions and to determine the main sided effects for an intrinsic validation, that is, to characterize the behavior of the analytical process. On the other hand, the inertia study in an extrinsic validation permits to check performance characteristics of the analytical process, obtaining useful information, which must be a fundamental part of a method validation.

REFERENCES

1. M. Valcárcel and A Ríos, A general approach to validation in analytical chemistry. *Quim. Anal.* 19 (2000):95–101.
2. ISO 9000, *Quality Management System. Fundamentals and Vocabulary* ISO, Geneva, 2000.
3. The United States Pharmacopeia USP XXIII. *Validation of Compendial Method*, United States Pharmacopeia Convention, Rockville, 1995.
4. EURACHEM-EAL G14 1997, (Previously Eurachem Guidance Document No 1/WELAC Guidance Document No. WGD 2). *Guidance on the interpretation of the EN 45001 series of standards and ISO/IEC Guide 25*, Laboratory of the Government Chemist, Teddington, UK.
5. M. Valcárcel and A. Ríos, The hierarchy and relationships of analytical properties. *Anal. Chem.* 65 (1993):781A–787A.
6. A. Ríos and M. Valcárcel, Representativeness of analytical results. *Analyst* 119 (1994):109–112.

7. ISO 5725-1, *Accuracy (Trueness and Precision) of Measurement Methods and Results. Part I: General Principles and Definitions*. ISO, Geneva, 1994.
8. A. García-Campaña, J. M. Bosque-Sendra, L. Cuadros-Rodríguez, and E. Almansa-López, A framework for in-house accuracy validation of analytical procedures. *Biomed. Chromatogr*. 14 (2000):27–29.
9. EURACHEM. *The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related topics*, Laboratory of the Government Chemist, Teddington, UK, 1998.
10. ICH-Q2A, *Validation of Analytical Procedures: Definitions and Terminology*. The International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use, 1995.
11. D. R. Jenke, Chromatographic method validation: A review of current practices and procedures. III. Ruggedness, re-validation and system suitability. *J. Liq. Chrom. & Rel. Technol*. 19 (1996):1873–1891.
12. M. Seartz and I. S. Krull, *Analytical method Development and Validation* (New York: Marcel Dekker, 1997, p. 67).
13. J. M. Green, A practical guide to analytical method validation. *Anal. Chem*. 68 (1996):305A–309A.
14. Guide de Validation Analytique, Rapport d'une Commision de la Société Française Sciences et Techniques Pharmaceutiques (SFSTP). I. Méthodologie. *STP Pharma Practiques* 2 (1992):205–226.
15. Y. Vander Heyden, The ruggedness of analytical methods. *Analusis* 22 (1994):M27–M29.
16. *Reviewer Guidance: Validation of Chromatographic Method*s. Center for Drug Evaluation and Research (CDER), FDA, 1994.
17. Y. G. Li, M. Li, G. X. Chou, Z. T. Wang, and Z. B. Hu, Ruggedness/robustness evaluation and system suitability test on United States Pharmacopeia XXVI assay ginsenosides in Asian and American ginseng by high-performance liquid chromatography. *J. Pharm. Biomed. Anal*. 35 (2004):1083–1091.
18. ICH-Q2B, *Text on Validation of Analytical Procedures: Methodology*. International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use, 1996.
19. J. Lang and S. M. Bolton, A Comprehensive Method Validation Strategy for Bioanalytical Applications in the Pharmaceutical Industry, in *Pharmaceutical and Biomedical Applications of Liquid Chromatography*, Fell, A.M. and Riley C.M., Eds.; (London: Pergamon, 1994, ch. 11).
20. P. Bruce, P. Minkkinen, and M. L. Riekkola, Practical method validation: validation sufficient for an analysis method. *Mikrochim. Acta* 128 (1998):93–106.
21. C. Vandenbosch, C. Vannecke, and D. L. Massart, Optimization of the separation of (6R)- and (6S)-leucovorin and evaluation of the robustness of the optimum. *J. Chromatogr. A*. 592 (1992):37–41.
22. M. W. Margriet, B. Hendriks, J. H. de Boer, and D. A. Doornbos, Multicriteria decision making. *Chemometr. Intell. Lab. Syst*. 16 (1992):175–191.
23. P. F. Vanbel, B. L. Tilquin, and P. J. Schoenmakers, Criteria for developing rugged high-performance liquid chromatographic methods. *J. Chromatogr. A*. 697 (1995):3–16.
24. S. Goga-Remont, S. Heinisch, and J. L. Rocca, Use of optimization software to determine rugged analysis conditions in high-performance liquid chromatography. *J. Chromatogr. A* 868 (2000):13–29.
25. M. R. Hadjmohammadi and F. Safa, Multi-criteria decision making in micellar liquid chromatographic separation of chlorophenols. *J. Sep. Sci*. 27 (2004):997–1004.
26. J. H. De Boer, A. K. Smilde, and D. A. Doornbos, Introduction of a robustness coefficient in optimization procedures: Implementation in mixture design problems. Part I: Theory. *Chemometr. Intell. Lab. Syst*. 7 (1990):223–236.
27. J. H. De Boer, A. K. Smilde, and D. A. Doornbos, Introduction of a robustness coefficient in optimization procedures: Implementation in mixture design problems. Part II: Some practical considerations. *Chemometr. Intell. Lab. Syst*. 10 (1991):325–336.
28. J. H. De Boer, A. K. Smilde, and D. A. Doornbos, Introduction of a robustness coefficient in optimization procedures: Implementation in mixture design problems. Part III: Validation and comparison with competing criteria. *Chemometr. Intell. Lab. Syst*. 15 (1991):13–28.
29. P. F. De Aguiar, Y. Vander Heyden, and D. L. Massart, Study of different criteria for the selection of a rugged optimum in high performance liquid chromatography optimisation. *Anal. Chim. Acta* 348 (1997):223–235.
30. H. Fabre, Robustness testing in liquid chromatography and capillary electrophoresis. *J. Pharm. Biomed. Anal*. 14 (1996):1125–1132.
31. S. Toasaksiri, D. L. Massart, and Y. Vander Heyden, Study of method validation criteria in capillary electrophoresis method for the separation of non-steroidal anti-inflammatory drugs. *Anal. Chim. Acta* 416 (2000):29–42.
32. H. Fabre and N. Mesplet, Robustness testing for a capillary electrophoresis method using the "short-end injection" technique. *J. Chromatogr. A* 897 (2000):329–338.
33. M. Jimidar, W. Van Ael, M. De Smet, and J. Cockaers, Method validation and robustness testing of an enantioselective CE methods for chemical quality control. *LC-GC. Europe* 15 (2001):230–242.
34. K. Maštovská, J. Hajšlová, and S. J. Lehotay, Ruggedness and other performance characteristics of low-pressure gas chromatography-mass spectrometry for the fast analysis of multiple pesticide residues in food crops. *J. Chromatogr. A* 1054 (2004):335–349.
35. Y. Vander Heyden, C. Hartmann, D. L. Massart, P. Nuyten, A. M. Hollands, and P. Schoenmakers, Ruggedness testing of a size-exclusion chromatography assay for low-molecular-mass polymers. *J. Chromatogr. A* 756 (1996):89–106.
36. S. Karmarkar, M. Koberda, J. Momani, D. Kotecki, and R. Garber, Validated ion-exclusion chromatographic method for citrate and acetate in medical fluids. *J. Chromatogr. A* 1039 (2004):147–153.
37. J. Vial and A. Jardy, Utilisation des plans d'expériences pour évaluer la robustesse d'une méthode d'analyse quantitative par chromatographie en phase liquide. *Analusis* 26 (1998):15–24.
38. E. Marengo, M. C. Gennaro, V. Gianotti, and S. Angelino, A test of robustness in IIR-RP-HPLC separation of nine priority pollutant phenols. *J. Liq. Chrom. Rel. Tech*. 24 (2001):341–353.
39. R. Romero, M. Sánchez-Viñas, D. Gázquez, M. G. Bagur, and L. Cuadros-Rodríguez, Robustness study for the determination of biogenic amines by HPLC. *Chromatographia* 53 (2001):481–484.
40. J. A. Day, M. Montes-Bayón, A. P. Vonderheide, and J. A. Caruso, A study of method robustness for arsenic speciation in drinking water samples by anion exchange HPLC-ICP-MS. *Anal. Bioanal. Chem*. 373 (2002):664–668.

41. J. Marcos, A. Rios, and M. Valcárcel, Assessment of analytical quality in automatic flow systems. *Fresenius J. Anal. Chem.* 354 (1996):140–149.
42. L. M. B. C. Álvares-Ribero and A. A. S. C. Machado, Usefulness of a ruggedness test in the validation of flow injection analysis systems. *Anal. Chim. Acta* 355 (1997):195–201.
43. L. Gámiz-Gracia and M. D. Luque de Castro, Development and validation of a flow-inyection method for the determination of albumin tannate, the active component of a pharmaceutical preparation. *J. Pharm. Biom. Anal.* 15 (1997):447–452.
44. L. Cuadros-Rodríguez, R. Blanc-García, A. García-Campaña, and J. M. Bosque-Sendra, A new approach to a complete robustness test of experimental nominal conditions of chemical testing procedures for internal analytical quality assessment. *Chemometr. Intell. Lab. Syst.* 41 (1998):57–68.
45. J. Aybar Muñoz, A. García-Campaña, and L. Cuadros-Rodríguez, Evaluating the significance threshold in robutness testing. A critical discussion on the influence of time in molecular florescence spectrometry. *Talanta* 56 (2002):123–136.
46. J. M. Bosque-Sendra, M. Nechar, and L. Cuadros-Rodríguez, Decision protocol for checking robustness with previous outlier detection in the validation of analytical method. *Fresenius J. Anal. Chem.* 365 (1999):480–488.
47. S. Furlanetto, S. Pinzauti, P. Gratteri, E. La Porta, and G. Calzeroni, Experimental design strategies in the optimization and robustness testing of adsorptive stripping voltammetric conditions for kynurenic acid determination. *J. Pharm. Biomed. Anal.* 15 (1997):1585–1594.
48. M. B. Sanz, L. A. Sarabia, A. Herrero, and M. C. Ortiz, A study of robustness with multivariate calibration. Application to the polarographic determination of benzaldehyde. *Talanta* 56 (2002):1039–1048.
49. M. Lavertu, Z. Xia, A. N. Serreqi, M. Berrada, A. Rodrigues, D. Wang, M. D. Buschmann, and A. Gupta, A validated ^{1}H NMR method for the determination of the degree of deacetylation of chitosan. *J. Pharm. Biomed. Anal.* 32 (2003):1149–1158.
50. F. J. Egea González, M. E. Hernández Torres, A. Garrido Frenich, J. L. Martínez Vidal, and A. M. García Campaña, Internal quality-control and laboratory-management tools for enhancing the stability of results in pesticide multi-residue analytical methods. *Trends Anal. Chem.* 23 (2004):361–369.
51. G. T. Wernimont, Ruggedness evaluation of test procedures. *ASTM Standardization News* 5 (1997):13–16.
52. W. J. Youden and E. H. Steiner, Statistical Manual of the Association of Official Analytical Chemists (Arlington: AOAC, 1975, pp. 33–36, 70–71 and 82–83).
53. M. Mulholland, Ruggedness testing in analytical chemistry. *Trends Anal. Chem.* 7 (1988):383–389.
54. Y. Vander Heyden, M. S. Khots, and D. L. Massart, Three-level screening designs for the optimisation or the ruggedness testing of analytical procedures. *Anal. Chim. Acta* 276 (1993):189–195.
55. Y. Vander Heyden, K. Luypaert, C. Hartmann, D. L. Massart, J. Hoogmartens, and J. De Beer, Ruggedness tests on the high-performance liquid chromatography assay of the United States Pharmacopeia XXII for tetracycline hydrochloride. A comparison of experimental designs and statistical interpretations. *Anal. Chim. Acta* 312 (1995):245–262.
56. A. Nijhuis, H. C. M. Van der Knaap, S. De Jong, and B. G. M. Vandeginste, Strategy for ruggedness tests in chromatographic method validation. *Anal. Chim. Acta* 391 (1999):187–202.
57. Y. Vander Heyden and D. L. Massart, Review of the use of robustness and ruggedness in Analytical Chemistry, in *Robustness of Analytical Methods and Pharmaceutical Technological Products*, A. Smilde, J. de Boer and M. Hendriks, Eds. (Amsterdam: Elsevier, 1996, pp. 79–147).
58. Y. Vander Heyden, A. Nijhuis, J. Smeyers-Verbeke, B. G. M. Vandeginste, and D. L. Massart, Guidance for robustness/ruggedness tests in method validation. *J. Pharm. Biomed. Anal.* 24 (2001):723–753. (Available also in *http://minf.vub.ac.be/∼fabi/validation/robust/index.html*).
59. J. A. van Leeuwen, L. M. C. Buydens, B. G. M. Vandegiste, G. Kateman, P. J. Schoenmakers, and M. Mulholland, RES, an expert system for the set-up and interpretation of a ruggedness test in HPLC method validation. Part 1: The ruggedness test in HPLC method validation. *Chemometr. Intell. Lab. Syst.* 10 (1991):337–347.
60. J. A. van Leeuwen, L. M. C. Buydens, B. G. M. Vandegiste, G. Kateman, P. J. Schoenmakers, and M. Mulholland, RES, an expert system for the set-up and interpretation of a ruggedness test in HPLC method validation. Part 2: The ruggedness expert system. *Chemometr. Intell. Lab. Syst.* 11 (1991):37–55.
61. J. A. van Leeuwen, L. M. C. Buydens, B. G. M. Vandegiste, G. Kateman, A. Cleland, M. Mulholland, C. Jansen, F. A. Maris, P. H. Hoogkamer, and J. H. M. van der Berg, RES, an expert system for the set-up and interpretation of a ruggedness test in HPLC method validation. Part 3: The evaluation. *Chemometr. Intell. Lab. Syst.* 11 (1991):161–174.
62. R. Romero, D. Gázquez, M. Sánchez-Viñas, L. Cuadros-Rodríguez, and M. G. Bagur, A geometric approach to robustness testing in analytical HPLC. *LC-GC North Am.* 20 (2002):72–80.
63. J. A. Van Leeuwen, B. G. M. Vandeginste, G. Kateman, M. Mulholland, and A. Cleland. An expert system for the choice of factor for a ruggedness test in liquid chromatography. *Anal. Chim. Acta* 228 (1990):145–153.
64. Y. Vander Heyden, F. Questier, and D. L. Massart, A ruggedness test strategy for procedure related factors: Experimental set-up and interpretation. *J. Pharm. Biomed. Anal.* 17 (1998):153–168.
65. Y. Vander Heyden, F. Questier, and D. L. Massart, Ruggedness testing of chromatographic methods: Selection of factors and levels. *J. Pharm. Biomed. Anal.* 18 (1998):43–56.
66. EUROCHEM/CITAC Guide, QUAM: 2000. P1, *Quantifying Uncertainty in Analytical Measurements*, 2nd edition, 2000.
67. Y. Vander Heyden, C. Hartmann, D. L. Massart, L. Michel, P. Kiechle, and F. Erni, Ruggedness tests on an HPLC assay: Comparison of tests at two and three levels by using two-level Plackett-Burman designs. *Anal. Chim. Acta* 316 (1995):15–26.
68. Y. Vander Heyden, S. Kuttatharmmakul, J. Smeyers-Verbeke, and D. L. Massart, Supersaturated designs for robustness testing. *Anal. Chem.* 72 (2000):2869–2874.
69. Y. Vander Heyden, M. Jimidar, E. Hund, N. Niemeijer, R. Peeters, J. Smeyers-Verbeke, D. L. Massart, and J. Hoogmartens, Determination of system suitability limits with a robustness test. *J. Chromatogr. A* 845 (1999):145–154.

70. E. Hund, Y. Vander Heyden, M. Haustein, D. L. Massart, and J. Smeyers-Verbeke, Comparison of several criteria to decide on the significance of effects in a robustness test with an asymmetrical factorial design. *Anal. Chim. Acta* 404 (2000):257–271.
71. R. Ragonese, M. Macka, J. Hughes, and P. Petcocz, The use the Box-Behnken experimental design in the optimization and robustness testing of a capillary electrophoresis method for the analysis of ethambutol hydrochloride in a pharmaceutical formulation. *J. Pharm. Biomed. Anal*. 17 (2002):995–1107.
72. Y. Vander Heyden, K. De Braekeleer, Y. Zhu, E. Roets, J. Hoogmartens, J. De Beer, and D. L. Massart, Nested designs in ruggedness testing. *J. Pharm. Biomed. Anal*. 20 (1999):875–887.
73. Y. Vander Heyden, A. Bourgeois, and D. L. Massart, Influence of the sequence of experiments in a ruggedness test when drift occurs. *Anal. Chim. Acta* 347 (1997):369–384.
74. E. Hund, D. L. Massart, and J. Smeyers-Verbeke, Robust regression and outlier detection in the evaluation of robustness tests with different experimental designs. *Anal. Chim. Acta* 463 (2002):53–73.
75. M. Mullholland and J. Waterhouse, Development and evaluation of an automated procedure for the ruggedness testing of chromatographic conditions in high-performance liquid chromatography. *J. Chromatogr*. 395 (1987):539–551.
76. E. Hund, Y. Vander Heyden, D. L. Massart, and J. Smeyers-Verbeke, Derivation of system suitability test limits from a robustness test on an LC assay with complex antibiotic samples. *J. Pharm. Biomed. Anal*. 30 (2002):1197–1206.
77. Y. Vander Heyden, D. L. Massart, Y. Zhu, J. Hoogmartens, and J. De Beer, Ruggedness test on the high performance liquid chromatography assay of the United States Pharmacopeia 23 for tetracycline·HCl: Comparison of different columns in a interlaboratory approach. *J. Pharm. Biomed. Anal*. 14 (1996):1313–1326.
78. M. Jimidar, N. Niemeijer, R. Peeters, and J. Hoogmartens, Robustness testing of a liquid chromatography method for the determination of vorozole and its related compounds in oral tablets. *J. Pharm. Biomed. Anal*. 18 (1998):479–485.
79. ISO, *Guide to the Expression of Uncertainty in Measurement (GUM)*, ISO, Geneva, 1995.