Standard line slopes as a measure of a relative matrix effect in quantitative HPLC–MS bioanalysis

B.K. Matuszewski*

Merck Research Laboratories, West Point, PA 19486, USA

Abstract

A simple experimental approach for studying and identifying the relative matrix effect (for example "plasma-to-plasma" and/or "urine-to-urine") in quantitative analyses by HPLC–MS/MS is described. Using as a database a large number of examples of methods developed in recent years in our laboratories, the relationship between the precision of standard line slopes constructed in five different lots of a biofluid (for example plasma) and the reliability of determination of concentration of an analyte in a particular plasma lot (or subject) was examined. In addition, the precision of standard line slopes was compared when stable isotope-labeled analytes versus analogs were used as internal standards (IS). Also, in some cases, a direct comparison of standard line slopes was made when different HPLC–MS interfaces (APCI versus ESI) were used for the assay of the same compound, using the same IS and the same sample preparation and chromatographic separation conditions. In selected cases, the precision of standard line slopes in five different lots of a biofluid was compared with precision values determined five times in a single lot. The results of these studies indicated that the variability of standard line slopes in different lots of a biofluid [precision of standard line slopes expressed as coefficient of variation, CV (%)] may serve as a good indicator of a relative matrix effect and, it is suggested, this precision value should not exceed 3–4% for the method to be considered reliable and free from the relative matrix effect liability. Based on the results presented, in order to assess the relative matrix effect in bioanalytical methods, it is recommended to perform assay precision and accuracy determination *in five different lots* of a biofluid, instead of repeat (*n* = 5) analysis in the same, *single* biofluid lot, calculate standard line slopes and precision of these slopes, and to use <3–4% slope precision value as a guide for method applicability to support clinical studies. It was also demonstrated that when stable isotope-labeled analytes were used as internal standards, the precision of standard line slopes in five different lots of a biofluid was ≤2.4% irrespective of the HPLC–MS interface utilized. This clearly indicated that, in all cases studied, the use of stable isotope-labeled IS eliminated relative matrix effect. Also, the utilization of the APCI interface instead of ESI led to the elimination of the relative matrix effect in all cases studied. When the precision of standard line slope values exceeds the 3–4% limit, the method may require improvements (a more efficient chromatography, a more selective extraction, a stable isotope-labeled IS instead of an analog as an IS, and/or a change in the HPLC–MS interface) to eliminate the relative matrix effect and to improve assay selectivity.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Bioanalysis; Relative matrix effect; Standard line slopes; HPLC; Tandem mass spectroscopy

1. Introduction

High-performance liquid chromatography (HPLC) with tandem mass spectrometric (MS/MS) detection has been demonstrated to be a powerful technique for the quantitative determination of drugs and metabolites in biological fluids. However, the common perception that utilization of HPLC–MS/MS

guarantees selectivity has been challenged by a number of reported examples of lack of selectivity due to ion suppression or enhancement caused by the sample matrix [1–6] and interferences from metabolites [7–9]. In light of these serious method liabilities, questions about how to develop and validate reliable HPLC–MS/MS methods are being raised. The central issue is what experiments, in addition to the validation data usually provided for bioanalytical methods, need to be conducted to demonstrate the absence of a relative ("lot-to-lot") matrix effect and to confirm HPLC–MS/MS assays selectivity. The current U.S. Food and Drug Administration's (FDA) Guidance for Industry on Bioanalytical Method Validation [10] and a Confer-

* Home address: 1604 Claudia Way, North Wales, PA 19454, USA.
Tel.: +1 215 699 8927.
  E-mail address: bogdan_matuszewski@comcast.net.

ence Report from the workshop held on the same subject [11], clearly indicate the need for the assessment of matrix effect during development and validation of HPLC–MS/MS methods "to ensure that precision, selectivity, and sensitivity will not be compromised" [10,11]. However, in both of these documents the experiments necessary to demonstrate the presence or absence of matrix effect in a given bioanalytical method are not described and/or suggested. In several recent papers, experiments confirming *qualitatively* the presence of matrix effect in biological matrices in comparison with the MS/MS response in neat solvents or HPLC mobile phases were described [12–15] but they do not provide guidance of how to evaluate and determine if an *existing* analytical method or a method under development is selective or suffers from the lack of selectivity due to the effect of matrix.

In our previous publications [4,6] we described strategies that allowed experimental demonstration, during assay development and validation, of the absence or presence of matrix effect in a newly developed bioanalytical method. This information then served as a guide for making changes and corrections, if any, to the original method that would allow the establishment of a truly selective method free of matrix effect interferences. These strategies were illustrated using as an example the experimental data obtained during development of bioanalytical methods for *one* of the drug candidates studied in our laboratories. These strategies included the determination of peak area ratios of a drug to an internal standard in different lots of a biofluid for samples spiked before and after extraction as opposed to multiple determination of these ratios in a single lot, determination of the "absolute" and "relative" matrix effect, and a quantitative assessment of matrix enhancement and ion suppression [4,6]. A similar method based on the assessment of the drug/IS ratios at single concentration in different lots of a biofluid was later applied for the assessment of differential ion suppression for two test compounds in plasma [16]. In addition, in our earlier paper [6] we have suggested, based on the data obtained for a *single* analyte, that determination of slopes of standard curves constructed in different biofluid lots may serve as an indicator of the relative matrix effect. This *relative* matrix effect is of primary concern in quantitative bioanalysis since in everyday practice, the standard curve samples are prepared in a *single* lot of a biofluid, and this curve is used next to determine concentrations of analytes in plasma samples originating from *many* different subjects/patients, at various time points (hours, days, weeks) after dosing, and from different population pools. If the relative MS/MS response is affected by the matrix, the pharmacokinetic data obtained using a bioanalytical method for which the absence of a relative matrix effect was not demonstrated may be erroneous. Therefore, it is of critical importance to study, identify, and eliminate the relative matrix effect in the bioanalytical assay procedures. A simple experimental approach for studying and identifying this effect is described. The method is based on the determination of slopes of standard lines constructed in five different lots of a biofluid during method validation.

In order to unequivocally eliminate the relative matrix effect uncertainty during post-dose sample analyses, standard lines for an analyte(s) in exactly the same post-dose biofluid samples containing the same endogenous compounds *and* metabolites but

without the presence of analyte(s) of interest should be generated. However, construction of such standard lines is impossible in practice. As a workable alternative, by eliminating matrix effect when at least control biofluids from different sources or subjects are evaluated during assay validation, may increase considerably the probability for the method to be much more reliable. In the experiments presented in this paper, assay validations were performed in five different lots of a biofluid instead in a single lot. The major consideration here was not to increase the number of experiments/injections usually performed in typical assay validation experiments but to get some critical insight into method reliability. By eliminating matrix effects in plasmas or urines originating from at least five different sources, the likelihood of providing more accurate bioanalytical and pharmacokinetic (PK) data may dramatically increase.

Using as examples a large number of methods developed in recent years in our laboratories, the relationship between the precision of standard line slopes constructed in five different lots of a biofluid and the reliability of determination of concentration of an analyte in a particular biofluid lot (or subject) was examined. In addition, the precision of standard line slopes was compared when stable isotope-labeled analytes versus analogs were used as internal standards (IS). Also, in some cases, a direct comparison of standard line slopes was made when different HPLC–MS interfaces (APCI versus ESI) were used for the assay of the same compound, using the same IS and the same sample preparation procedure, and under the same chromatographic separation conditions. In selected cases, the precision of standard line slopes in five different lots of a biofluid was compared with precision values determined five times in a single lot. The results of all these studies indicated that the variability of standard line slopes in different lots of a biofluid [precision of standard line slopes expressed as coefficient of variation, CV (%)] may serve as a good measure of a relative matrix effect and that this precision value should be very high for the method to be considered reliable and free from relative matrix effect liabilities.

2. Experimental

2.1. Materials

All compounds studied and their respective internal standards were synthesized at Merck Research Laboratories (Rahway, NJ, and West Point, PA, USA). All solvents and reagents were of HPLC or analytical grade and were purchased from Fisher Scientific (Fair Lawn, NJ, USA). The different lots of drug-free human, heparinized plasma originated from Biological Specialties Corporation (Lansdale, PA, USA). Human urine samples originated from laboratory volunteers. Cerebrospinal fluid (CSF) originated from subjects participating in clinical studies. Nitrogen (99.999%) was purchased from West Point Supply (West Point, PA, USA).

2.2. Instrumentation

Perkin-Elmer (PE) Sciex (Thornhill, Ontario, Canada) API 3000 and Applied Biosystem/MDS Sciex (Concord, Ontario,

Canada) API 4000 tandem mass spectrometers equipped with an atmospheric pressure chemical ionization (APCI, heated nebulizer, HN) or an electrospray ionization (ESI, ion spray, ISP, or turbo-ion spray, TISP) interfaces, a Perkin-Elmer Series 200 (Norwalk, CT, USA), Shimadzu SIL-HTC (Kyoto, Japan) or similar autoinjectors, and a PE 200, Shimadzu LC-10ADVP or similar quaternary pumps were used for HPLC–MS/MS analyses. The data were processed using MacQuan or Analyst™ software (PE Sciex).

2.3. Standard solutions

Stock solutions of standards were prepared in mobile phases. These stock solutions were then diluted further with the mobile phase to give a series of working standards that were used for the preparation of standard curve samples. A similar procedure was used for the preparation of stock solutions of internal standards. These stock solutions were serially diluted with the mobile phase to yield working standards that were used for spiking all standard curve samples.

2.4. Chromatographic conditions

Chromatographic separation was performed on a variety of analytical columns with mobile phases consisting of acetonitrile (ACN) and water containing formic acid and/or ammonium acetate pumped, depending on the column diameter, at a flow rate of 0.2–1 mL/min. The chromatographic conditions were adjusted in such a way that the capacity factors (k') of analytes and internal standards in all cases studied were $k' > 2$. When the direct comparison between APCI and ESI interface was made, the total eluent from the column (for example, 1 mL/min) was directed to the APCI interface, whereas in the case of the ESI interface, the effluent was split (for example, 95:5) and the flow directed to the ESI interface was 50 µL/min.

2.5. HPLC–MS/MS conditions

A PE Sciex triple quadrupole mass spectrometer (Sciex API 3000 or API 4000) was interfaced via Sciex HN, ISP or TISP probes with the HPLC system. The HN probe was maintained at 500 °C and gas phase chemical ionization was effected by a corona discharge needle using positive and/or negative ion atmospheric pressure chemical ionization (APCI). The nebulizing gas (N_2) pressure was set for the HN and ISP interfaces at 80 and 40 psi, respectively. The curtain gas flow (N_2) was 0.9 L/min, and the sampling orifice potential and other mass spectrometric compound-specific parameters were optimized for each compound. The dwell time was 250–400 ms and mass analyzers Q1 and Q3 were operated at unit mass resolution. The mass spectrometer was programmed to admit the protonated $[M + H]^+$ or deprotonated $[M - H]^-$ molecules of analytes via the first quadrupole filter (Q1). Collision induced fragmentation at Q2 yielded the product ions at Q3 that were monitored. Peak area or height ratios obtained from selective reaction monitoring of

the analytes versus respective internal standards were utilized for the construction of calibration lines, using weighted ($1/x^2$) linear least-square regression of the plasma concentrations and measured peak area ratios. Samples prepared in the first lot of a biofluid were injected first followed by samples from consecutive lots.

2.6. Sample preparation

Samples for construction of standard lines were prepared in *five different lots* of a biofluid and were used to evaluate the assay accuracy, precision, and the absence or presence of a "relative" matrix effect. This approach was different than a common practice of evaluating assay precision and accuracy that involves repeat analyses ($n = 5$ at selected concentrations on the standard line) of standard curve samples prepared in the *same, single lot* of a biofluid. By comparing slopes of standard lines between these five different sets, the absence or presence of a "relative" matrix effect on the quantification of analytes was assessed. Slopes of standard lines were determined from the linear regression analysis of the peak area (height) ratios of drug/IS versus analyte concentrations. Since standard curve samples were prepared by spiking a biofluid *before* extraction, the peak areas of drug and the IS reflected overall "process efficiency" (PE) of the procedure [6], a combination of efficiency of analyte recovery and the effect of matrix on ionization. Although the absolute peak areas (heights) of analytes at the same concentration may be different in different biofluid lots (due to differences in extraction efficiency and/or matrix effect on ionization or both), the ratios of drug/IS in different biofluid lots (and slopes of the standard lines derived from these ratios) should not be affected. Therefore, slopes of standard lines in different biofluid lots may serve as a good measure of the "relative" matrix effect (defined here as the combination of the effect of matrix on both recovery of analytes from different lots *and* ion suppression or enhancement between different lots).

Different sample preparation procedures were utilized for various compounds and included solid phase (SPE) and liquid–liquid extraction (LLE) in a 96-well format. These procedures were partially automated using Tomtec Quadra 96 (Hamden, CT, USA) and Packard Multiprobe Robotic Systems (Downers Grove, IL, USA). When comparison between different interfaces was made, samples were analyzed first using the ESI interface and as soon all samples were analyzed, the same samples were injected again into the same HPLC–MS/MS system equipped with the HN interface.

2.7. Precision, accuracy and recovery

The precision of the method was determined by the replicate analyses ($n = 5$, in five different lots) of a biofluid containing an analyte at all concentrations utilized for the construction of calibration curves. The linearity of each standard curve was confirmed by plotting the peak area ratio of an analyte/IS versus drug concentration. The accuracy of the method was expressed by [(mean observed concentration)/(spiked concentration)] × 100.

2.8. Selectivity and "cross-talk" effect issues

Selectivity of all methods was confirmed by analyzing control biofluid blanks from six different sources without the presence of internal standards. No response in MS/MS channels used for monitoring both drugs and internal standards was observed. In addition, the absence of "cross-talk" between channels used for monitoring a drug and an IS was confirmed by injecting separately samples containing an IS at the concentration used in the assay and monitoring the response in the MS/MS channel used for detecting the drug, and by injecting samples containing a drug at the highest concentration on the standard line and monitoring the response in the IS channel. No "cross-talk" was observed in all cases studied.

3. Results

The precision values of standard line slopes constructed in five different lots of biofluids for the 52 methods recently developed in our laboratories for various drug candidates entering clinical development are listed in Table 1 (column A). These values are listed in the descending order from the highest precision methods (small slope precision values) to the lowest precision methods (large slope precision values). In addition, the percent difference between the lowest and highest slope values obtained in these five different biofluid lots for each compound are also included (column B). In column C, the range of precision values, calculated from the peak area (height) ratios of drug/IS in five different biofluid lots at all concentrations on the standard line are also tabulated. Entries in "bold" highlight cases when stable isotope-labeled internal standard was used in the assay. In all other cases, analogs of compounds under study were used as internal standards. When direct comparison was made between different interfaces used for analysis of samples from otherwise the same method, the entries in Table 1 for these cases are underlined. The type of interface utilized (column D) is indicated only in cases when an ESI (ISP, TISP) interface was used. In all other cases, the APCI (HN) interface was applied. Only selected additional details that are relevant to the discussion of the relative matrix effect, for example when measurements were made in the same, single lot of a biofluid instead of in five different lots (entries in "italics"), are included in Table 1. For simplicity, all other experimental details of methods are omitted. The vast majority of compounds studied were early developmental candidates of proprietary nature and their chemical structures cannot be disclosed at this time. However, for the purpose of the discussion of the relative matrix effect, the knowledge of chemical structures of compounds under study is not of critical importance.

4. Discussion

4.1. General observations

For the vast majority of methods developed in our laboratories and listed in Table 1 (entries 1–45) the precision (CV, %) of standard line slopes constructed in five different lots of a

biofluid was high and did not exceed the value of 3.4%. Similar data for methods described in the literature is not available since in all cases reported, validation experiments (precision and accuracy determination) are based on the repeat analysis in the same, single biofluid lot. The reason for high precision values of standard line slopes in methods developed in our laboratories is largely due, in my opinion, to our early awareness [4–6] of the need to develop methods free from relative matrix effect liabilities. As recommended earlier [4,6], we routinely try to develop methods that require an effective extraction of analytes from biological fluids using SPE or LLE extraction and avoid highly ineffective, "generic" sample extraction procedures, dilute-and-shoot approaches, acetonitrile precipitation and injection of the supernatant techniques, etc. In addition, all methods listed in entries 1–45 (Table 1) required an effective HPLC separation and retention of all analytes on the HPLC columns with the retention factor (k') of at least $k' > 2$. Under such conditions, the likelihood of observing both an absolute and a relative matrix effect is minimized [6]. In addition, in 17 out of 45 "good" cases (entries 1–45), the stable isotope-labeled (SIL) internal standard was used. The utilization of SIL internal standards effectively eliminates any relative matrix effect liability with the assumption that isotopic purity and stability issues of SIL standards are addressed. Also, in the majority of cases, an APCI interface (39 compounds) rather than the ESI interface (6 compounds) was used. The APCI interface is known to be less prone for exhibiting a matrix effect in comparison with ESI-type interfaces [4,6,13,17,18].

All assays listed in entries 1–45 are considered to be fully adequate for supporting long-term human pharmacokinetic studies and may be considered as free from relative matrix effect issues. In addition to high precision values of standard line slopes (<3.4%) constructed in five different lots of a biofluid (column A, Table 1), the precision values at all concentrations used for the preparation of standard curves and determined in five different lots of a biofluid, did not exceed 8.7% (column C, Table 1). Our internal validation procedures are significantly more demanding than recommended in references [10,11] and require the CV values at all concentrations on the standard lines *in five different lots* of a biofluid to be <10% (<15% at the lower limit of quantification, LLOQ) as opposed to <15% (<20% at LLOQ) in a *single lot* [10,11]. The comparison of data in columns A and C (Table 1) indicates that the CV values listed in column C may not reflect fully the extent of a relative matrix effect. No correlation between the increase in values in column A and a range of CV values listed in column C is apparent. It was suggested earlier [4,6,16] that the large CV values in different lots of a biofluid may be indicative of matrix effect, but these values may also reflect the overall reproducibility of various sample preparation steps utilized in the assay, the performance and ruggedness of the HPLC system, and other assay method variables. For example, the assay CV range (column C) at different concentrations was the same in entry 7 (1.1–7.9%) as in entry 43 (2.1–7.5%) whereas the CV of the standard line slope (column A) was significantly higher (3.2) in entry 43 than in entry 7 (0.7). This may indicate that the determination of CV values in different lots of a biofluid, especially at a single concentration on the standard line [16] may not be a good and

Table 1
Summary of bioanalytical method validation data for selected compounds under development

Compound (entry) #	A Slopes CV (%)[a]	B Slope difference (%)[b]	C Assay CV range (%)[c]	D Other details[d]
1	0.19	0.5	0.5–4.7	U, LLE
2	**0.40**	**0.9**	**0.8–2.7**	**P, LLE**
3	0.40	1.1	1.0–3.2	**P**
4	**0.45**	**0.9**	**0.7–2.7**	**U, TISP, LLE**
5	<u>0.53</u>	<u>1.3</u>	<u>0.7–5.4</u>	<u>U</u>
6	**0.7**	**1.7**	**1.1–5.7**	**U, TISP, LLE**
7	**0.7**	**1.7**	**1.1–7.9**	**P, LLE**
8	**0.8**	**2.1**	**0.9–3.5**	**U, TISP, SPE**
9	0.80	1.9	0.9–3.0	**P**
10	0.90	2.2	0.5–3.8	U, LLE
11	0.93	2.3	0.6–3.1	P, SPE
12	0.94	2.3	2.5–4.7	P
13	**1.0**	**1.9**	**1.7–6.1**	**P, LLE**
14	1.0	2.3	1.4–3.5	P, classical LLE
15	1.0	2.5	0.8–3.2	P, LLE
16	1.0	3.0	1.3–8.5	P, LLE
17	1.2	3.0	1.5–2.9	P
18	**1.2**	**2.7**	**0.9–7.8**	**P, LLE**
19	**1.2**	**3.3**	**1.3–3.8**	**P, LLE**
20	**1.3**	**3.0**	**0.9–6.0**	**S**
21	**1.3**	**3.6**	**0.5–2.1**	**U, DI**
22	1.4	2.9	0.8–4.1	U
23	1.5	4.4	0.9–8.0	P, LLE
24	**1.6**	**4.6**	**1.6–6.9**	**P, LLE**
25	1.6	4.3	1.8–6.8	P, LLE, TISP
26	1.6	4.8	1.7–6.9	U, LLE, TISP
27	**1.7**	**4.2**	**0.7–3.6**	**P**
28	1.8	4.0	1.1–3.3	P
29	**1.8**	**4.7**	**1.6–4.2**	**U, TISP, LLE**
30	1.8	4.5	1.0–5.5	U, LLE
31	**1.9**	**3.7**	**1.8–8.1**	**U, LLE, DI**
32	1.9	5.1	2.1–3.7	U
33	**2.0**	**5.5**	**1.2–4.2**	**P, TISP, LLE**
34	2.0	6.0	3.0–4.9	P
35	2.2	5.3	1.2–3.0	P
36	**2.3**	**4.8**	**1.5–8.1**	**U, LLE**
37	2.3	5.6	3.1–7.4	U
38	*0.31*	*0.8*	*1.7–6.9*	*CSF, SL*
39	**2.4**	**5.0**	**2.4–6.4**	**P, TISP**
40	2.5	6.6	2.0–3.7	P, LLE
41	2.6	5.8	2.0–5.1	P, LLE
42	2.9	8.0	2.2–4.6	U
43	3.2	7.3	2.1–7.5	U
44	3.3	8.0	5.1–8.7	U
45	3.4	9.8	2.6–5.4	U, SPE
46	<u>4.6</u>	<u>11.5</u>	<u>4.1–7.3</u>	<u>U, TISP</u>
47	5.0	15.0	5.8–8.2	P
48	6.5	18.3	1.5–7.7	U, DI
49	8.7	20.5	4.0–9.3	P, TISP[e]
50	13.2	34.3	11.1–27.8	P, ISP[f]
51	*2.4*	*6.5*	*4.2–8.5*	*P, ISP, SL[f]*
52	16.1	46.3	7.5–14.9	P, TISP[e]

[a] Precision value (coefficient of variation, CV) of slopes of standard lines constructed in five different lots of a biofluid.

[b] Maximum difference between the highest and the lowest slope values divided by the lowest slope value and multiplied by 100.

[c] Range of coefficient of variation values (method precision) determined at all concentrations used for constructing standard lines.

[d] The following abbreviations were used: P: plasma; U: urine; S: serum; ESI: ISP or TISP interface; SPE and/or LLE: solid phase and/or liquid–liquid extraction in 96-well format; SL: single lot of a biofluid; DI: direct injection, no sample extraction; CSF: cerebrospinal fluid.

[e] Exploratory assay, four different plasma lots, validation data limited.

[f] Reported earlier [6].

sensitive indicator of the presence and/or absence of a relative matrix effect. In order to eliminate method variability reflected in the CV values obtained at a *single, representative* concentration in five different lots [16], these CV values should rather be compared with the analogous CV values obtained by repeating the analysis ($n = 5$) in a single lot. Any difference between these two CV values may be indicative of the contribution of the relative matrix effect to the CV values obtained in different lots. Also, the CV values determined in different biofluid lots but at a *single* concentration may be highly concentration dependent. In addition, the reproducibility of determination of CV values at a *single* concentration may be variable on a day-to-day basis. Instead, the determination of CV values of standard line slopes encompassing all concentrations within the standard curve range is quite reliable and reproducible. These standard line slope CV values determined in different biofluid lots may constitute the best indicator of the presence/absence of the relative matrix effect in a bioanalytical method.

4.2. Range of standard line slope values in different lots of a biofluid

The difference in slope values for standard curves prepared in different lots of a biofluid (column B, Table 1) represents a difference between concentration obtained in an assay when an analyte present in one lot of a biofluid (for example, originating from a subject participating in a clinical study) is analyzed and its concentration calculated using a standard line prepared in a different biofluid lot (control, blank). This "slope difference" (column B) corresponds to the *maximum* difference in the calculated concentration of an analyte in five lots studied that originates from the relative matrix effect. In an "ideal" situation, and in the absence of a relative matrix effect, the slopes of standard lines constructed in different biofluid lots should be the same. The larger the values in column B are, the more pronounced the relative matrix effect becomes. For example, the difference in slope value 9.8% (entry 45, column B) indicates that due to a relative matrix effect, the concentration of an analyte in one lot (for example lot "a") is 9.8% higher/lower than in a different lot (lot "b") in which the standard line was constructed, *even if the concentrations of an analyte in both lots was the same*. Such a difference (9.8%) may already be of some significance in the overall PK evaluation of the drug (drug interaction studies, comparison of different subject groups, formulation comparison, etc.). Larger slope difference values (11.5–46.3%), as listed in entries 46–52 (Table 1), are clearly indicative of the presence of a significant relative matrix effect, and as such, methods exhibiting such relatively high slope differences should not be recommended for use in support of human PK studies. For example, if method in entry 50 was used in support of an interaction study, the 34.3% difference (column B) in AUC values between two different drug treatments may have nothing to do with drug interaction but may originate from the relative matrix effect due to the difference in the chemical make up of the biofluid between two treatments. Inspection and comparison of columns A and B (Table 1) clearly indicate that the smaller the CV value of standard line slopes is, the smaller is the difference in slope values and the method is

more reliable and free from relative matrix effect liability. The currently proposed cut-off value of <3–4% (CV of standard line slopes in different lots of a biofluid, column A) for acceptance of an analytical method as practically free from a significant relative matrix effect liability is based on comparison of entries 1–45 versus 46–52 in columns A and B (Table 1). For example, at 5% CV (entry 47, column A), the difference in the determined concentration solely due to the matrix effect may be as high as 15%, and even higher (18.3–46.3%) at higher slopes CV values (8.7–16.1%) (entries 48–50, 52, columns A and B). Based on the data obtained for a *single* analyte [6], we have proposed earlier a cut-off value of <4–5% for the method to be acceptable. However, now, based on a large number of assay data (Table 1) accumulated and presented in this paper for 52 different methods, the <3–4% cut-off value may be considered as more appropriate.

4.3. Comparison of precision (CV) of standard line slopes constructed in different lots versus repeated analysis in a single lot

The comparison between CV values of standard line slopes obtained in five different biological fluid lots with the CV of similar five slopes obtained in a single lot may serve as an excellent indicator of the relative matrix effect [6]. This is illustrated by comparing data presented in entries 50 and 51 (Table 1). In both cases, the same method in plasma (extraction, chromatography, the same MS/MS interface) was used for the same compound. The only difference was that in entry 50 the standard line slopes were generated in five different lots of plasma whereas in entry 51 they were generated in single lot ($n = 5$). The considerable increase in the CV of standard line slopes (column A) from 2.4% (in a single lot) to 13.2% (in five different lots) together with a significant increase in values listed in columns B and C are indicative of the presence of a significant relative matrix effect. Clearly, a method characterized by such parameters needs to be modified until values in columns A–C obtained in five different lots are similar to those obtained in a single lot. As a simpler alternative, instead of comparing values in columns A–C in single versus multiple sources of a biofluid, the CV values of slopes in different sources of a biofluid (column A) may be determined and a method needs to be modified until the slope CV values are in the recommended range of <3–4%.

Another indirect comparison of CV slopes values in a single versus multiple sources of a biofluid is presented in entries 37 and 38 (Table 1). The same method was used for the determination of a compound but in one case (entry 38) the slope CV values were obtained in a single lot of CSF, whereas in the other case (entry 37) slope CV values were obtained in five different lots of urine. The high precision (small CV values, 0.31%) of slopes obtained in a single lot of CSF (entry 38) is clearly indicative of the excellent method reproducibility, whereas a slightly higher value obtained in five different lots of urine (2.3%, entry 37) may be reflective of the presence of some, albeit insignificant, relative matrix effect. Analogous data in a single lot of urine was not generated since the relative matrix effect (entry 37) based on the analysis in five different lots was considered minor

(maximum difference in standard line slopes was 5.6%, column B). In the case of CSF (entry 38) analyses were performed in a single lot due to the difficulties in obtaining CSF samples from multiple sources (subjects).

4.4. Use of stable isotope-labeled internal standard

In 17 (44%) of methods (Table 1) characterized by the high precision (≤2.4%) of the slope values in different biofluid lots (entries 1–39), SIL analytes were used as internal standards. When SIL internal standards were utilized, the precision of standard line slopes (column A) in different lots varied from 0.4% (entry 2) to 2.4% (entry 39). These data clearly indicated that the utilization of stable isotope-labeled internal standards effectively eliminated relative matrix effect liability. However, in all cases, the issues of isotopic purity of compounds, "cross-contamination" or "cross-talk" between MS/MS channels used for monitoring the drug and IS, and isotopic stability of an IS (absence of isotope exchange) needed to be addressed [19].

4.5. Choice of an interface and a relative matrix effect

As reported earlier [4,6], the utilization of an APCI instead of a ESI interface may lead to the elimination of, or a decrease of the matrix effect. Even when analogs were used as internal standards, in practically all cases presented in Table 1, a relative matrix effect was not observed (entries 1–24, 27–45) when the APCI interface was utilized. On the other hand, methods that used ESI (ISP or TISP) interfaces (entries 46, 49–52) exhibited a well-pronounced relative matrix effect. Clearly, these latter methods require significant changes (more efficient and selective sample preparation procedures, more efficient chromatography, utilization of a SIL internal standard) before they are applied to support clinical PK studies. Inspection of data in Table 1 also indicates that SIL internal standards were used in practically all "good" (entries 1–45) methods (except two) that employed an ESI (TISP) interface. Due to the greater propensity for exhibiting matrix effect by ESI versus APCI interfaces, the need for SIL internal standards appears to be much higher when ESI rather than APCI interfaces are utilized. The reason for the observed differences in ion suppression/enhancement between APCI and ESI interfaces is unclear but it is probably due to a different mechanism of ionization (ion evaporation from the liquid droplet in ESI versus gas phase ionization in APCI) [1,13,20]. It has to be emphasized that both ESI and APCI interfaces may produce reliable methods. The major point is that irrespective of the interface utilized, the absence of a "relative" matrix effect needs to be confirmed and the approach proposed in the paper may be highly useful in making that determination.

4.6. Additional comments

In the course of utilization of a method for long-term bioanalytical support, hundreds or even thousands of different subjects may participate in these studies and the molecular content of their plasmas and/or urines (for example) may be widely different. The determination of standard line slopes in widely different biofluid lots originating from multiple population (when available) should be considered in assay validation experiments. As an alternative, these determination may be performed using control biofluid lots from normal volunteers and later compared with standard line slopes constructed in biofluid lots from other subject groups (healthy versus sick, young versus old, subjects on different diets, etc.). If these slopes are practically the same, the "relative" matrix effect may be considered as absent or negligible. Again, by eliminating matrix effects in plasmas or urines originating from at least five different sources, the likelihood of providing more accurate bioanalytical and PK data may dramatically increase.

5. Conclusion

The data accumulated clearly indicate that the determination of standard line slopes in different sources (lots) of a biofluid may serve as a good *quantitative* indicator of the presence/absence of a relative matrix effect in bioanalytical methods. Based on the data presented, it is proposed that the precision (CV) value of standard line slopes constructed in five different lots of a biofluid should not exceed 3–4% for the method to be considered practically free from the relative matrix effect liability. Comparison of the CV values of standard line slopes in five different lots of a biofluid with the analogous values obtained by repeated analysis ($n = 5$) in a single lot may also serve as an excellent measure of a relative matrix effect. Such experiments clearly fulfill the requirements of the FDA Guidance [10,11] requiring the assessment of "relative" matrix effect in bioanalytical methods. In addition, comparison of standard line slopes determined in different subjects/population under study may serve as an excellent measure of the relative matrix effect between different subjects/patients.

Use of stable isotope-labeled analogs as internal standards effectively eliminates relative matrix effect liability and, when available, is highly recommended. In all other cases, a careful choice of the MS interface (APCI versus ESI), sample preparation conditions, and effective HPLC separation is needed to avoid relative matrix effect complications.

References

[1] P. Kebarle, L. Tang, Anal. Chem. 65 (1993) 972A.
[2] S.D. Clarke, H.M. Hill, T.A.G. Noctor, D. Thomas, Pharm. Sci. 2 (1996) 203.

[3] D. Buhrman, P. Price, P. Rudewicz, J. Am. Soc. Mass Spectrom. 7 (1996) 1099.

[4] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Anal. Chem. 70 (5) (1998) 882.

[5] I. Fu, E.J. Woolf, B.K. Matuszewski, J. Pharm. Biomed. Anal. 18 (1998) 347.

[6] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Anal. Chem. 75 (2003) 3019.

[7] M.L. Constanzer, C.M. Chavez, B.K. Matuszewski, J. Carlin, D. Graham, J. Chromatogr. B 693 (1997) 117.

[8] B.K. Matuszewski, C.M. Chavez, M.L. Constanzer, J. Chromatogr. B 716 (1998) 195.

[9] M. Jemal, E.-Q. Xia, Rapid Commun. Mass Spectrom. 13 (1999) 97.

[10] Food and Drug Administration, Guidance for Industry on Bioanalytical Method Validation, Federal Register, vol. 66, no. 100, 23 May 2001, p. 28526.

[11] V.P. Shah, K.K. Midha, J.W.A. Findlay, H.M. Hill, J.D. Hulse, I.J. McGilveray, G. McKay, K.J. Miller, R.N. Patnaik, M.L. Powell, A. Tonelli, C.T. Viswanathan, A. Yacobi, Pharm. Res. 17 (12) (2000) 1551.

[12] R. Bonfiglio, R.C. King, T.V. Olah, K. Merkle, Rapid Commun. Mass Spectrom. 13 (1999) 1175.

[13] R. King, R. Bonfiglio, C. Fernandez-Metzler, C. Miller-Stein, T. Olah, J. Am. Soc. Mass Spectrom. 11 (11) (2000) 942.

[14] B.K. Choi, D.M. Hercules, A.I. Gusev, J. Chromatogr. A 907 (2001) 337.

[15] B.K. Choi, D.M. Hercules, A.I. Gusev, Fresenius' J. Anal. Chem. 369 (2001) 370.

[16] M.J. Avery, Rapid Commun. Mass Spectrom. 17 (2003) 197.

[17] J. Schuhmacher, D. Zimmer, F. Tesche, V. Pickard, Rapid Commun. Mass Spectrom. 17 (2003) 1950.

[18] T. Sangster, M. Spence, P. Sinclair, R. Payne, C. Smith, Rapid Commun. Mass Spectrom. 18 (2004) 1361.

[19] C.M. Chavez-Eng, M.L. Constanzer, B.K. Matuszewski, J. Chromatogr. B 767 (2002) 117.

[20] P. Kebarle, J. Mass Spectrom. 35 (2000) 804.