Validation of qualitative analytical methods

Esther Trullols, Itziar Ruisánchez, F. Xavier Rius

This article reviews the state of the art in validating qualitative analytical methods. After introducing the scope of these qualitative methods, their main characteristics and how they differ from quantitative analysis methods, we propose a classification according to the detection system. We discuss the institutions, programs and documents dealing with the validation of qualitative methods, and we present the performance parameters- false positive and negative, sensitivity and specificity rate, cut-off, unreliability region, ruggedness and cross-reactivity. We also briefly describe the various strategies used to validate qualitative analytical methods — Contingency Tables, Bayes' Theorem, Statistical Hypothesis Tests and Performance Characteristic Curves.

Keywords: Binary results; Performance parameters; Qualitative analysis; Validation

Esther Trullols, Itziar Ruisánchez*, F. Xavier Rius
Universitat Rovira i Virgili, Departament de Química Analítica i Química Orgánica, Plaça Imperial Tàrraco 1, E-43005 Tarragona, Spain

*Corresponding author.
Tel.: +34 977 558170;
Fax: +34 977 559563;
E-mail: ruisan@quimica.urv.es

1. Introduction

One of the trends in modern analytical chemistry is the development of new analytical techniques and methods that can reliably identify and quantify the components in complicated samples, such as those related to environmental problems or food protection. Hyphenated techniques, such as the combination of chromatography with mass spectrometry or various spectroscopic techniques, are just some of the examples of these developments. These powerful tools have involved a considerable investment in expensive instruments and require analysts to be properly trained.

However, from a practical point of view, many users find it increasingly important to reconsider whether quantitative results are really necessary. In routine laboratories, e.g., it is quite usual for the first stage to determine whether one or more analytes are present/absent in a sample and, if so, for the second step to estimate their concentration level, e.g., in assessing if a sample of drinking water is free from pollutants. Therefore, instead of trying to quantify the pollutants in the sample as the first goal, it could be enough just to assure if they are present above or below the permitted concentration level. Qualitative methods are used in these cases. They are commonly used as screening techniques before quantification with the routine method, and that enables both time and cost of analysis to be reduced.

The quality of the results provided by these qualitative methods is of the utmost importance. Users of these analytical methods must make sure that the results obtained in their laboratory are fit for their purpose. This means that the analytical requirements must be defined and the values of the performance parameters assessed before they are used as routine methods in the laboratory. In other words, qualitative methods must also be validated [1]. Usually, validation of analytical methods has been developed and applied to quantitative methods. As a consequence, nowadays there are many validation guidelines that are either accepted by regulatory bodies or by communities of practitioners in specific fields. There is, however, no general validation guideline available for qualitative analytical methods.

This review discusses the state of the art of validation in qualitative methods. We try to fill a gap by clarifying the concepts related to qualitative analytical methods. First, we review the various programs provided by the organizations that deal with qualitative method validation, and then we define and discuss some terms. We go on to explain some performance parameters and how they are calculated, and, finally, we describe the strategies used to validate qualitative analytical methods.

0165-9936/$ - see front matter © 2003 Published by Elsevier B.V. doi:10.1016/S0165-9936(04)00201-8

2. Qualitative methods of analysis

The idea of qualitative method is by no means new. It has been defined by the European Community as ''the assessment of the presence or absence of one or more analytes in a sample due to its physical and chemical properties'' [2].

Association of Official Analytical Chemists (AOAC International) defines qualitative methods as a ''method of analysis whose response is either the presence or absence of the analyte, detected either directly or indirectly in a certain amount of a sample'' [3].

It can be concluded from the definitions that a qualitative analytical method is used to find out if a sample contains one or more specific analytes. In these cases, the result of the analysis can be binary only: presence/absence or YES/NO.

As can be easily inferred, presence/absence is not considered to be an absolute measure related to a concentration level of zero but to a specific concentration level. Below this limiting level, the concentration of analyte is considered insignificant. The detection of the analyte may require either an instrument or the human senses, but whatever the way the response is recorded, it is converted into a YES/NO result.

It is well known that quantitative methods make it possible to quantify one or more analytes in a sample by using calibration curves that transform the instrumental response into the measurand, often expressed as the concentration of analyte. Between qualitative and quantitative methods, there is still room for semi-quantitative methods of analysis. These methods provide an approximate response that enables the analyte to be roughly quantified, and they usually assign the test sample to a given class (e.g., the concentration could be high, medium, low or very low). This means that the estimate of the true concentration has a large associated uncertainty. Even so, they are useful because quantification does not always have to be accurate. A representative example would be the test strips for pH measurements. These methods usually cost less than quantitative methodologies, are easier to handle and have other practical performance parameters.

One of the main drawbacks when dealing with qualitative methods is the terminology used because there is no internationally accepted vocabulary, so several names are commonly used in the bibliography. Although terms such as screening systems, test kits, field tests or immunoassays are traditionally used when referring to qualitative methods, they could also be used when dealing with quantitative and semi-quantitative methods. Consequently, here we shall try to put into context the terms that are usually found in the literature.

To start with, it is interesting to consider the term ''screening'' in this regard. In an analytical problem, a screening analysis separates or discriminates samples from a large group that contain, e.g., one or more analytes above or below a pre-set value (Fig. 1). This value is often expressed as a concentration level, and can be set by an official agency, internal quality control or a client, among other possibilities. This pre-set concentration is also called the specification limit, threshold value or maximum permitted level, among other names.

Nowadays, it is quite usual for the term ''screening method'' to be used as a synonym for ''qualitative method'' [4]. However, often the term ''screening'' is also used to describe a step that comes before the calibration stage in a quantitative method. Therefore, screening is not always related to qualitative but sometimes to quantitative analysis [5].

Another similar term is ''screening test'' that gives a reliable indication that the analytes of interest are present/absent in the sample at a level that is hazardous or not permitted [6]. Usually, screening tests are

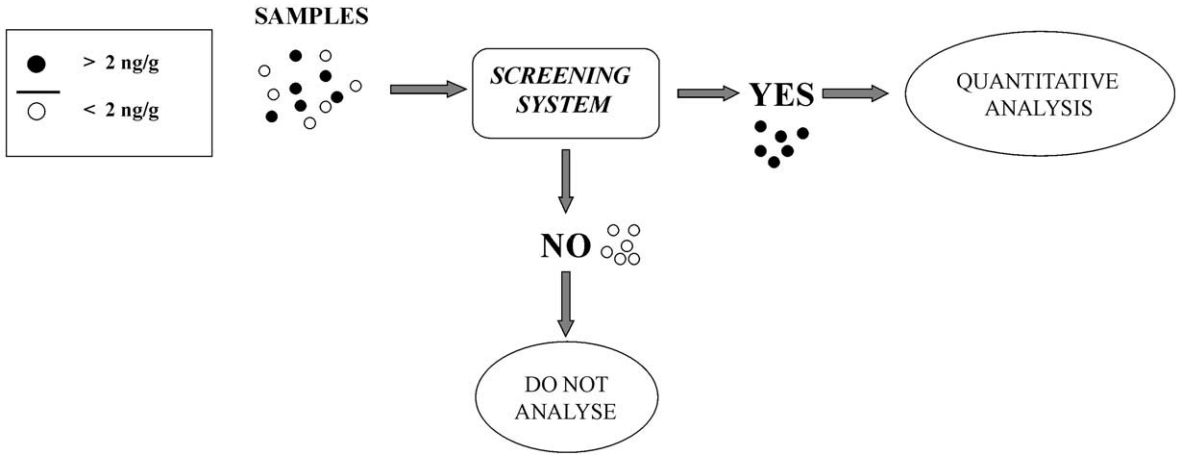


Figure 1. Scheme for a screening system of samples. ● Samples containing more than 2 ng/g of analyte. ○ Samples containing less than 2 ng/g of analyte.

commercially available in a package containing all the reagents and sometimes the instrumentation for the analysis, and they are also known as ''test kits'' [7]. These kits are used for ''rapid and direct analyses'', because they are easy to handle, cheap to purchase and to run, and quick. They also provide results on site.

Another widely used, synonymous term in some fields is ''immunoassay'' [8], an analytical technique that uses an antibody molecule as a binding agent to detect and quantify substances in a sample. Immunoassays have been shown to detect and to quantify many compounds of environmental interest, such as pesticides, industrial chemicals or drug residues, so some specific forms of immunoassay [9] can be considered as quantitative methods. Some of the most important advantages of immunoassays are their rapidity, sensitivity, specificity and cost-effectiveness; they can be designed as rapid, field-portable, qualitative methods or as standard quantitative laboratory procedures; and, they can also be used as screening methods to identify samples that need to be analyzed further by classical analytical methods.

2.1. Classification of qualitative methods

As often happens in many disciplines, there is no generally accepted classification of qualitative methods, although several schemes with a diversity of criteria have been proposed by various authors.

Valcárcel et al. [4] suggest quite a broad classification based on a variety of criteria: the physical state of the sample (i.e., whether it is solid or liquid); the detection system (either sensorial or instrumental); etc. The authors discuss the integration of the chromatographic techniques and the qualitative methods, so the resulting analytical systems can be classified as sensors, systems that use separate laboratory steps or methods that integrate the body of operations.

More intuitive sorting exists, e.g., Unger-Heumann [7] considers test kits as adaptations of well-known analytical methods, so the classification takes into account if test kits are based on chemical, physicochemical, biochemical or biological methods.

Throughout this article, we have classified qualitative methods of analysis according to the type of detection system so as to differentiate between sensorial and instrumental detection.

2.2. Qualitative methods based on sensorial detection

The main feature of these qualitative methods is that human senses are used to record and interpret the response. As might be expected, vision is the sense that is most used (e.g., the response can be a signal, such as a colored solution, a spot on a test strip or the appearance of turbidity). In order to obtain this response, these methods are based on the reaction between the analyte of interest in the sample and specific reagents involved in the procedure. The magnitude of this response can be either directly or indirectly related to the concentration of the analyte. The reaction follows different principles, mainly chemical and immunological. The most commonly used chemical reactions are complexation and precipitation. However, in immunological methods, in particular those of the ELISA (enzyme-linked immunosorbent assay) type, the appearance of the colored spot requires the addition of an enzyme that recognizes the analyte-antibody binding.

In addition to visual inspection, color development can be measured and color intensity related to analyte concentration. One way of doing so is to compare the color to a color card or wheel with a predefined correspondence between color intensity, either in solution [10] or test strip [11], and concentration.

2.3. Qualitative methods based on instrumental detection

These methods provide an instrumental response, which, in many cases, measures absorbance, although in principle any instrument can be used. There are considerable differences between the way instruments are used in qualitative and quantitative analysis. The final decision is made by comparing the response of a test sample and the response of a sample containing the target analyte at the specification level. We call this the reference sample. Instead of working in the concentration domain, these methods work in the response domain. They can also be used to quantify the analyte in the sample, if necessary.

Their basis is that an instrumental response is used to decide whether the analyte is above or below a specific concentration level. No calibration curve is prepared, however; the test-sample response is simply compared to the response provided by the reference sample, so this reference sample, which should ideally be a reference material, is measured and its response (r_{SL}) recorded. Subsequently, the recorded test sample response (r_i) is compared to r_{SL}. If r_i is larger than r_{SL}, it can be concluded that the test sample contains the analyte at a concentration level higher than the reference sample. However, if r_i is lower than r_{SL}, then the conclusion is that the test sample contains less analyte than the reference. Thus the instrumental response is converted into a binary response of the type YES/NO.

Using this procedure, Waters et al. [12] compared the test-sample response with the reference-sample response but did not consider either probabilities of type α or type β errors. These probabilities of error are used by Pulido et al. [13] to calculate the so-called cut-off value, a limiting value in the response domain, at which the decision about whether the analyte is above or below the specific concentration level must be taken.

As in the previous case (sensorial detection), chemical and immunological based reactions are commonly

used. ELISA-based methods can be considered to be special cases because a specific detection tool is sometimes required (e.g., when a 96-microtiter-plate format is used). This tool enables the calibration standards and some samples to be measured simultaneously. Although the calibration curve can be computed, it need not be used if the only thing required is a comparison between the response of the reference sample and the test sample.

3. Method validation in qualitative analysis

As is well known, before any analytical method is applied to test samples on a routine basis, it should be validated, so its performance characteristics should be defined and properly assessed. The ISO/IEC 17025 standard [14] describes the importance of method validation and its application in the analytical laboratory.

There is general agreement about the concept of method validation. The ISO defines method validation as a ''confirmation with an examination and provision of objective evidences that particular requirements for a specified use are met'' [1], so the first thing to be done is to define these particular requirements that depend on the specific determination ahead and are, therefore, particular to each case. This is very much related to the concept of ''fitness for purpose'' [15] and can also be applied to qualitative analysis methods.

The validation of these methods must follow the same philosophy as that of quantitative methods, although there are some differences in the methodology, as described below. In recent years, some organizations have published guidelines or documents about the validation of qualitative analytical methods. The aim of the next section is to give an overall view of the institutions involved in this subject.

3.1. Organizations that deal with qualitative method validation

All organizations that deal with qualitative method validation focus on the concept of fitness for purpose, and therefore on evaluating the relevant performance parameters. Among the different possibilities, the general recommendation is that participation in collaborative studies is the preferred way of validating methods. The strongest exponent of this idea is AOAC International [16]. Like the ''Peer-Verified Methods Program'' for quantitative in-house methods [17,18], AOAC International has the ''Performance Tested Methods Program'' [19] specifically addressing test kits. This validation program makes it possible for the quality parameters claimed by the manufacturer or end user to be assessed by a third laboratory. Similarly, the ''International Seed Testing Association'' (ISTA) [20] has a program called ''Performance Validated Method'' in which a third laboratory proves the quality parameters of the test kits based on immunological reactions.

The US Environmental Protection Agency (EPA) [21] also has a specific document called ''Guidance for Methods Development and Methods Validation for the RCRA Program'' [22]. This ensures that established, validated immunoassays are available for measuring and monitoring needed for the RCRA (Resource Conservation and Recovery Act) Program and it is addressed to developers of qualitative and quantitative methods in general.

In ''The Fitness for Purpose of Analytical Methods'' document [15], EURACHEM specifies that the qualitative performance parameters that should be evaluated are: confirmation of identity; sensitivity; selectivity/specificity; and, precision. Precision may be expressed as true and false positive (and negative) rates and it has to be taken into account that these rates are related to sensitivity and specificity. To avoid problems of nomenclature, the same guide clarifies the meaning of these two parameters in chemical usage. AOAC International also proposes and defines what it calls the four performance indicators: sensitivity, specificity, false negative and false positive rates [3].

Similarly, in its official bulletin [2], the European Union (EU) defines and proposes the evaluation of the following qualitative parameters: limit of detection (CCβ); selectivity/specificity; stability; applicability; and, robustness. The EU also states that screening methods can be used as long as they are properly validated and the percentage of false complaints (probability of β error) is lower than 5% at the concentration level of interest.

Finally, the European Cooperation for Accreditation of Laboratories (EAL) has a guide entitled ''Validation of Test Methods'' [23], which emphasizes that the uncertainty associated with the method is the most important quality parameter. This guide also makes specific reference to qualitative methods that deal with sensorial responses, in the sense that not all known validation procedures are applicable. It has to be clarified that, in this guide, ''test methods'' refers to any analytical method (quantitative and qualitative).

According to the above, the definition of method validation is applicable to both quantitative and qualitative methods of analysis, although there are differences in the validation process. The different meanings of the performance parameters used in qualitative and quantitative methods and the disparity in their definitions require changes in the ways that they are calculated.

3.2. Use of references

References are essential in method validation, as trueness has to be assessed, so, if we try to use references from quantitative analysis in a qualitative method, we

can follow an established hierarchical order. The hierarchy ranges from primary methods to recovery studies, and it includes certified reference materials (CRMs), participation in collaborative studies and the use of confirmatory methods.

Unfortunately, there are considerably fewer possibilities for qualitative analytical methods. For these cases, there is still no primary method. Moreover, CRMs are rather complicated to use. It should be emphasized that any qualitative method claimed to work at the specification level will provide positive and negative results about the test samples. But, as a result of experimental or random error, false rates (either positive or negative) are obtained close to this concentration level, so the CRM should contain the analyte at a concentration level that is near to the specification limit. If the concentration level is either far below or far above the specification limit, we will be able to check only if the method correctly classifies the samples as negative or positive. For CRM concentrations close to this concentration level, we have to compute the probabilities of false positive and false negative responses, so the comparison with a CRM has to be in terms of probabilities, and cannot be in terms of concentration.

As a result, whenever possible, comparison with a reference method is the best option. The analysis must be made using both the reference method (usually quantitative) and the qualitative method [24,25]. To assess whether the qualitative method is performing well, the proportions of positive results obtained by both methods have to be compared by means of a suitable hypothesis test, such as the Chi-square test (χ^2) [3].

Participation in collaborative studies is also recommended. However, as with CRMs, basic statistics, such as mean and standard deviation, cannot be computed. Each laboratory will report its own results (positive and negative test samples). The positive or negative rates can be computed both individually, for each participating laboratory, and globally, for the study as a whole [26,27]. Again the probabilities obtained by each laboratory can be compared by means of the Chi-square test. If any these possibilities is impracticable, spiked samples can be used as a first approximation for the validation process.

3.3. Qualitative performance parameters

The definition of the performance parameters is an important aspect to consider when dealing with qualitative analysis. Table 1 shows some of the most common parameters, according to whether the type of analytical method chosen is quantitative or qualitative.

Although some performance parameters have the same name, the concepts attached to them and their evaluation can be different, e.g., sensitivity can be differently considered depending on the analytical method. If a quantitative method is used, sensitivity should be a numerical value that indicates how the response changes whenever there is a variation in the concentration of the analyte. However, this parameter will be evaluated in a different way if a qualitative method is used. The same occurs with the specificity, detection limit, cut-off value and uncertainty or unreliability region.

The following parameters have to be considered when dealing with qualitative responses.

3.3.1. False positive and negative rates. The false positive rate is ''the probability that a test sample is a known negative, given that the test sample has been classified as positive by the method'' [3].

$$False\ \ positive\ \ rate = \frac{fp}{tn + fp} \tag{1}$$

where fp are false positive test samples and tn are known true negative test samples.

Similarly, the false negative rate is ''the probability that a test sample is a known positive, given that the test sample has been classified as negative by the method'' [3].

$$False\ \ negative\ \ rate = \frac{fn}{tp + fn} \tag{2}$$

where fn are false negatives samples and tp known true positive test samples.

3.3.2. Sensitivity and specificity. Generally speaking, when dealing with qualitative methods, sensitivity is ''the ability of a method to detect truly positive samples as positive'' [6], so the sensitivity rate ''is the probability, for a given concentration, that the method will classify the test sample as positive, given that the test sample is a 'known' positive'' [28]. It can be calculated as:

$$Sensitivity\ rate = \frac{test\ positives}{total\ number\ of\ known\ positives} = \frac{tp}{tp + fn} \tag{3}$$

Table 1. Quality parameters for both quantitative and qualitative analysis methods

Quantitative method	Qualitative method
Accuracy: trueness, precision	False positive and negative rates
Uncertainty	Sensitivity and specifity
Sensitivity and specifity	Selectivity: interferences
Selectivity: interferences	Detection limit
Range and linearity	Cut-off limit
Detection limit	Unreliability region
Ruggedness or robustness	Ruggedness or robustness

where *tp* are truly positive test samples and *fn* are false negative test samples.

The same occurs with specificity, which is defined as ''the ability of a method to detect truly negative samples as negative'' [6]. In the same way, the specificity rate ''is the probability, for a given concentration, that the method will classify the test sample as negative, given that the test sample is a 'known' negative'' [28], so it can be expressed as:

$$Specificity\ rate = \frac{test\ negatives}{total\ number\ of\ known\ negatives} = \frac{tn}{tn+fp} \quad (4)$$

where *tn* are truly negative test samples and *fp* are false positive test samples.

3.3.3. Unreliability region. In quantitative analysis, the uncertainty is the numerical value related to the interval in which the measurand may be found with a given probability. However, for qualitative methods, having binary responses of the YES/NO type, there is no meaning for a number associated with the result and expressed as a semi-interval that is attached to it, so uncertainty is expressed not as a numerical value but as a region of probabilities of committing error. Moreover, following the nomenclature used until now, it corresponds to the region in which false responses are obtained (either false positive or false negative).

As we are dealing with a region where there are certain probabilities of error, some authors prefer to call it an unreliability region rather than an uncertainty region [29]. This region is defined by an upper and a lower concentration limit [30], between which the qualitative method can provide false responses. As these false responses can be either positive or negative, the upper and lower limits that define this unreliability region depend on the probability of obtaining these false responses, which is fixed by the analyst.

3.3.4. Detection limit and cut-off value. The term detection limit was defined by the IUPAC [31] in 1995 for quantitative analysis. According to this definition, it can be calculated when the response is a numerical value and when a value is assigned to the two probabilities of α- and β-type errors. When the response is of the binary-sensorial type, however, the standard deviation of the blank samples cannot be calculated, and the probabilities of α- and β-type errors cannot be considered at the same time, although they are both set by the analyst. Depending on the interest of the analyst and the problem in hand, either the probability of committing an α-type error or that of committing a β-type error will be considered.

The detection limit has also been defined as ''the lowest concentration of the analyte which the test can reliably detect as positive in the given matrix'' [6]. This implies that we should consider only the probability of a β-type error or false negative rate, usually at 5%. This definition is presented in the context of assessing a maximum permitted concentration level, but, if it is extrapolated to the case of assessing a minimum concentration level, we should consider only the probability of an α-type error or false positive rate, also at 5%. Therefore, both probabilities of committing error cannot be considered simultaneously. In the first case, the limit of detection coincides with the upper limit of the unreliability region, where the sensitivity rate is 95% and it also coincides with the cut-off value. However, in the second case, the limit of detection coincides with the lower limit of the unreliability region.

The cut-off value is a special performance parameter, since it has been widely studied and used in qualitative analytical methods that use instrumental responses [13]. Regarding the qualitative methods with sensorial responses, this value means the concentration level where the qualitative method differentiates the samples with a certain probability of error, usually 5%. In the particular case of problems related to the maximum permitted level, the cut-off value is related to the sensitivity, as it corresponds to the concentration level at which the sensitivity rate is 95%, when the β-type error probability has been set at 5%.

Other parameters should also be considered. Ruggedness is an important parameter related to how the method performs under variations in the operational, environmental, etc., conditions. In quantitative methods, it must be evaluated [2,15], but, in qualitative methods, it need not be. According to some authors [3], it is not a ''formal part of the validation protocol'' and ''it is not a submission requirement'' when submitting a method for evaluation.

Another parameter to be considered is cross-reactivity or the presence of interferences. For test kits, in particular, it is recommended to check whether the presence of analytes of the same family as the one under study might modify the result of the analysis. These checks are mandatory for manufacturers of the test kits.

3.4. Evaluation of the qualitative performance parameters

There are various ways of evaluating the performance parameters in qualitative analysis. Recently, Pulido et al. [32] showed that Contingency Tables [33], Bayes' Theorem [34], Statistical Hypothesis Tests [13] and Performance Characteristic Curves [35] are the four main ones, each of which has advantages and drawbacks. However, depending on whether or not the type of response obtained is instrumental and on the number of analyses that the analyst wants to

perform, etc., we will have to choose one methodology or another.

3.4.1. Contingency Tables. Contingency Tables have been widely used in bioassays [36,37]. They are based on the calculation of probability. Although other formats are possible, the simplest and most commonly used are those that give a two-category classification: positive or negative, above or below a regulatory concentration level, etc. Then, the qualitative method result is compared with the results obtained using the confirmatory method (see Fig. 2). From this table, it is possible to calculate only four performance parameters (false positive, false negative, sensitivity and specificity rates) and two predictive values (positive, PPV, and negative, NPV).

One of the main features of this approach is that it gives an overall vision of how the qualitative method performs, but it does not give individual information, as a probability of error for each sample is not computed. This means that it is assumed that the unknown sample has the same statistical behavior as the samples used to build the Contingency Table. One of the drawbacks is that the capacity of the Contingency Table depends on the total number of analyzed samples used to build it and the experimental design. It should also be pointed out that all samples must be analysed using both the qualitative and confirmatory methods.

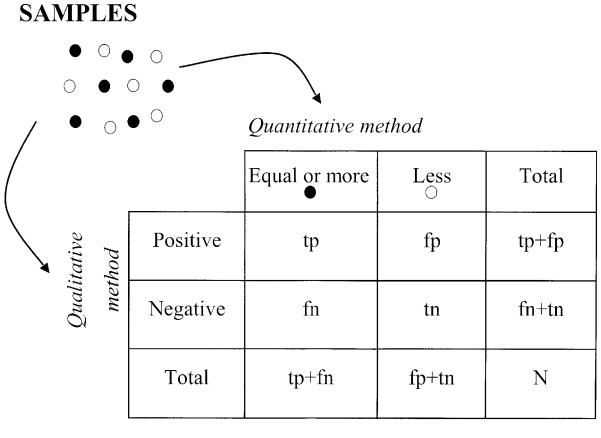
3.4.2. Bayes' Theorem. This methodology is based on the well-known Bayes' Theory of Probability. Several intermediate probabilities must be computed and evaluated. Bayes' Theorem calculates the probability of giving a correct result (either positive or negative) when it is indeed correct, P(a/p). This probability is called the conditional probability, so many analyses are required in order to achieve a good uncertainty estimation or a better error probability. The main feature of this methodology is that, unlike Contingency Tables, the probability of giving a wrong result is estimated individually, because the conditional probability is calculated for each analysed sample. And, again, only the same four parameters can be calculated: false positive, false negative, sensitivity and specificity rates.

3.4.3. Statistical Hypothesis Tests. These Hypothesis Tests compare the response of the sample with that of a pre-set reference [13] (Fig. 3). As was said above, this reference sample contains the analyte at a specific concentration level.

The main advantages of these Hypothesis Tests derive from the use of the well-known probability of an α-type error (the probability of committing false positives) and the increasingly used probability of a β-type error (the probability of committing false negatives). This method makes it easy to evaluate uncertainty when using qualitative methods that provide an instrumental response. Traceability can also be verified and the detection limit computed. However, if the test kit does not provide an instrumental response, or if the response is based on a visual observation that cannot be quantified, Hypothesis Tests cannot be used.

3.4.4. Performance Characteristic Curves. Performance Characteristic Curves are a plot of the probability of having a positive result versus the concentration level of the analyte. The result is a sigmoidal type of curve the slope and the amplitude of which are particular for each qualitative method (see Fig. 4).

The main advantage is that considerable information is provided. In addition to false positive and false negative rates, these Curves make it possible to calculate sensitivity and specificity rates and other performance characteristics of qualitative methods, such as the



	Equal or more ●	Less ○	Total
Positive	tp	fp	tp+fp
Negative	fn	tn	fn+tn
Total	tp+fn	fp+tn	N

Figure 2. Example of a 2×2 Contingency Table. *tp*, true positive samples; *fp*, false positive samples; *fn*, false negative samples; *tn*, true negative samples.

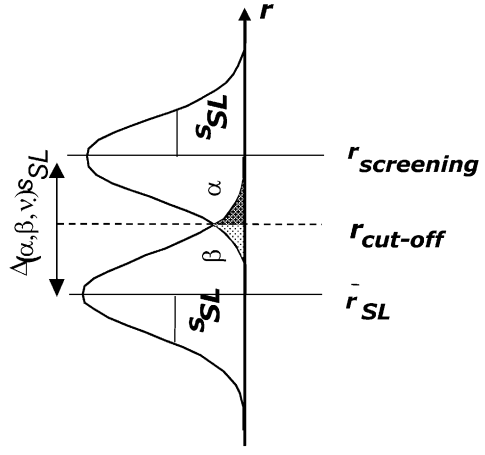


Figure 3. Statistical Hypothesis Tests for qualitative analysis methods providing instrumental response.

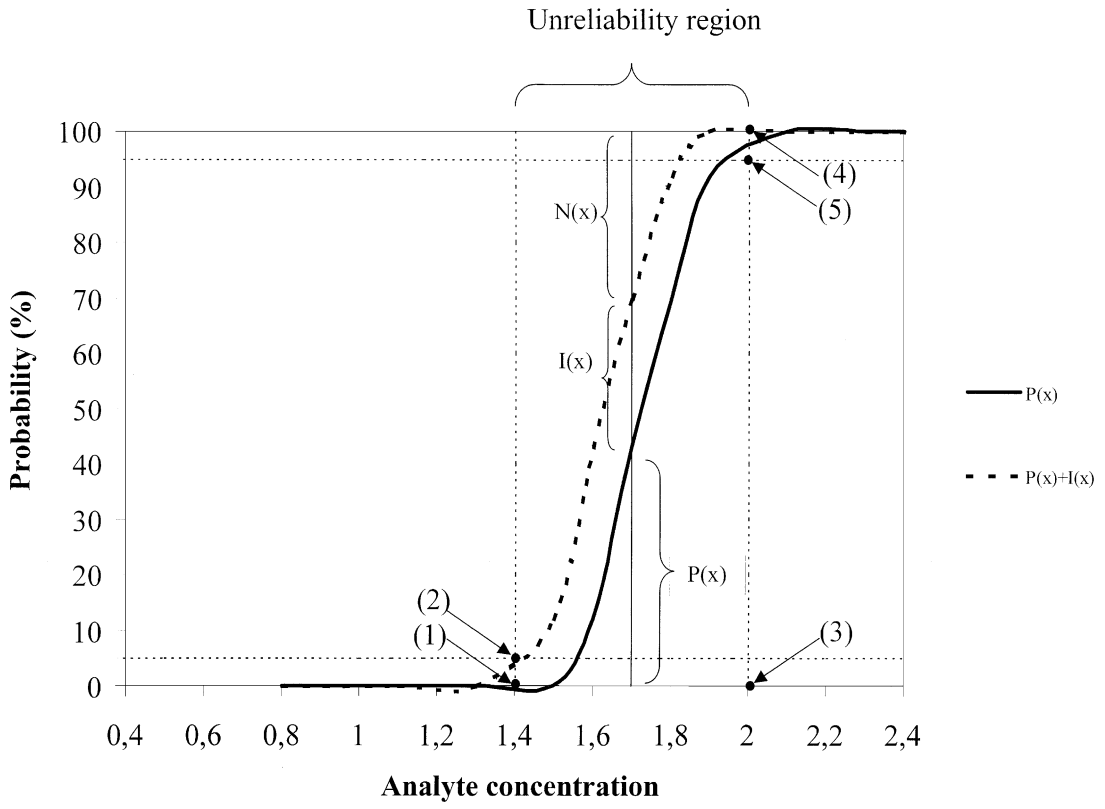


Figure 4. Performance Characteristic Curve. Probability of positive responses, P(x), and probability of positive plus inconclusive responses, P(x) + I(x), were plotted versus concentration levels tested. (1) FP = P(x); (2) $X_{0,05}$, where specificity = N(x) = 100−(P(x) + I(x)); (3) $X_{0,95}$, Cut-off limit, detection limit; (4) FN = 100−(P(x) + I(x)); (5) Sensitivity = P(x) = 100−β.

detection limit and the cut-off limit or the unreliability region. The main drawback is that it is necessary to perform several analyses for each concentration level.

4. Conclusion

Demand for qualitative analysis methods is increasing and they are becoming more and more important. However, some aspects still need to be developed and clarified. For users, one of the most confusing is the nomenclature used to refer to qualitative analysis, since there are many different terms that often have different meanings. Similar confusion occurs with the classification of qualitative methods, where there are several possibilities, according to different authors. Although this may be of no practical importance for many users, some work should be done to structure the criteria for classification.

Validation of qualitative analysis methods is an important issue to consider so as to provide confidence to the analysts. Although several organizations are working on this task, very few of them have defined validation protocols and their own validation programs for method developers. It has to be said that there is still confusion regarding how this validation process should be performed generally. Performance parameters are quite well defined, but, even so, a way of evaluating them has yet to be established. In this article, we have briefly described some possibilities. As far as the use of references in qualitative analysis methods is concerned, the possibilities are considerably fewer than for quantitative analysis methods. Consequently, the references available should be examined more intensively.

Acknowledgements

The authors acknowledge the economic support from the MCyT (Project No. BQU2000-1256).

References

[1] International Organization for Standardization, ISO 8402, Quality management and quality assurance. Vocabulary, ISO, Geneva, Switzerland, 1994.
[2] /657/EC: Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results.
[3] P. Feldsine, C. Abeyta, W. Andrews, J. AOAC Int. 85 (2002) 1187.

[4] M. Valcárcel, S. Cárdenas, M. Gallego, Trends Anal. Chem. 18 (1999) 685.
[5] A. Sanz-Medel, B. San Vicente de la Riva, J.M. Costa-Fernández, R. Pereiro, Anal. Chim. Acta 451 (2002) 203.
[6] J.J. O'Rangers, R.J. Condon, in: J.F. Kay, J.D. MacNeil, J.J. O'Rangers (Editors), Current Issues in Regulatory Chemistry, AOAC Int., Gaithersburg, Maryland, USA, 2000, p. 207.
[7] M. Unger-Heumann, Fresenius' J. Anal. Chem. 354 (1996) 803.
[8] D. Barceló, M.-C. Hennion, Anal. Chim. Acta 362 (1998) 3.
[9] D. Barceló, A. Oubiña, J.S. Salau, S. Pérez, Anal. Chim. Acta 376 (1998) 49.
[10] Merck Farma y Quæmica, S.A (http://www.merck.es).
[11] C. Heiss, M.G. Weller, R. Niessner, Anal. Chim. Acta 396 (1999) 309.
[12] L.C. Waters, R.R. Smith, J.H. Stewart, A. Jenkins, J. AOAC Int. 77 (1994) 1664.
[13] A. Pulido, I. Ruisánchez, R. Boqué, F. Xavier Rius, Anal. Chim. Acta 455 (2002) 267.
[14] International Organization for Standardization, ISO/IEC 17025, General requirements for the competence of testing and calibration laboratories, ISO, Geneva, Switzerland, 1999.
[15] EURACHEM, The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics, EURACHEM Secretariat, Teddington, Middlesex, UK, 1998 (http://www.eurachem.ul.pt).
[16] AOAC International, The cornerstone for online analytical methods (http://www.aoac.org).
[17] AOAC International, Peer Verified Methods Program — Manual on policies and procedures, AOAC Int., Gaithersburg, Maryland, USA.
[18] AOAC International, Method Validation Programs (http://www.aoac.org/vmeth/page1.htm).
[19] AOAC International, Rapid Test Kits/Performance Tested Methods (http://www.aoac.org/testkits/perftestedmtd.html).
[20] International Seed Testing Association (http://www.seedtest. org).
[21] Environmental Protection Agency (http://www.epa.gov/).
[22] Test Methods: Methods Development and Approval Process (http://www.epa.gov/epaoswer/hazwaste/test/methdev.htm).
[23] EA (European co-operation for Accreditation) (http://www.european-accreditation.org).
[24] D. Barceló, B. Ballesteros, A. Dankwardt, P. Schneider, M.P. Marco, Anal. Chim. Acta 475 (2003) 105.
[25] R.W. Sheets, Sci. Total Env. 219 (1998) 13.
[26] F.D. McClure, J. Assoc. Off. Anal. Chem. 73 (1990) 953.
[27] S. De Saeger, L. Sobanda, A. Desmet, C. Van Peteghem, Int. J. Food Microbiol. 75 (2002) 135.
[28] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Data Handling in Science and Technology 20A, Handbook of Chemometrics and Qualimetrics: Part A, Elsevier Science, Amsterdam, The Netherlands, 1997, p. 436.
[29] A. Ríos, D. Barceló, L. Buydens, S. Cárdenas, K. Heydorn, B. Karlberg, K. Klemm, B. Lendl, B. Milman, B. Neidhardt, R.W. Stephany, A. Townshend, A. Zschunke, M. Valcárcel, Accred. Qual. Assur. 8 (2003) 68.
[30] E. Trullols, I. Ruisánchez, F.X. Rius, J. AOAC Int. (accepted for publication).
[31] L. Currie, Pure Appl. Chem. 67 (1995) 1699.
[32] A. Pulido, I. Ruisánchez, R. Boqué, F.X. Rius, Trends Anal. Chem. 22 (2003) 647.
[33] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Data Handling in Science and Technology 20A. Handbook of Chemometrics and Qualimetrics: Part A, Elsevier Science, Amsterdam, The Netherlands, 1997, p. 475.
[34] R.M. McFall, T.A. Treat, Ann. Rev. Psychol. 50 (1999) 215.
[35] R. Song, P.C. Schlecht, K. Ashley, J. Hazard. Mater. 83 (2001) 29.
[36] B.J. Neil, E. Keeler, S.J. Adelstein, New Engl. J. Med. 293 (1975) 267.
[37] N.E. Hawass, Brit. J. Radiol. 70 (1997) 360.