ELSEVIER

Journal of Chromatography A, 921 (2001) 121–133

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Initial evaluation of quantitative performance of chromatographic methods using replicates at multiple concentrations

Marcela Alejandra Castillo[a,b], Reynaldo César Castells[a,b,*]

[a]*División de Química Analítica, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, 47 esq. 115, 1900 La Plata, Argentina*

[b]*CIDEPINT, 52e/121y 122, 1900 La Plata, Argentina*

Received 12 December 2000; received in revised form 23 February 2001; accepted 12 April 2001

Abstract

The introduction of a novel analytical method must be supported by consistent information about its quantitative potentialities; this is critical for whoever considers its utilization for an specific application. Unfortunately, literature abounds in papers proposing excellent chromatographic methods of analysis that have been subjected to comparatively poor quantitative evaluation. The methodology proposed in the present work makes use of some of the performance characteristics whose measurement is recommended in validation protocols; pertinent to this stage of method development are the detection and quantitation limits, the linear range and the repeatability. All this information can be calculated from the results of a calibration with several replicates at each analyte level. Replicates enable the calculation of reproducibility at several analyte levels and the estimation of the linear range; more important, replicates are necessary to detect changes in peak area standard deviation with analyte amount. Regression of calibration data by means of unweighted least-squares (ULSR) can only be performed in those cases in which homoscedasticity has been previously verified; heteroscedastic calibration data demand regression by means of weighted least-squares (WLSR), since ULSR results in gross overestimation of prediction limits at low analyte concentration. The proposal is used for the preliminary quantitative evaluation of a method for the determination of nine biogenic amines by means of pre-column derivatization with dabsyl chloride and separation of derivatives by RPLC. Limits of detection are calculated by a regression approach and by the classical signal-to-noise ratio method (S/N approach). No significant difference was detected for the amines limits of detection estimated by WLSR and by the S/N approach; ULSR estimated limits of detection are between 7 and 78 times larger than those obtained by the other two methods, as a consequence of the heteroscedasticity of calibration data.

Keywords: Performance characteristics; Repeatability; Linearity; Limit of detection; Limit of quantitation; Biogenic amines

1. Introduction

The development of a novel analytical method may be undertaken to solve a specific problem, as the determination of one or several analytes in a given type of sample; in other circumstances the objective is a method that after suitable adaptations might be used in the analysis of different materials. In any case the development work is hardly initiated by working on real samples; measurements during the initial stages are usually performed on solutions of well characterized substances, on analyte stan-

*Corresponding author. Fax: +54-221-427-1537.

E-mail address: rcastell@dalton.quimica.unlp.edu.ar (R.C. Castells).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)00867-6

dards if available. In chromatographic analysis, for instance, conditions (column packing and dimensions, mobile phase, isocratic or gradient operation, etc.) to separate in a reasonable time all the analytes among themselves and from other substances whose presence can be expected in real samples is usually the first aspect to be defined; although there are well known strategies and even commercial software to attain this objective [1,2], considerable experimental work must usually be invested to optimize the separation. When analytes are not detectable by available means or when a more selective detection is desirable for complex samples, a derivatization stage is often included [3,4]; conditions maximizing the yield of the derivatization process must then be studied. This study necessarily precedes the optimization of derivatives separation, although some degree of superposition is usually unavoidable.

Once derivatization and separation conditions have been defined the method is subjected to a preliminary quantitative evaluation with a twofold objective. First, to check that the method fits to analyte levels predictable for the type of materials motivating the study, bearing in mind that the performance observed in the work with standards very probably constitutes the most that can be expected from it. In the second place because the information gathered at this stage is important to detect matrix effects at the time of whole method validation. It is debatable whether a quantitative evaluation, performed at this stage of the development, can be included within definitions of validation. Massart et al. ([5], Chapter 13) although not rejecting the possibility of validating only the actual determination, state that the first golden rule in validation is ''validate the whole method''; this includes pre-measurement operations, matrices and range of analyte concentration. EURACHEM [6] stresses that the limits between method development and validation are diffuse and that many method performance parameters associated with validation are often evaluated, at least roughly, as part of method development. To avoid confusions, the term validation shall not be employed in this paper; this is not an obstacle for using validation performance criteria and for following (with some restrictions) validation measurement guidelines.

A method for the determination of the nine biogenic amines more frequently occurring in foods and beverages is under study in our laboratory; it consists in derivatizing the amines with dabsyl chloride (4-dimethylaminoazobenzene-4′-sulfonyl chloride), separation of derivatives by RPLC and detection by means of UV spectrophotometry. Derivatization and separation, as well as some initial tests on meat samples, are treated in a separate paper; experimental procedures and calculations pertinent to the preliminary evaluation of the quantitative performance of the method [7] are the subject of the present paper.

2. Performance characteristics to be evaluated and overview of measurement methods

Not all the performance characteristics measured in validation protocols are pertinent at the present preliminary evaluation study. Those considered as indispensable in this work are treated subsequently.

2.1. Precision

Repeatability is the only relevant precision criterion at this preliminary evaluation; it expresses the closeness of agreement between a series of independent measurements performed by a single analyst on a given equipment over a short interval of time, and represents the first or most basic level of precision evaluation [6,8]. Other precision criteria, as *reproducibility* (precision between laboratories) and *intermediate precision* (different analysts and equipment, over extended time scale, within a given laboratory), are reserved for whole validation protocols.

Since precision is often concentration dependent, repeatability should be measured at several analyte levels; ICH [9] suggests a minimum of nine independent measurements covering the procedure range. Results are reported as standard deviation, as relative standard deviation (RSD) or as percentage standard deviation (%RSD). Derivatization implies pre-chromatographic manipulations that unavoidably result in precision losses; some estimations [1] raise %RSD from below 1–2% for methods involving only elemental pre-chromatographic operations (sample weighing and dissolving) to 5–10% for methods including derivatization steps, and even higher at trace levels.

2.2. Linearity and linear range

Linearity is the ability of a method to elicit responses that are proportional to analyte amount; the linearity of the calibration line (graphical representation of peak areas as a function of measurement standard [6] amount) needs to be checked. The *linear range* is the interval of analyte amount over which the method behaves linearly. The *quantitation limit* defines the lower end of the linear range; the upper end is usually imposed by instrumental factors, as wavelength bandwidth and stray radiation in UV detection [10]. Derivatization can considerably shorten the linear range at its upper end, for instance when analytes or their derivatives enter into non-linear regions of distribution processes (extraction, adsorption) or when excess reagent concentration is limited by its solubility in the reaction medium. The linear range is estimated from the regression analysis of the calibration plot. Correlation coefficient is misleading in testing linearity [11]; visual inspection of calibration line and residuals can be illustrative, but more objective tests are available, as the F-test for lack-of-fit or the test of significance of the quadratic regression coefficient obtained on fitting calibration data to a second degree polynomial ([5], Chapters 8 and 13). An iterative procedure making use of prediction limits was recently proposed [12]. Linearity is not imperative, but it is highly desirable; therefore, if non-linearity is detected by means of any of these tests, the working range is usually reduced at the upper end and the test is repeated through the remaining points.

2.3. Lowest limits of the method

The *detection limit*, L_D, is the lowest amount of analyte in a sample that can be detected with a specified degree of certainty, but not necessarily quantified. The *quantitation limit*, L_Q, is the lowest amount of analyte in a sample that can be quantitatively determined with an acceptable level of precision and accuracy.

IUPAC rigorous definition of L_D [13,14] is based on the statistical theory of hypothesis testing. L_D depends on blank signal (μ_{bl}), on response dispersion at the blank level (σ_{bl}) and on the definition of two risks: risk α of detecting analyte when it is absent and risk β of not detecting analyte when it is present; illustrations on the relationships between these parameters can be consulted in [5], Chapter 13. Assuming that measurements distribute normally, that σ at L_D is equal to σ_{bl} and defining $\alpha = \beta = 0.05$, method response at the detection limit, Y_D, shall be given by [13,14]:

$$Y_D = \mu_{bl} + 3.29\,\sigma_{bl} = \mu_{bl} + b_1 L_D$$
$$\therefore L_D = 3.29\,(\sigma_{bl}/b_1) \quad (1)$$

where b_1 is the slope of the calibration line. On the same lines, method response at the quantitation limit, Y_Q, is given by:

$$Y_Q = \mu_{bl} + k_Q \sigma_{bl} = \mu_{bl} + b_1 L_Q$$
$$\therefore L_Q = k_Q(\sigma_{bl}/b_1) \quad (2)$$

IUPAC [13] proposes a default value $k_Q = 10$, thus making $\sigma_{bl}/(Y_Q - \mu_{bl}) = 1/k_Q = 0.10$ for the RSD at the L_Q. If the variance of the response changes with analyte amount, Eqs. (1) and (2) are substituted by:

$$Y_D = \mu_{bl} + 1.65\,(\sigma_{bl} + \sigma_{LD}) = \mu_{bl} + b_1 L_D$$
$$\therefore L_D = 1.65\,(\sigma_{bl} + \sigma_{LD})/b_1 \quad (3)$$

$$Y_Q = \mu_{bl} + k_Q \sigma_{LQ} = \mu_{bl} + b_1 L_Q$$
$$\therefore L_Q = k_Q(\sigma_{LQ}/b_1) \quad (4)$$

The standard deviation of blanks cannot be measured in chromatography; alternative methods must be applied and ICH recommendations [9] include two of them:

(a) Based on signal-to-noise ratio: base line peak-to-peak noise, h_N, is measured on the chromatogram of a sample blank processed by the analytical method for a specified interval of time before and after the analyte retention time, and compared to analyte peak heights, $h(X)$, measured under the same conditions from samples containing known low analyte amounts, X. L_D and L_Q are then obtained as $L = k\,h_N\,X/h(X)$, where $k = 2$ or 3 for L_D and $k = 10$ for L_Q. This is the most popular method, and has been in use from the early years of chromatographic analysis; it stems from analysts' experience and common sense and lacks any formal relationship to statistical theory.

(b) Based on the calibration curve: either the residual standard deviation of a regression line, s_e, or the standard deviation of its intercept, s (b_0), are used as estimators of σ_{bl} in Eq. (1) or Eq. (2). Miller and Miller [15] made a similar proposal in 1988; their suggestion, however, was the use of the intercept, b_0, and s_e as estimators of μ_{bl} and σ_{bl}, respectively. When the precision of the measurements is independent of concentration (homoscedasticity) and the model is correct (no lack-of-fit detected), s_e is an estimate of the standard deviation of the measurements ([5], Chapter 8). On the other hand s (b_0) can differ significantly from s_e, and its value depends on the distribution of concentration values within the calibration range. When the precision of the measurements changes with analyte concentration (heteroscedasticity) regression must be performed by means of weighted least squares regression (WLSR), specially when we are interested in the low concentration region; the residual standard deviation calculated with this statistical tool, s_{ew}, is a weighted average for the calibration range of the standard deviation of the measurements, and cannot be used as an estimate of σ_{bl}. Neglect of these basic statistical principles led recently [16] to suggest as a global proposal the use of s_{ew} within ICH method to estimate L_D; neither ICH nor Miller and Miller mention WLSR in connection with detection limits.

Several years ago Hubaux and Vos [17] demonstrated that IUPAC prescriptions about detection limit can be implemented by means of a calibration plot. Besides the calibration line, their approach makes use of the prediction limits, i.e. the lines that define the region where as yet not measured signals shall fall with a probability of (1-α-β): 100α% of the points are predicted to fall above the higher limit and 100β% of the points below the lower limit. A calibration graph is sketched in Fig. 1; two important points can be located on it by means of the following relations, in which the abbreviation *Pr* is used for probability:

$$Pr\,[(Y > \hat{Y}_C) \text{ when } X = 0] = \alpha \quad (5)$$

$$Pr\,[(Y \leq \hat{Y}_C) \text{ when } X = L_C] = 0.5 \quad (6)$$

$$Pr\,[(Y \leq \hat{Y}_C) \text{ when } X = L_D] = \beta \quad (7)$$

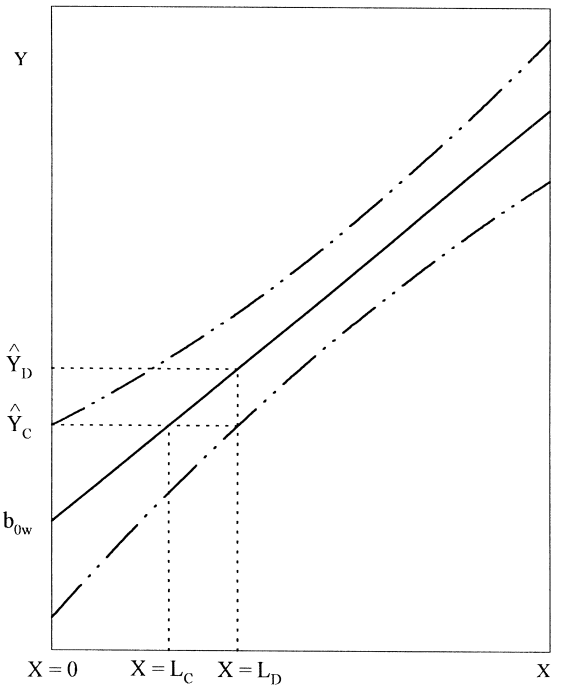

Fig. 1. Sketch of a linear calibration line (—) with its upper and lower prediction limits (— ·· —). L_C is the critical level and L_D is the detection limit.

Relations (5) and (7) are used in a recent IUPAC document [13] to define the *critical level*, L_C, and the detection limit, L_D. $\widehat{Y}_C$ was called *decision limit* by Currie [18], since it represents the signal above which the presence of analyte can be reliably decided. $\widehat{Y}_C$ is the signal predicted by the regression equation for $X = L_C$; however, 50% of the signals measured at L_C shall fall below $\widehat{Y}_C$. In other words, measuring at L_C the probability of a false positive is low but the chances of a false negative are high; it is not an adequate criteria to express the lowest limits of a method, but it is essential to define L_D. The regression approach has been recommended by AOAC [19], by IUPAC [20] and by the German Normative Institute [21].

Hubaux and Vos made use of unweighted least squares regression, ULSR, in their presentation of the regression approach. A basic hypothesis of ULSR is that the variance of the error distribution of signals about their expectation remains constant [22]; in terms of calibration, this requirement means that the signal variance must be independent of analyte

amount. Numerous cases of heteroscedastic calibration data reported by analysts working with different instrumental techniques have been cited by Oppenheimer et al. [23] and by Zorn et al. [24]. Garden et al. [25] were categorical with respect to this subject: ''It is reasonable to expect that much analytical data will not show constant variance nor would we expect the variance to be a simple function of concentration''. When ULSR is performed on heteroscedastic data effects of the larger signal variability at high analyte levels are specially deleterious at low analyte levels; prediction bands are artificially widened at this region and lower limits calculated from them do not reflect the assay capabilities.

Two main options are available to deal with heteroscedastic data. One is data transformation, for instance changing to a log Y versus log X relationship; variance variability can be drastically reduced, but data linearity may be seriously affected. The second option is WLSR, that differs from ULSR in that a statistical weight w_i is assigned to responses Y_i; usually the inverse variance $(1/s_i^2)$ at each level is used as weight. This signifies that WLSR demands a larger number of replicates than ULSR; estimates about the minimum number of replicates by different authors range from six to twenty. This level of replicate is frequently difficult to attain in practice because of different motives, as costs or availability of calibration standards and reagents, or time demanded by previous operations or by the chromatographic run. References to the determination of L_D by means of WLSR are scarce [23,24,26].

In order to estimate the quantitation limit, signal standard deviation is measured at different analyte levels, signal RSD values are calculated at each level and L_Q is defined as the analyte amount for which RSD reaches a pre-established value, for instance RSD = 0.10. This procedure has been recommended by EURACHEM [6], and can be used both with homo- and heteroscedastic data. It can be implemented graphically, for instance by plotting RSD against X.

All the performance characteristics that were considered in the present paper as pertinent to fulfil a preliminary quantitative evaluation can be calculated from data measured in a calibration with several replicates at each concentration. A drawback of this approach is that measurements must be distributed over the (presumed) whole method range, since linear range is one important criteria to be estimated while distributing the measurements in the low concentration region is advisable in estimation of L_D or L_Q. As mentioned, present objectives are different from those prevailing in validation and is taken for granted that the whole procedure is going to be validated on real samples.

3. Experimental

Dabsyl chloride and amine hydrochlorides were purchased from Aldrich (St. Louis, MO, USA). HPLC-grade acetonitrile, methanol, acetone and benzene were from E.M. Science (Gibbstown, USA). Sodium hydrogen carbonate, hydrochloric acid and sodium hydroxide were from Merck Química Argentina (Buenos Aires). Distilled water was purified in a Milli-Q System (Millipore Co., Bedford, MA, USA). Nylon filtration membranes were from Micron Separations (Westborough, MA, USA).

Amines name, abbreviation and purity are given in Table 1. Two series of standard solutions, series S1 (Phe, Ca, Tyr and Spd) and series S2 (Try, Pu, Hist, Ser and Sp), with seven concentrations each were prepared in 0.1 *M* hydrochloric acid. Two stock solutions were prepared by dissolving carefully weighed solid standards of each series in 0.1 *M* HCl; the remaining standards in each series were prepared by diluting both stocks by careful weighing of solutions. Highest and lowest concentrations are listed in Table 1.

Details on derivatization and separation procedures shall be shortly reported in a separate paper. For the present purposes it is sufficient to mention that 50 μl of each standard solution were mixed with 450 μl of buffer (sodium carbonates 0.15 *M*, pH 9.5) and 500 μl of reagent (6 m*M* dabsyl chloride in acetonitrile) and heated at 70°C for 30 min. Two ml of benzene were added after cooling in an ice bath, followed by sonication and centrifugation until two homogeneous phases were obtained; 1.8 ml of the organic extract were separated and dried at room temperature under nitrogen current. The residue was dissolved in 0.9 ml of acetone and filtered through

Table 1
Biogenic amines abbreviations, purities and concentration ranges

Amine	Abbreviation	Purity[a] (% w/w)	Concentration (nmol/ml)	
			Lowest	Highest
Tryptamine	Try	99	15.47	4866
Phenethylamine	Phe	99	17.29	5238
Putrescine	Pu	99	16.79	5281
Cadaverine	Ca	99	17.13	5188
Histamine	Hist	98	10.34	3251
Serotonine	Ser	99	10.34	3251
Tyramine	Tyr	98	10.63	3220
Spermidine	Spd	99	10.64	3224
Spermine	Sp	98	10.33	3248

[a] As given by the supplier.

0.2 μm Nylon membranes; 5 μl of this solution was injected.

Chromatography was performed in a Shimadzu LC-10A instrument, equipped with a Sil-10A auto-injector and an SPD-M10A diode array detector (Shimadzu Co., Kyoto, Japan). Shim-Pack SBC-ODS, 150×2.5 mm I.D., 5 μm spherical particles columns (Shimadzu Co.) thermostated to 30°C were used. Separation was obtained with a gradient beginning with a 12 min isocratic period with 73% v/v of a methanol–acetonitrile mixture (1.6:1) and 27% v/v water and ending with 95% v/v of the same organic mixture and 5% v/v water. Flow rate was 0.3 ml/min and detection was at 450 nm. Separation of the nine amines and restoration of the conditions to those prevailing at beginning of the run demanded 50 min; chromatograms for solutions of both series are shown in Fig. 2. Fourteen standard solutions (seven per each series) covering the whole concentration range for each amine were derivatized and chromatographed following a randomized sequence on each of seven successive days.

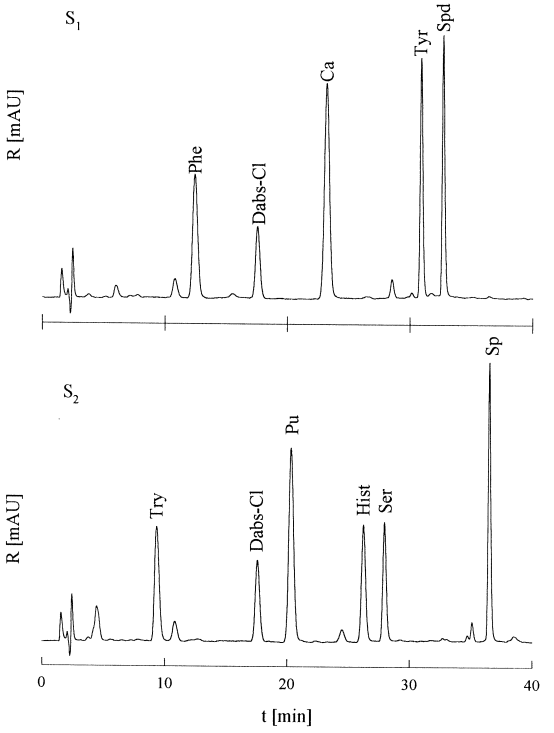

Fig. 2. Separation of the amines contained in standard solutions of series S_1 and S_2. See Experimental for chromatographic conditions and Table 1 for abbreviations.

4. Results

4.1. Signal variance

The number of analyte levels as well as that of replicates was a compromise between time demanded by the assay and reliability in the measurements. The variance of the measurements increases with analyte concentration in all the cases, and heteroscedasticity was confirmed by means of Cochran's test ([5], Chapter 6). Some of the s_i versus X_i behaviors can be fitted to second order polynomials (as is the case of Spd, Fig. 3) while in other cases the points are scattered about what can be envisaged as linear relationships (as Hist, Fig. 3). Some of these

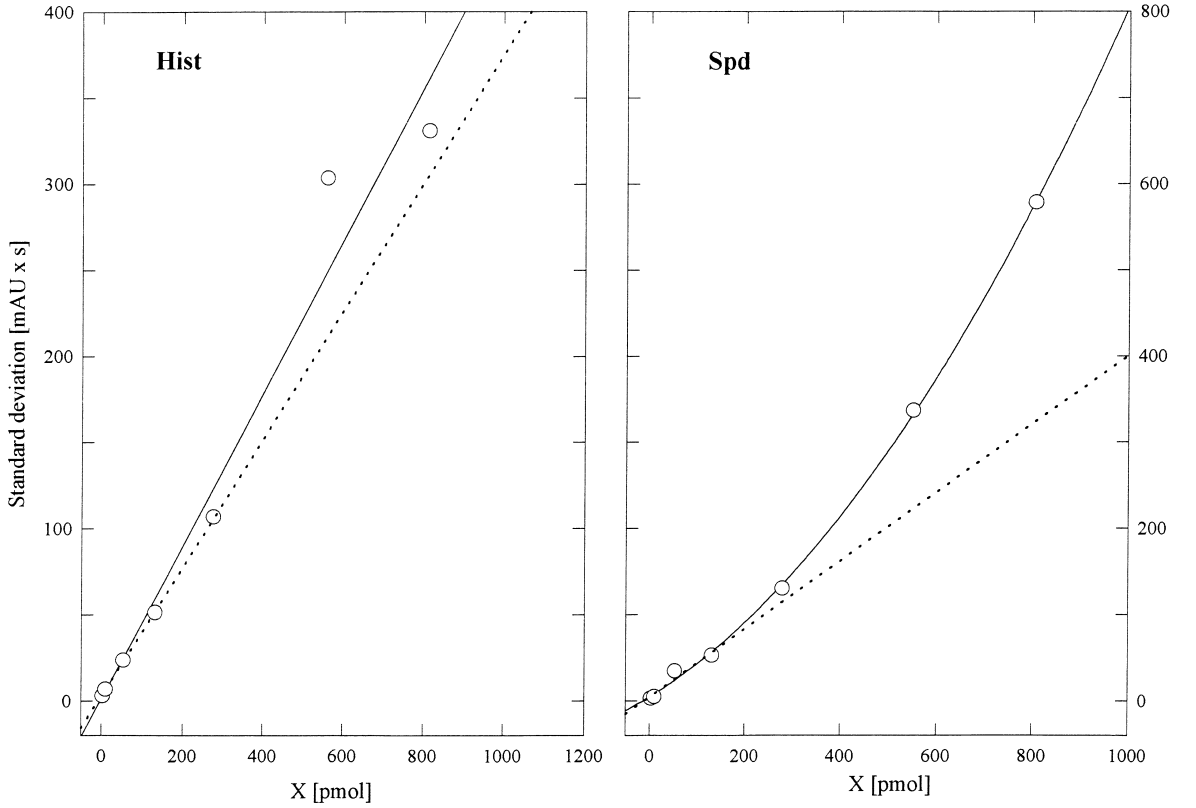

Fig. 3. Peak area standard deviation s_i as a function of injected analyte amount X_i. Full lines correspond to attempts of fitting s_i to X_i through all the experimental range. Dotted lines correspond to fitting s_i to Eq. (11).

functions, however, displayed inappropriate behaviors below the lowest experimental analyte level (as a minimum, or extrapolation to negative values at $X=0$). Since as a general rule no anomalous points and a regular increase of s_i with X_i was observed, statistical weights necessary to perform WLSR through the whole range were calculated from raw standard deviations; this was considered as a more realistic attitude than smoothing standard deviation results by fitting to arbitrary functions. However, at the low concentration region (corresponding to X_i between about 2.5 and 150 pmol) s_i and X_i are linearly related, with correlation coefficient that in all cases is higher than 0.997; these equations are subsequently used to estimate L_D and L_Q. As shown in Fig. 4, the prediction limits calculated by means of these linear equations are coincident in the low concentration region with those obtained from the fitting of s_i to X_i for the whole concentration range.

4.2. Regression analysis

Peak areas (Y_i) were fitted to a model of first order in the injected analyte amount (X_i/pmol, as calculated from derivatization procedure) both by means of ULSR and of WLSR. Basic equations for WLSR are summarized in Table 2. Normalized weights were adopted [Eq. (T.2)], with $\Sigma w_i = N$, the total number of independent measurements. ULSR equations are obtained from those in Table 2 by making $s_i = s$ (i.e. $w_i = 1$) for all i.

Initial regressions were performed on the complete data set, covering the whole concentration range, and the F-test at a 0.05 significance level was then used to check the fit of experimental data to the linear model. Peak areas measured at the highest analyte level were removed in those cases where lack-of-fit was detected and a regression through the remaining points was performed, followed by a new application

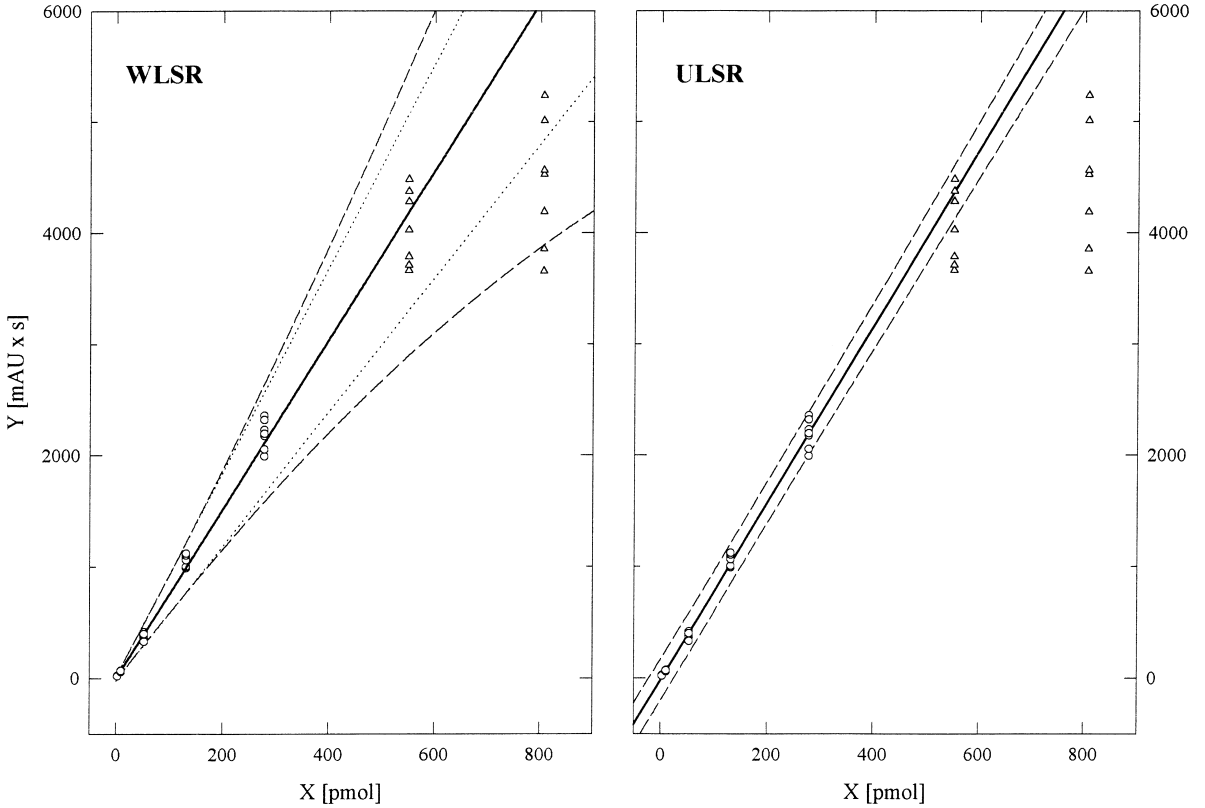

Fig. 4. Weighted and unweighted least-squares regression analysis for Spd. Includes the calibration line (—); prediction limits (– – –), w_i calculated by fitting s_i to X_i by means of a second degree polynomial (see Fig. 3); prediction limits at the low concentration region (·····), w_i estimated by means of Eq. (11). Circles: Points within the linear range. Triangles: Points outside the linear range.

Table 2
Summary of equations for WLSR calculations

$\hat{Y}_{iw} = b_{0w} + b_{1w}X_i$	(T.1)	$w_i = N(1/s_i^2)/\sum(1/s_i^2)$	(T.2)
$\bar{X}_w = \sum w_i X_i / \sum w_i$	(T.3)	$\bar{Y}_w = \sum w_i Y_i / \sum w_i$	(T.4)
$S_{xxw} = \sum w_i (X_i - \bar{X}_w)^2$	(T.5)		
$b_{1w} = \sum w_i Y_i (X_i - \bar{X}_w)/S_{xxw}$	(T.6)	$b_{0w} = \bar{Y}_w - b_{1w}\bar{X}_w$	(T.7)
$s^2(b_{1w}) = s_{ew}^2/S_{xxw}$	(T.8)	$s^2(b_{0w}) = s_{ew}^2 \sum w_i X_i^2 / S_{xxw} \sum w_i$	(T.9)
$s_{ew}^2 = \sum w_i (Y_i - \hat{Y}_i)^2/(N-2)$			(T.10)
$Y_i = \hat{Y}_{iw} \pm t_{(1-\alpha,\phi)} s_{ew} \left[1/w_i + 1/\sum w_i + (X_i - \bar{X}_w)^2/S_{xxw}\right]^{1/2}$			(T.11)

For ULSR, make $w_i = 1$ for all i; X_i: Analyte amount (pmol); Y_i: Peak area (μAU×s); N: Total number of calibration points; s_i: Peak area standard deviation at the i-level; w_i: Normalized statistical weight; $\hat{Y}_{iw}$: Predicted peak area at the i-level; b_{0w}, b_{1w}: Weighted intercept and slopes estimates; $\bar{X}_w$, $\bar{Y}_w$: Weighted mean X and Y; s_{ew}: Weighted residual standard deviation.

of the F-test. The iterative process was continued until no lack-of-fit was detected and the highest level remaining at that step was considered as the upper limit of the linear range. Application of this criterion resulted in the removal of no level from Try data, of the highest level for Phe, Pu and Ca and of the two highest levels for the remaining amines. Scaled residuals for the weighted regression, defined as $(Y_{ij} - \hat{Y}_i)/s_i$, are shown in Fig. 5 for two representative cases within their respective linear ranges; residuals distribute homogeneously over the whole linear range and no trend is apparent.

Results of the weighted and unweighted regression analysis are given in Table 3. As usual, the slopes are only marginally affected by weighting, but the intercepts and their errors can drop by more than one order of magnitude. Weighting has also an important effect on the residual standard deviation, this resulting in a considerable narrowing of the prediction band at the low concentration region.

Method repeatability at two analyte levels, calculated from seven individually derivatized aliquots, is listed in columns A and B of Table 4. Instrumental repeatability, calculated from five chromatograms of the same derivatized aliquot, is listed under column C.

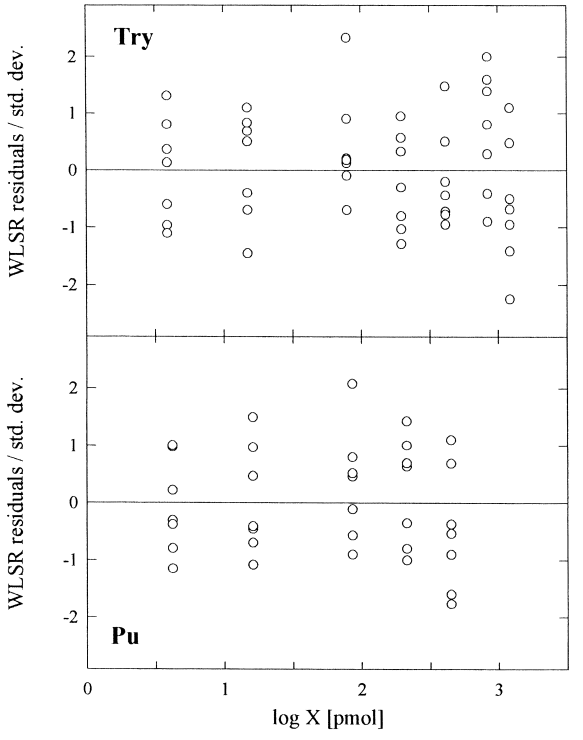

Fig. 5. Plot of WLSR scaled residuals $(Y_{ij} - \hat{Y}_{iw})/s_i$ against log X_i.

4.3. Detection limit by the regression approach

From Fig. 1 and Eq. (T.11):

$$\hat{Y}_C = b_{0w} + t_{(1-\alpha,\phi)} s_{ew} [1/w_0 + 1/\sum w_i + \bar{X}^2_w / S_{xxw}]^{1/2}$$

$$= -\hat{Y}_D - t_{(1-\beta,\phi)} s_{ew} [1/w_D + 1/\sum w_i + (L_D - \bar{X}_w)^2 / S_{xxw}]^{1/2} \quad (8)$$

where w_0 and w_D are the statistical weights at $X = 0$ and at $X = L_D$, respectively, and $t_{(1-\alpha,\phi)}$ is Student's factor with a probability α of been exceeded on a distribution with $\phi = N - 2$ degrees of freedom; $t_{(1-\beta,\phi)}$ is analogously defined at the lower limit. By combining Eq. (8) with Eq. (T.1), we get:

$$L_C = [t_{(1-\alpha,\phi)} s_{ew} / b_{1w}][1/w_0 + 1/\sum w_i + \bar{X}^2_w / S_{xxw}]^{1/2} \quad (9)$$

and

$$L_D = L_C + [t_{(1-\beta,\phi)} s_{ew} / b_{1w}][1/w_D + 1/\sum w_i + (L_D - \bar{X}_w)^2 / S_{xxw}]^{1/2} \quad (10)$$

$\alpha = \beta = 0.01$ was chosen for all the calculations. As mentioned in the former section, s_i can be satisfactorily expressed at low concentrations by means of:

$$s_i = a_0 + a_1 X_i \quad (11)$$

where a_0 and a_1 are constants. Eqs. (11) and (T.2) are used to estimate w_0 and w_D. Then L_C is calculated by means of Eq. (9), but calculation of L_D by means of Eq. (10) requires iteration; convergence is rapid, the root being attained in 5–10 iterations. Eqs. (9) and (10), after making $w_i = 1$ for all i, are used to calculate L_D from results obtained in ULSR. Results calculated by means of both regression techniques can be read in Table 5.

4.4. Quantitation limit

Eq. (4) can be estimated as:

$$L_Q = k_Q (s_{LQ} / b_{1w}) \quad (12)$$

Table 3
Least squares parameters of the regression of peak area (Y_j/μAU×s) against injected analyte amount (X_j/pmol). For each amine data on the first line correspond to WLSR and those on the second line to ULSR

Amine	N[a]	Intercept±s[b]	Slope±s[b]	Residual standard deviation	Upper limit of the linear range (pmol)
Try	49	4358±1319	4811±55	7392	1200
		36788±45245	4694±77	197722	
Phe	42	643±2421	4758±28	12098	900
		−10190±12125	4856±29	49865	
Pu	42	−1943±2117	8052±108	10154	900
		89399±55433	7335±131	228510	
Ca	42	1411±2484	8260±94	12839	900
		−21827±26250	8456±63	107959	
Hist	35	6036±1720	6431±133	7904	280
		31998±15980	5956±115	58516	
Ser	35	−425±1489	4508±93	6721	280
		16255±10425	4303±75	38176	
Tyr	35	995±1600	7401±116	7084	280
		−306±15644	7423±112	57430	
Spd	35	−1978±2159	7609±192	9433	280
		−14798±17661	7950±127	64832	
Sp	35	−7008±2167	8241±213	9249	280
		13501±25185	8598±182	92223	

[a] Number of calibration data points.
[b] Standard deviation.

If a function $s(X)=f(X)$ has been previously evaluated, L_Q can be determined, for instance, by plotting $b_{1w}\ X/f(X)$ against X; L_Q is the value of X corresponding on the plot to an ordinate k_Q. For the particular case of this work the simplicity of Eq. (11) facilitates an analytical solution:

$$L_Q = k_Q a_0/(b_{1w} - k_Q a_1) \quad (13)$$

Values of L_Q calculated for IUPAC default value, $k_Q=10$, can be read in Table 5. The results were confirmed by repeated derivatization and chromatography of standard solutions with concentrations close to those resulting in the injection of $X=L_Q$ pmol.

Table 4
Results for the repeatability of the method at two analyte levels (A and B) and for instrumental repeatability at one level (C)

Amine	% RSD		
	A	B	C
Try	5.2 (80)	4.3 (400)	0.96 (400)
Phe	2.6 (85)	2.2 (200)	1.7 (200)
Pu	5.7 (85)	4.3 (200)	2.2 (200)
Ca	4.9 (85)	3.5 (200)	1.8 (200)
Hist	6.4 (50)	6.0 (130)	2.9 (130)
Ser	7.1 (50)	6.2 (130)	1.7 (130)
Tyr	6.3 (50)	3.2 (130)	1.5 (130)
Spd	9.5 (50)	5.0 (130)	2.5 (130)
Sp	8.5 (50)	6.5 (130)	1.6 (130)

A and B: Seven aliquots, individually derivatized, one injection each; C: Five injections of the same derivatized aliquot; % RSD$=(s_i/\bar{Y}_i)\times 100$, with s_i and $\bar{Y}_i$ expressed in peak area units; Numbers between parentheses indicate injected analyte amount (pmol).

4.5. Detection and quantitation limits by the signal-to-noise ratio (S/N ratio approach)

Peak-to-peak noise, h_N, was measured on chromatograms of blank samples run on different days. Digital data (means of absorbance readings taken for periods of 0.64 s) collected for 3 min before and 3 min after each analyte retention time were exported and linearly fitted versus time by means of ULSR; influence of base line drift was removed in this way and the root mean square noise, rsmn [27] was obtained as the residual standard deviation of the fit.

Table 5
Detection and quantitation limits calculated by the regression approach (ULSR and WLSR) and by the signal-to-noise ratio method

Amine	Detection limit/pmol			Quantitation limit/pmol	
	ULSR	WLSR	S/N ratio[a]	Eq. (13)	S/N ratio[a]
Try	210 (840)[b]	2.7 (10.8)	6.2 (24.8)	7.8 (31.2)	33.2 (132.8)
Phe	52 (208)	7.4 (29.6)	6.7 (26.8)	12.2 (48.8)	36.0 (144.0)
Pu	55 (220)	3.1 (12.4)	4.2 (16.8)	8.5 (34.0)	20.5 (82.0)
Ca	65 (260)	4.6 (18.4)	4.4 (17.6)	7.7 (30.8)	21.1 (84.4)
Hist	51 (204)	4.0 (16.0)	4.2 (16.8)	11.5 (46.0)	20.5 (82.0)
Ser	46 (184)	4.5 (18.0)	3.9 (15.6)	12.9 (51.6)	25.1 (100.4)
Tyr	40 (160)	4.5 (18.0)	2.3 (9.2)	8.3 (33.2)	12.4 (49.6)
Spd	42 (168)	6.2 (24.8)	2.6 (10.4)	12.7 (50.8)	11.7 (46.8)
Sp	55 (220)	3.4 (13.6)	3.1 (12.4)	10.8 (43.2)	11.8 (47.2)

[a] Calculated by means of $L = k\,[h_N/h(X)]\,X$ where $h(X)$ is peak height at the analyte level X and h_N is peak-to-peak noise, using $k=2$ for L_D and $k=10$ for L_Q.
[b] Numbers between parentheses indicate the concentration (nmol/ml) of the standard solution subjected to the procedure.

Peak-to-peak noise is between three and six times larger than rmsn, depending upon noise characteristics; visual inspection of our base lines indicated that making $h_N = 5 \times$ rmsn was a safe choice.

Peak heights, $h(X)$, were measured by means of instrument software from the base line defined for integration purposes after careful selection of peak detect, top and end points. Then $h(X)/h_N$ ratios were fitted to a model of first order in X by means of WLSR using the inverse variance of peak heights as weighting factors. L_D and L_Q, calculated from the regression equation as the X values making $h(X)/h_N = 2$ or 10, respectively, can be read in Table 5. It should be understood that this methodology was applied in order to obtain the best that can be expected from the approach; as normally used, the S/N ratio approach consists in making a short number of measurements in the low concentration region and interpolating between them.

5. Discussion

Deleterious effects of derivatization on precision are clearly shown in columns B and C of Table 4, where whole method precision and purely chromatographic precision at the same analyte level are compared. The results gathered in the table fall within the ranges estimated by other authors [1].

The upper limit of the linear range (Table 3) decreases as the analyte retention time increases; this last depends on derivative hydrophobicity, which in his turn is strongly affected by the number of dabsyl groups linked to one analyte molecule. Therefore, it is hardly possible to state if the deviations from the linear behavior at the highest concentrations result from derivatives solubilities or from limitations in the extent of reaction imposed by the low solubility of the reagent in the reaction mixture.

The unweighted prediction intervals at low concentrations are wider than the weighted intervals in situations of non-constant signal variance (see Fig. 4). Therefore, detection limits are generally overestimated when heteroscedastic calibration data are treated by the ULSR-approach [24–26]; unweighted detection limits listed in Table 5 range from 7 to 78 times the weighted values and from 8 to 34 times those estimated by means of the S/N ratio approach.

Closer agreement is observed on comparing detection limits estimated by means of WLSR and S/N ratio approaches. Paired differences are positive in some cases and negative in others, and range from less than 5% to more than 100% of the estimated values; however, no significant difference between the results of the two methods was detected by means of the Wilcoxon signed rank non-parametric test ([5], Chapter 12).

Quantitation limits calculated by the S/N ratio approach with $k_Q = 10$ are larger than those obtained by means of Eq. (13), corresponding to %RSD that range from 4.5 to 10.5.

It is not easy to decide if these differences, not

negligible but not catastrophic, justify to change from the long used, intuitive and familiar to chromatographers S/N ratio approach to the more elaborate, with more solid theoretical basis (but also experimentally more demanding) WLSR-approach. Our suggestions can be summarized in the following terms:

(a) The regression approach in the form recommended by ICH [9] (i.e. estimating the blank standard deviation by means of the residual standard deviation of the regression line or by means of the slope standard deviation) is not advisable by the reasons given in Section 2.3.

(b) The S/N approach, as described by the ICH [9], can be used in those situations in which the estimation of the lowest limits of the method constitutes the only objective.

(c) In the more usual situations in which reproducibility and linear range estimations are as important as those of lowest limits, the methodology proposed in the present paper constitutes a more reliable option. On the other side, once the chromatograms were run their results can be used to apply the S/N approach, providing a confirmatory result.

6. Conclusions

A proposal for the preliminary evaluation of the quantitative performance of a chromatographic method is presented and discussed. Detection and quantitation limits, linear range and repeatability are calculated from the results of a calibration with several replicates at each analyte level. Replicates are essential in order to detect changes in precision with analyte concentration, since application of unweighted least-squares regression in situations of non-constant signal variance yields overestimated detection limits; on the other side replicates are also essential to check response linearity and repeatability. An statistically acceptable agreement is observed for the nine analytes assayed between detection limits estimated by weighted least-squares regression and by means of the signal-to-noise approach. Quantitation limits estimated as the analyte amount resulting in peak heights ten times higher than the base line peak-to-peak noise correspond to RSD values (calculated as peak area standard deviation/mean peak area ratio) smaller than 0.10, the default value prescribed by IUPAC.

Acknowledgements

This work was sponsored by CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina), by Agencia Nacional de Promoción Científica y Tecnológica (contract BID 802/OC-AR) and by CICPBA (Comisión de Investigaciones Científicas de la Provincia de Buenos Aires). M.A.C. acknowledges contract UNLP-FOMEC for a fellowship.

References

[1] L.R. Snyder, J.J. Kirkland, J.L. Glach, Practical HPLC Method Development, 2nd ed., Wiley, New York, 1997.
[2] P.J. Schoenmakers, Optimization of Chromatographic Selectivity, Elsevier, Amsterdam, 1986.
[3] H. Lingeman, W.J.M. Underberg (Eds.), Detection — Oriented Derivatization Techniques in Liquid Chromatography, Marcel Dekker, New York, 1990.
[4] G. Lunn, L.C. Hellwing, Handbook of Derivatization Reactions for HPLC, Wiley, New York, 1998.
[5] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, 1997.
[6] EURACHEM Guide. The Fitness for Purpose of Analytical Methods, LGC Ltd, Teddington, UK, 1998.
[7] M.E. Swartz, I.S. Krull, Analytical Methods Development and Validation, Marcel Dekker, New York, 1997.
[8] ICH Harmonized Tripartite Guideline. Text on Validation of Analytical Procedures, ICH4, October 1994.
[9] ICH Harmonized Tripartite Guideline. Validation of Analytical Procedures: Methodology, ICH4, November 1996.
[10] Chapter 13 D.A. Skoog, F.J. Holler, T.A. Nieman, in: Principles of Instrumental Analysis, 5th ed., Harcourt Brace and Co, Philadelphia, 1998.
[11] Analytical Chemistry Committee, Royal Society of Chemistry, London, Analyst 119 (1994) 2363.
[12] J. Vial, A. Jardy, J. High Resol. Chromatogr. 22 (1999) 217.
[13] IUPAC, prepared by L.A. Currie, IUPAC Commission on Analytical Nomenclature, Recommendations in Evaluation of Analytical Methods including Detection and Quantification Capabilities, Pure Appl. Chem. 67 (1995) 1699, Transcription in Anal. Chim. Acta 391 (1999) 105.
[14] L.A. Currie, Anal. Chim. Acta 391 (1999) 127.

[15] J.C. Miller, J.N. Miller, Statistics for Analytical Chemistry, 2nd ed., E. Horwood, London, 1988.
[16] J. Vial, A. Jardy, Anal. Chem. 71 (1999) 2672.
[17] A. Hubaux, G. Vos, Anal. Chem. 42 (1970) 849.
[18] L.A. Currie, Anal. Chem. 40 (1968) 586.
[19] G.T. Vernimont, in: W.E. Spendley. (Ed.), Use of Statistics to Develop and Evaluate Analytical Methods, AOAC, Arlington, VA, 1985.
[20] IUPAC Committee on Analytical Nomenclature, Pure Appl. Chem. 66 (1994) 595.
[21] German Normative Institute, DIN 32645. Beuth, Berlin, 1994.
[22] N.R. Draper, H. Smith, Applied Regression Analysis, 2nd ed., J. Wiley and Sons, New York, 1981.
[23] L. Oppenheimer, T.P. Capizzi, R.M. Weppelman, H. Mehta, Anal. Chem. 55 (1983) 638.
[24] M.E. Zorn, R.D. Gibbons, W.C. Sonzogni, Anal. Chem. 69 (1997) 3069.
[25] J.S. Garden, D.G. Mitchell, W.N. Mills, Anal. Chem. 52 (1980) 2310.
[26] J.R. Burdge, D.L. MacTaggart, S.O. Farwell, J. Chem. Ed. 76 (1999) 434.
[27] J.V. Hinshaw, LC–GC 18 (2000) 126.