ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
28 (2002) 421–429

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

www.elsevier.com/locate/jpba

Validation of HPLC and GC–MS systems for bisphenol-A leached from hemodialyzers on the basis of FUMI theory

Yuzuru Hayashi *, Rieko Matsuda, Yuji Haishima, Takeshi Yagami, Akitada Nakamura

National Institute of Health Sciences, 1-18-1 Kami-Yoga, Setagaya, Tokyo 158, Japan

Received 28 May 2001; accepted 6 July 2001

Abstract

This paper proposes a method for the validation of chromatography systems in which many experiments to estimate SD or RSD are difficult or impossible to carry out because of time, cost, etc. HPLC systems with UV–Vis and fluorescence detectors and GC–MS system for bisphenol-A leached from hemodialyzers are taken as an example. Examined as validation characteristics are not only the ordinary quantities (precision, accuracy, range, limit of detection (LOD), limit of quantitation (LOQ), specificity and linearity) but also precision plots (measurement RSD vs. concentration), 95% confidence intervals of calibration lines and LOD signals over baselines. The precision plots, calibration confidence intervals and LOD signals are shown to be advantageous to validate and compare the analytical performance of the systems. The LOD, LOQ, precision plots and 95% confidence intervals of calibration lines are all derived from the SD of measurements and the reliability of these quantities and plots depends totally on the reliability of the SD estimates. This paper uses a probability theory, called the FUMI theory, to estimate as exact a measurement SD as possible without the replication. The precision of the HPLC and GC–MS systems is shown to coincide with the repeatability obtained by the repetition of measurements.

Keywords: Precision; Chromatography; Validation; FUMI theory

1. Introduction

Recently, attention has been paid to the validation of an analytical procedure in many societies. The main objective of the validation is to demonstrate that the analytical procedure is suitable for the purpose of the analysis [1]. Its social aim will be that analytical measurements made in one location should be consistent with those made elsewhere [2].

The ICH guideline proposes the analytical characteristics that are criteria to judge the suitability and capabilities of the analytical procedures [1]. Among them are specificity, linearity, range, accuracy, precision, limit of detection, limit of quantitation and robustness. The precision which is described as the relative standard deviation (RSD)

* Corresponding author. Tel.: +81-3-3700-1141; fax: +81-3-3707-6950.

0731-7085/02/$ - see front matter
PII: S0731-7085(01)00698-7

or standard deviation (SD) of measurements is of special importance in that significant quantities such as limit of detection (LOD) and limit of quantitation (LOQ) are calculated directly from the measurement SD.

An exact value of the precision, however, is not easy to estimate in practice. As is well-known, a reliable SD estimate cannot be obtained until the same experiments are repeated many times. As an example, the 95% confidence intervals of the SD estimate from five measurements are about $\pm$ 65% of the true value. Furthermore, situations often arise in which time, cost or sample amount makes it difficult or almost impossible to repeat experiments many times.

This paper focuses on the analytical performance of the HPLC and GC–MS systems which have already been used in our laboratory for bisphenol-A leached from hemodialyzers [3]. The precision is estimated by a probability theory called the Function of Mutual Information (FUMI) theory [4–6]. The prominent advantages of the theory are that the repeated measurements are dispensed with and that the measurement SD estimated from a single measurement (chromatogram) has the 95% confidence intervals corresponding to those for about 30 measurements ($\pm$ 20% of the true value) [6]. The high reliability of the theoretical SD is not strange, because the FUMI theory utilizes 1024 or 2048 data points of a baseline instead of 30 measurements.

Bisphenol-A is an environmental estrogen-like chemical and has been shown to affect reproduction in wild life and public health [7–18]. This phenolic xenoestrogen is a major component of polycarbonate, polysufonate and epoxy resins. Human exposure can derive from medical appliances or leachate of foodstuffs, if they are made up of or packaged in the resins. Bisphenol-A elution from hemodialyzers is critical in that the potentially hazardous material is taken immediately into blood circulation.

2. Materials

Hemodialyzers which were composed of polycarbonate casing and cellulose acetate or polysulfonate hollow-fiber were offered from three makers, A–C [3]. All the reagents were of analytical grade or equivalent and used without further purification.

3. Methods

3.1. Whole chemical analysis

The detailed procedures for the analysis were given elsewhere [3]. The HPLC and GC–MS analyses were designed and performed according to literature [19–23].

The entire steps for the bisphenol-A analysis consist of the sampling, preparation and measurement (see Fig. 1). The sampling is carried out by circulating the bovine serum in the hemodialyzers. The aim of the preparation is to enrich the analyte. In Fig. 1, the calibration is involved in the chromatography step for convenience.

A Hitachi AS-2000 liquid chromatograph was equipped with an F-1080 fluorescence detection system (FL HPLC) or L-4250 UV–Vis detector (UV HPLC) and was controlled by a D-7000 system manager. For the FL HPLC system, the excitation and emission wavelengths were 280 and 310 nm, respectively, and for the UV HPLC system, the detection wavelength was 235 nm. A JEOL gas chromatograph (Automass, GC–MS) was used with a DB-5MS column (0.25 mm $\times$ 30 m) and electron impact ionization (positive-ion mode) of the mass spectrometer.

3.2. Uncertainty prediction

The detailed explanation of the FUMI theory has already been given in Ref. [4], but the theory is briefly reviewed for understanding this paper. The applicability of the theory has been verified with experiments to the HPLC systems with the UV–Vis photo-multiplier and photo diode array detectors [24], but the HPLC with the fluorescence detector and GC–MS were first examined in this paper.

First, the measurement SD and RSD are estimated from the stochastic properties of signal and noise of a chromatogram. Then, LOD, LOQ, the

precision plots and 95% confidence intervals of a calibration line are calculated based on the theoretical SD or RSD [5,6].

In the FUMI theory, the time variation in the baseline (noise) is described as the mixed random processes of white noise and the Markov process. The white noise is a time-independent process with one parameter, but the Markov process has a time-correlation (also called the auto-correlation) with two parameters. The three noise parameters can uniquely be determined by the least-squares fitting of the theoretical power spectrum of the white noise plus Markov process to the observed noise power spectrum of the baseline and are in turn used for calculating the measurement SD in the time space. Coupled with the information on the signal shape over an integration domain, the FUMI theory can provide the measurement RSD (= (measurement SD)/(averaged measurements)).

At high concentrations, the measurement uncertainty originates mainly from the injection error. The FUMI theory also takes this error into account to estimate the precision over a wide concentration range. For example, the RSD value (0.59%) at the upper limit of the range (100 ppb) for the UV HPLC system in Table 3 was calculated from the uncertainty caused by the baseline fluctuation and the injection error (0.25% RSD). The UV and FL HPLC systems are assumed to have the same injection error (0.25% RSD) according to literature [4]. The error of the manual injection used in the GC–MS system (7% RSD) was experimentally determined.

The calculation of the measurement SD and RSD was performed on a computer software

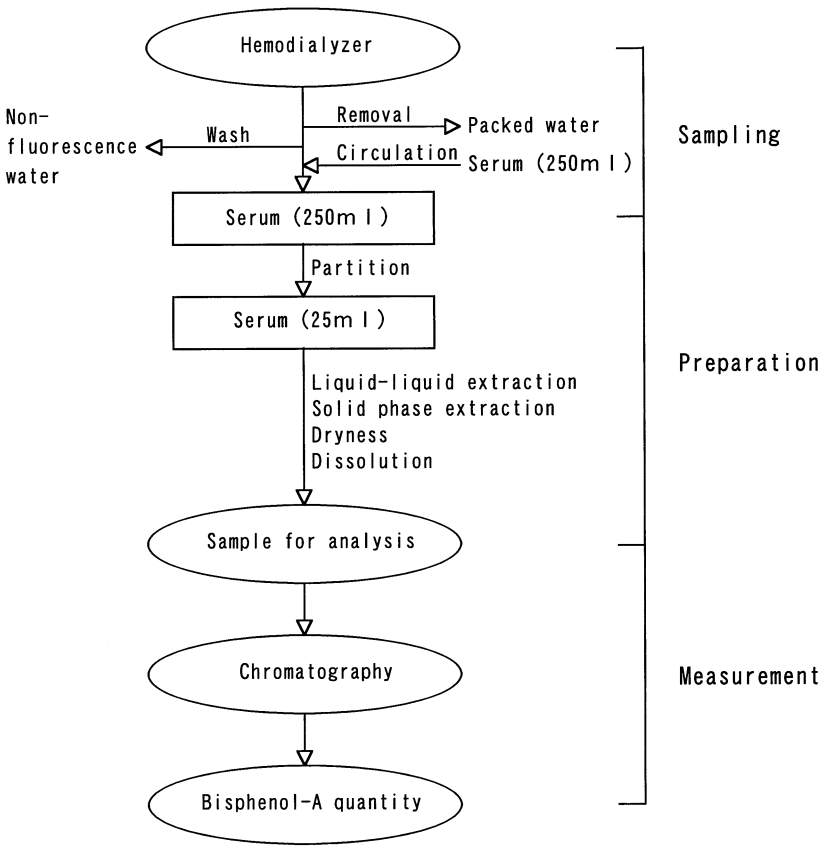

Fig. 1. Steps of the bisphenol-A analysis. The detailed procedures were written elsewhere [3].

Table 1
Precision (% RSD) of the hemodialyzer samples in the analytical instruments

Hemo-dialyzers	Bisphenol-A quantity (ppb)	% RSD		
		HPLC (FL)	HPLC (UV)	GC–MS
A	13.2	2.0	33	7.0
B	18.3	1.5	25	7.0
C	94.9	0.3	5.5	7.0

called TOCO version 2.0 (Total Optimization of Chemical Operations). Figs. 3–5 and Tables 1–3 were also made by the TOCO.

The 95% confidence intervals of a calibration line were also calculated according to the probability theory of calibration [6]. In the well-known statistical approach, the right confidence intervals of a calibration line cannot be obtained without appropriate weights of the least-squares fitting of the calibration line in the hetero-scedastic situations [25]. However, the theory can provide the right intervals, without any weight, in both the homo- and hetero-scedastic situations, if the measurement SD of all the calibration samples can be known before the calibration process [6]. In this paper, the SD values estimated theoretically as mentioned above are used for the 95% intervals. The uncertainty of the instruments examined here is hetero-scedastic.

The experimental SD or RSD of area measurements was obtained with the same samples in the consecutive runs ($n = 3$ or 5) of the same analytical instrument. This SD or RSD corresponds to the repeatability by the definition of ICH [1].

As far as instrumental analyses are concerned, some methods for predicting the measurement SD from signal and noise in the instruments have been put forward so far [4–6,26–36].

4. Results

First, the concentration range is examined. The chromatograms for the samples from makers A–C are shown in Fig. 2. The lowest sample concentration is 13.2 ppb and the highest 94.9 ppb. The analysis to be conducted here is determined to range from 10 to 100 ppb.

The analytes are extracted in the bovine serum (see Fig. 1) and there could exist various interfering peaks in the chromatograms. However, the purity and specificity of the bisphenol-A peaks in the chromatograms were verified with GC–MS, LC–MS and NMR [3]. The analytical systems can be concluded to separate the analyte from the matrix components of the serum. The accuracy of the analyses can also be considered acceptable, since the recovery of the samples and linearity of the calibration lines (see below) are both satisfactory and no interfering peaks or ghost peaks could be found over the peak domains in the chromatograms [3].

Fig. 3 shows the calibration lines of bisphenol-A in the measurement systems (FL, UV and GC–MS). The lines are drawn by the least-squares fitting with no weight. The linearity of the calibration can be evaluated by the visual inspection. Although not shown in Fig. 3, the linearity range for the UV HPLC system is even wider than the other systems. The calibration data of the GC–MS appeared slightly curved and are grouped into three linear regions. However, this does not matter to the quantitative analysis.

The 95% confidence intervals of a calibration line represent the error between the calibration lines obtained under the same experimental conditions. The narrower the intervals, the more precise

Table 2
LOD and LOQ of bisphenol-A in the analytical instruments

Instruments	LOD (ppb)	LOQ (ppb)
HPLC (FL)	0.65	2.0
HPLC (UV)	16.0	48.0
GC–MS	0.16	1.9

Table 3
Validation results

Objects	Characteristics				Judgment
	Range (ppb)	Repeatability (% RSD)	LOD (ppb)	LOQ (ppb)	
	Acceptable limits				
	10–100	<10	<1	<10	
	Experimental results				
HPLC (FL)	10–100	0.38–2.7	0.65	2.0	Accepted
HPLC (UV)	10–1000	0.59–43	16.0	48.0	Rejected
GC–MS	0.2–1, 2–10, 20–100	7.0	0.16	1.9	Accepted

Accuracy is confirmed by the recovery of the instrumental analyses and by the linearity of the calibration lines. Specificity is regarded as satisfactory by GC–MS, LC–MS and NMR [3]. The experimental results for the repeatability are specified over the range (10–100 ppb).

the calibration. The 95% confidence intervals of the calibration lines are drawn above and below the calibration lines in Fig. 3. The 95% intervals for the GC–MS system can be spotted clearly, but those for the HPLC systems with the UV and FL detectors are difficult or too narrow to see. Fig. 3 demonstrates that if the quantitative results from many different calibration lines are compared, the results of the GC–MS system are not so precise as the FL and UV HPLC systems. The wide calibration intervals for the GC–MS arise from the error of the manual injection.

The measurement RSD estimated by the FUMI theory is listed in Table 1 for the real samples shown in Fig. 2. The precision of the FL HPLC is even higher than the other systems. The precision at one or several concentrations is useful, but the continuous change in the precision across the entire concentration range will provide more beneficial information about the uncertainty structure of the analysis (see below).

In Fig. 4, the theoretical precision curves (—) agree well with the experimental results (●) over the wide concentration range. We should note that the theoretical curves are not the least-squares fittings to the experimental results, but are obtained, independently of the observed RSD, from the signal and noise of the instrumental output. The repeated measurements (●) were carried out in order to confirm the reliability of the FUMI theory. If the prediction of the measurement RSD is successful in the instrumental analyses, the time- and effort-consuming replication can be dispensed with.

In general, the measurement RSD decreases with increasing analyte concentration. All the precision curves of the analytical systems in Fig. 4 observe this rule. Difference appears in the details of the precision (—). The precision of the FL HPLC is higher than that of the UV HPLC over the whole concentration range. The UV HPLC will be more useful for concentrated samples of 200–1000 ppb because of the narrow linearity range of the FL HPLC (see Table 3). At concentrations below 2 ppb, the GC–MS can give more precise measurements than the FL HPLC, but the precision above 2 ppb goes the other way around due to the serious injection error (RSD = 7%).

Another advantage of the precision plots is the visual estimation of LOD and LOQ. In this paper, the IUPAC definition is adopted: the LOD signal is three times the blank SD and the LOQ is ten times the blank SD [37]. The blank SD can be replaced by the measurement SD at low concentrations. The definitions also imply that the LOD is the concentration or signal magnitude of the sample which shows 33% RSD of measurements and that the LOQ corresponds to the sample of 10% measurement RSD. From the precision plot of the UV HPLC in Fig. 4, we can guess that the

LOD is about 20 ppb and the LOQ is about 50 ppb.

The theoretical LOD and LOQ estimates are listed in Table 2. The concentrations of the samples of makers A and B (13.2 and 18.3 ppb) are below the LOQ of the UV HPLC (48.0 ppb) and we can conclude that the UV system cannot quantify samples A and B with enough precision.

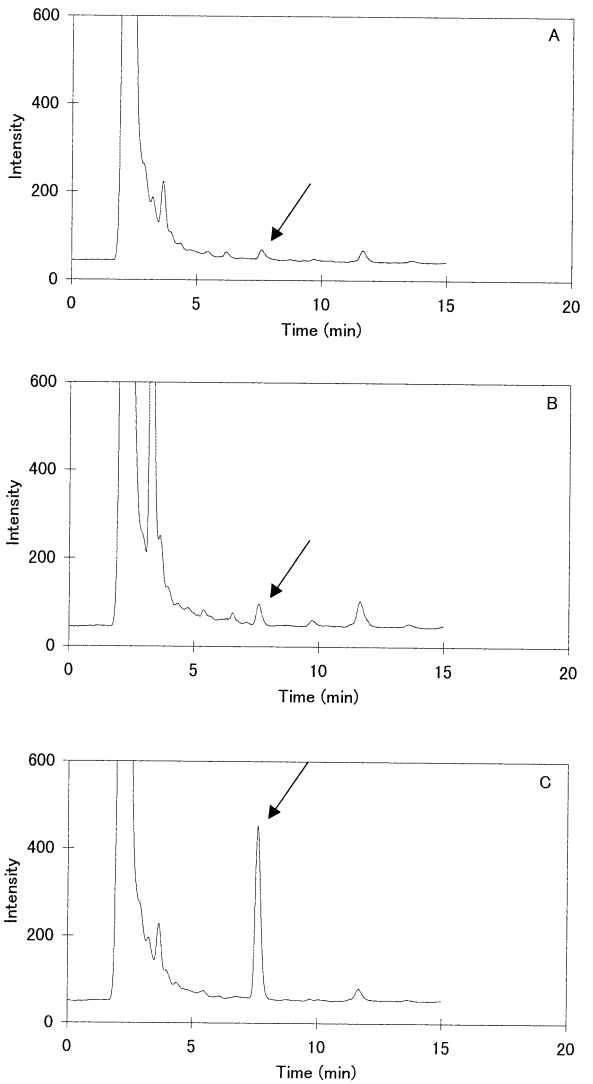

Fig. 2. Chromatograms for three samples. The chromatograms are obtained in the HPLC system with the fluorescence detector. (A) The sample from maker A; (B) from maker B; and (C) from maker C. The arrows show the bisphenol-A peaks and the concentration estimates are: A, 13.2; B, 18.3; C, 94.9 ppb.

However, the sufficient precision is available for samples A and B on the FL HPLC and GC–MS, because the LOQ (2.0 and 1.9 ppb) is lower than the concentration range of the real samples (10–100 ppb). The LOQ is almost the same for the FL HPLC and GC–MS, but the LOD of the FL HPLC is four times that of the GC–MS.

The LOD signals which are derived based on the theoretical SD are shown in Fig. 5 with the arrows. The signal-to-noise ratio (S/N) is often used as a yardstick of LOD [19,20,22]. In the FUMI theory, however, the LOD signals are determined from the SD of false measurements (false areas or heights) created by the baseline fluctuation, and not according to the S/N. Therefore, the S/N is different for the three LOD signals in Fig. 5. The LOD signal for the GC–MS looks noisy, having the lower S/N than that for the FL or UV HPLC.

The white noise looks like a high-frequency vibration (see the bottom figure of Fig. 5) and the Markov process a slowly changing fluctuation (see the middle figure of Fig. 5). The most significant feature of the random processes is that the white noise can be more effectively eliminated by the signal processing (integration or summation over the signal region) than the Markov process. That is, the Markov process has more critical effects on the precision than the white noise. Since more Markov process is included in the baselines of the FL and UV HPLC systems, the LOD signals for the HPLC systems look less noisy than that for the GC–MS system.

The visual inspection of the LOD signals will often be effective to find a failure of the theoretical LOD estimation. The LOD signals of Fig. 5 would fit the chromatographer's sense. If the noise impossible to treat in the FUMI theory such as spike noise or pulsation is involved in almost all parts of the baseline, larger LOD signals would result mistakenly.

For the sake of demonstration, the precision plots (—) in Fig. 4 and LOD signals in Fig. 5 are both derived from the averaged power spectra of the baselines contained in the chromatograms used for the calibration lines. If the plots and LOD signals are made from a single chromatogram, the uncertainty of them would be $\pm$

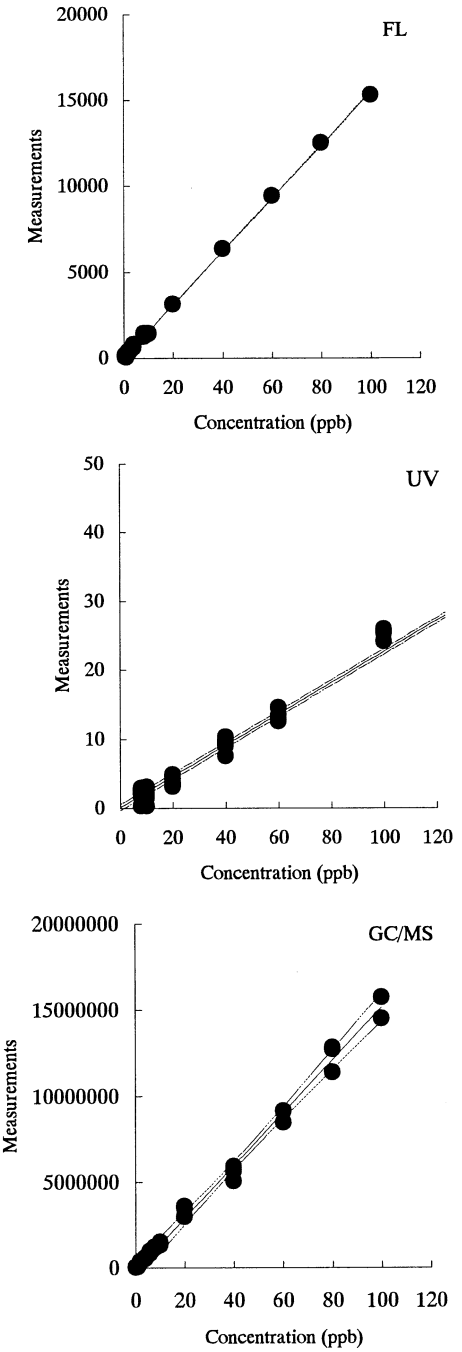

Fig. 3.

20% of the values in the figures (see Section 1 or Ref. [6]).

5. Discussion

Table 3 lists the results of the validation of the instrumental analyses for bisphenol-A leached from hemodialyzers. The acceptable limits of the validation characteristics should be determined according to the purpose of the analysis. Here, the range is set at 10–100 ppb, since the bisphenol-A concentrations of all the samples are covered in the range. The samples should be quantified with the required precision ($< 10\%$ measurement RSD). Therefore, the LOQ should be below the lower limit of the range (10 ppb). The acceptable LOD is set at a tenth of the lower limit of the range so that a more diluted sample down to 1 ppb can be detected.

Consequently, the FL HPLC and GC–MS are concluded to be suitable for the purpose of the analysis, but the UV HPLC is not acceptable because of the poor precision ($> 10\%$ RSD) in the range below 40 ppb. The FL HPLC is recommendable more than the GC–MS thanks to the higher precision over the range (see Table 3).

The validation results of this paper do not provide general aspects of instrumental analyses and are restricted within the systems examined here. The validation should be performed in each laboratory and for each system of a laboratory. However, the validation scheme proposed here will be valid to many instruments of a laboratory.

In this paper, the precision of the sample preparation was not examined. If the preparation error is regarded as significant, the precision should be estimated from the entire analysis including the preparation and instrumental analysis. However,

Fig. 3. Calibration lines and 95% confidence intervals. Upper figure: HPLC with the fluorescence detector (FL); middle figure: HPLC with the UV detector (UV); bottom figure: GC–MS. The slope and y-intercept of the calibration lines and correlation coefficient are: 154.3, 35.42, 0.9998 for the FL HPLC; 0.2256, 0.1753, 0.9997 for the UV HPLC; 1.538×10^5, -1.754×10^5, 0.9922 for the GC–MS. These values for the GC–MS result from the fitting of a line over the three linear regions (see the text).

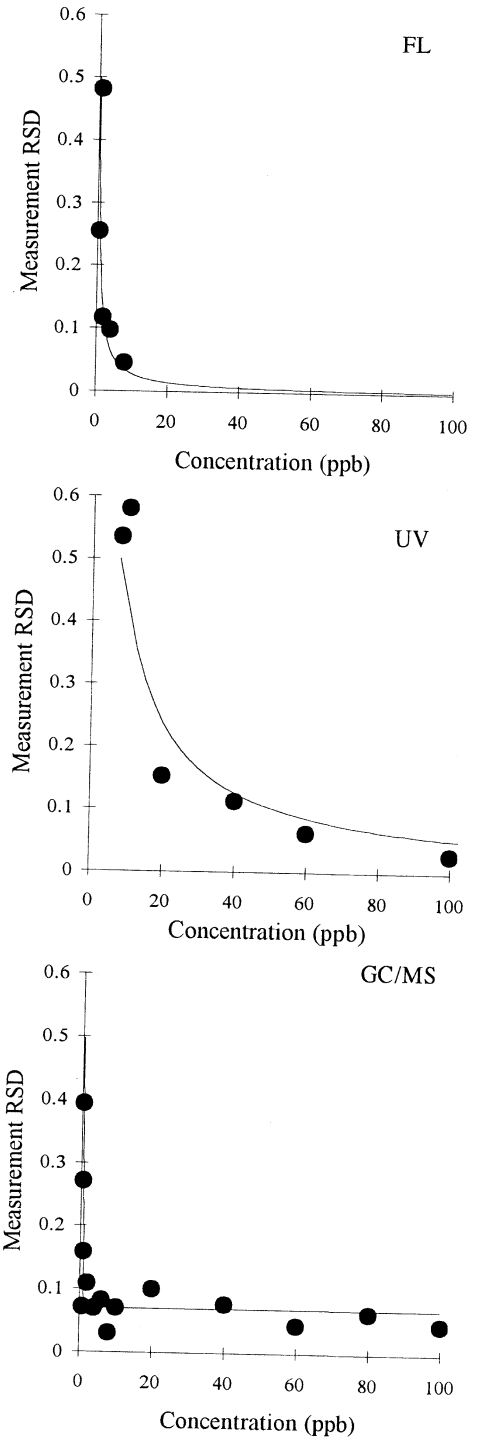

Fig. 4.

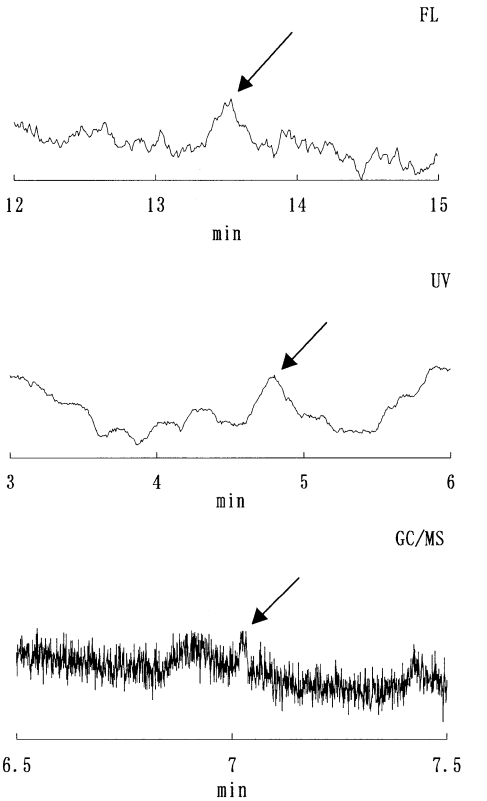

Fig. 5. LOD signals for bisphenol-A. The arrows show the LOD signals over the observed baselines. The LOD signals are created by reducing smooth, large peaks to the LOD concentrations and by adding them to the observed baselines.

as long as the quantitative analysis is concerned, the RSD for the preparation should be less than the upper limit (= 10% RSD) by definition [37]. Fortunately, in this case, the LOD for the entire analysis well agrees with the LOD estimated from the instrumental analysis alone. This is because RSD^2 for entire analysis = RSD^2 for preparation + RSD^2 for instrumental analysis (error propagation rule) and $34.5 = (10^2 + 33^2)^{1/2}$ (note that the RSD at the LOD is ca. 33%). In the same manner, if the RSD for the preparation is less

Fig. 4. Precision plots for the bisphenol-A analysis. The circles denote the RSD from the repeated measurements and the lines are the precision predicted by the FUMI theory. The theoretical lines are not the least-squares fitting to the experimental results (see the text). The replication number is five for the FL and UV HPLC and three for the GC–MS.

than 3%, the LOQ for the entire analysis is almost equal to the LOQ for the instrumental analysis alone, since $10.4 = (3^2 + 10^2)^{1/2}$.

The FUMI theory treats the error from the instrumental analysis only (see Section 3). The above consideration leads to the following conclusions:

(A) The LOD for the entire quantitative analysis is almost the same as the LOD estimated by the FUMI theory; and

(B) The LOQ for the entire quantitative analysis is almost the same as the LOQ estimated by the FUMI theory, if the preparation error is less than 3% RSD.

If the RSD for the entire analysis is estimated by the repeated experiments, we can divide the total error into the preparation error and measurement error by using the error propagation rule and FUMI theory. Then, a precision plot for the entire analysis would be obtained as a validation material. In this study, the total error magnitude was not assessed because of time for the preparation procedures. However, we believe that the FUMI theory has been shown to be useful as a validation method for quantitative analysis.

References

[1] ICH Topics Q2B, 1996.
[2] EURACHEM Guide, 1998.
[3] Y. Haishima, Y. Hayashi, T. Yagami, A. Nakamurua, J. Biomed. Mater. Res. Appl. Biomater. 58 (2001) 209–215.
[4] Y. Hayashi, R. Matsuda, Anal. Chem. 66 (1994) 2874–2881.
[5] R. Matsuda, Y. Hayashi, S. Sasaki, K. Saito, K. Iwaki, H. Harakawa, M. Satoh, Y. Ishizuki, T. Kato, Anal. Chem. 70 (1998) 319–327.
[6] Y. Hayashi, R. Matsuda, R.B. Poe, Analyst 121 (1995) 591–599.
[7] X. Long, R. Steinmetz, N. Ben-Jonathan, A. Caperell-Grant, P.C. Young, K.P. Nephew, R.M. Bigsby, Environ. Health Perspect. 108 (2000) 243–247.
[8] J. Ashby, P.A. Lefevre, J. Appl. Toxicol. 20 (2000) 35–47.
[9] T. Takao, W. Nanamiya, I. Nagano, K. Asaba, K. Kawabata, K. Hashimoto, Life Sci. 65 (1999) 2351–2357.
[10] F. Farabollini, S. Porrini, F. Dessi-Fulgherit, Pharmacol. Biochem. Behav. 64 (1999) 687–694.
[11] J.J. Segura, A. Jimenez-Rubio, R. Pulgar, N. Olea, J.M. Guerrero, J.R. Calvo, J. Endod. 25 (1999) 341–344.
[12] A. Mariotti, K.J. Soderholm, S. Johnson, Eur. J. Oral. Sci. 106 (1998) 1022–1027.
[13] D. Roy, J.B. Colerangle, K.P. Singh, Front Biosci. 6 (1998) d913–d921.
[14] T. Tsutsui, Y. Tamura, E. Yagi, K. Hasegawa, M. Takahashi, N. Maizumi, F. Yamaguchi, J.C. Barrett, Int. J. Cancer 75 (1998) 290–294.
[15] J.A. Dodge, A.L. Glasebrook, D.E. Magee, D.L. Phillips, M. Sato, L.L. Short, H.U. Bryant, Steroid Biochem. Mol. Boil. 59 (1996) 155–161.
[16] A. Atkinson, D. Roy, Biochem. Biophys. Res. Commun. 210 (1995) 424–433.
[17] A. Atkinson, D. Roy, Environ. Mol. Mutagen. 26 (1995) 60–66.
[18] S. Steiner, G. Honger, P. Sagelsdorff, Calcinogenesis 13 (1992) 969–972.
[19] U. Bolz, W. Körner, H. Hagenmaier, Chemosphere 40 (2000) 929–935.
[20] J. Sajiki, K. Takahashi, J. Yonekubo, J. Chromatogr., B 736 (1999) 255–261.
[21] J. Salafranca, R. Batlle, C. Nerin, J. Chromatogr., A 864 (1999) 137–144.
[22] S.N. Pedersen, C. Lindholst, J. Chromatogr., A 864 (1999) 17–24.
[23] C. Lambert, M. Larroque, J. Chrom. Sci. 36 (1997) 57–62.
[24] Y. Hayashi, R. Matsuda, Anal. Sci. 10 (1994) 725–730.
[25] J.N. Miller, Analyst 116 (1991) 3–14.
[26] E.D. Prudnikov, J.W. Elgersma, H.C. Smit, J. Anal. At. Spectrom. 9 (1994) 619–622.
[27] J.D. Ingle Jr., S.R. Crouch, Spectrochemical analysis, Prentice-Hall, New Jersey, 1988.
[28] L.D. Rothman, S.R. Crouch, J.D. Ingle Jr., Anal. Chem. 47 (1975) 1226–1233.
[29] N.W. Bower, J.D. Ingle Jr., Anal. Chem. 48 (1976) 686–692.
[30] N.W. Bower, J.D. Ingle Jr., Anal. Chem. 49 (1977) 574–579.
[31] P.W.J.M. Boumans, Anal. Chem. 66 (1994) 459A–467A.
[32] P.W.J.M. Boumans, Spectrochim. Acta, Part B 46 (1991) 917–939.
[33] C.Th.J. Alkemade, W. Snelleman, G.D. Boutilier, B.D. Pollard, J.D. Winefordner, T.L. Chester, N. Omenetto, Spectrochim. Acta, Part B 33 (1978) 383–399.
[34] G.D. Boutilier, B.D. Pollard, J.D. Winefordner, T.L. Chester, N. Omenetto, Spectrochim. Acta, Part B 33 (1978) 401–415.
[35] C.Th.J. Alkemade, W. Snelleman, G.D. Boutilier, J.D. Winefordner, Spectrochim. Acta, Part B 35 (1980) 261–270.
[36] B.D. Pollard, A.H. Ullman, J.D. Winefordner, Anal. Chem. 53 (1981) 330–336.
[37] G.L. Long, J.D. Winefordner, Anal. Chem. 55 (1983) 712A–724A.