ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Pharmaceutical and Biomedical Analysis
35 (2004) 887–893

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

www.elsevier.com/locate/jpba

Bioanalytical method validation: a risk-based approach?

C.A. James*, M. Breda, E. Frigerio

Pharmacokinetics, Dynamics and Metabolism, Pharmacia Italia S.p.A., Gruppo Pfizer Inc., Viale Pasteur 10, 20014 Nerviano, MI, Italy

Received 23 December 2003; received in revised form 1 February 2004; accepted 16 February 2004

Available online 10 April 2004

Abstract

Bioanalysis frequently involves the measurement of very low analyte concentrations in complex and potentially variable matrices. It is not possible to test in validation every possible circumstance that may be encountered when analyzing study samples; logically, therefore, some risk of obtaining erroneous results exists when validated methods are applied to study samples. An initial attempt has been made to apply a risk management tool to the bioanalytical situation, with the hope that this will stimulate further discussion on the idea of more formally addressing "risk" with regards to bioanalytical method validation.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Bioanalysis; Validation; Risk assessment

1. Introduction

Bioanalysis, perhaps more that many other types of analysis, involves unknown factors as analyte concentrations are low, and matrices are complex and may contain unknown and variable components. The reliability of modern methods is therefore remarkable given that analytes are often measured at concentrations of one in a billion or less, and even after extraction and chromatography the analyte of interest may often represent only one ion in a million entering the interface of an LC–MS–MS instrument.

The industry consensus on bioanalytical validation, reflected in FDA guidelines and other publications [1–4] has done much to improve the quality of bioanalysis and to establish generally accepted validation specifications. However, almost all quantitative acceptance criteria applied to validation and quality control of bioanalytical methods rely on data generated using spiked control biofluid samples. Whilst the limitations of such samples as a model for "real" samples have been acknowledged for many years, "spiked samples" remain a mainstay for validation and quality control in bioanalysis. Whilst the potential for ion-suppression/enhancement (matrix) effects with LC–MS–MS methods was described as long ago as 1993 [5,6], recent publications [7–12] have emphasized the possibility for such effects to result in major and unseen errors in LC–MS–MS bioanalysis, due primarily to differences between sources of biological matrix, and between the components of "real" study samples and spiked samples used to prepare calibration and QC samples. Whilst ion-suppression effects are not the only way in which spiked calibration and

* Corresponding author. Tel.: +39-024-838-3196; fax: +39-024-838-3012.
E-mail address: chris.a.james@pfizer.com (C.A. James).

0731-7085/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2004.02.028

QC samples may be an unreliable model for study samples, such results have served to remind us of the potential limitations of bioanalytical validation experiments. These potential problems appear to relate largely to the complexity and potential variability of biomatrix samples. In addition to various endogenous substances [13], biological matrices may contain drug metabolites, co-administered drugs and/or metabolites, anti-coagulants, formulation excipients and dietary substances and supplements [14]. The presence and/or concentration of matrix components may vary in individual subjects, in different races, or patient groups, or due to the circumstances of different studies; disease state and the pharmacological action of the drug or co-medications may also alter the matrix.

Given that it is impossible to examine every eventuality that might be met in future application of a bioanalytical method during validation, even "highly validated" methods must logically carry some risk of generating erroneous results when applied to study samples. Whilst these risks may be implicitly recognized, they are not discussed explicitly in validation guidelines. The following article therefore attempts to apply a risk management approach to bioanalytical validation. Encouragement to apply such an approach has been taken from the FDA announcement of a new "risk-based approach for cGMP" [15,16] even if there is no direct indication of the applicability of this approach to bioanalysis at this time.

2. Ion-suppression

In contrast to bioanalytical methods using HPLC with detectors such as UV and fluorescence, LC–MS–MS has been perceived as a highly selective technique in which interfering peaks are rarely seen. This has encouraged analysts to increase sample throughput [17,18] often by minimizing sample preparation procedures and running very rapid chromatographic separations with limited retention of the analyte. However, as indicated in the introduction, increasing reports of matrix (ion-suppression/enhancement) effects are changing the perception of LC–MS–MS as a highly selective bioanalytical technique, to one that is selective for "visible" interferences, but that may be non-selective for "unseen" interferences [8]. Fig. 1 gives an example of an ion-suppression effect due to a formulation component in rat plasma study samples, which was resolved by amending the chromatographic conditions to separate the compound from the formulation components (P. Larger, personal communication). In early time-point samples, this resulted in a reduction in the measured analyte concentration of several fold. Fig. 2 shows single reaction monitoring (SRM) and total ion current (TIC) traces corresponding to this example. With the isocratic system, the TIC plot (Fig. 2, plot C) is clearly at baseline levels where the analyte peak elutes. However, with the gradient elution system where ion-suppression was observed,

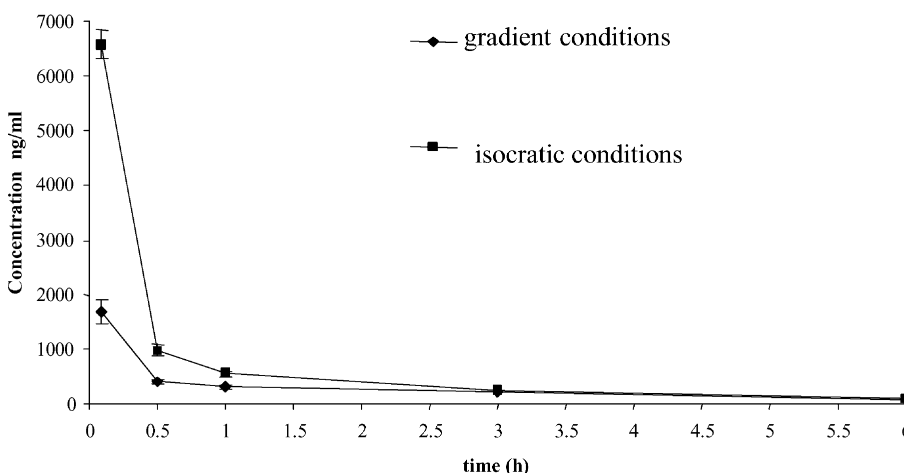

Fig. 1. Plasma concentration–time (mean, $n = 3$) profiles following IV administration of a compound in a formulation containing Tween 80 to rats. Samples were analyzed by two methods, the method using gradient elution suffers from ion-suppression.

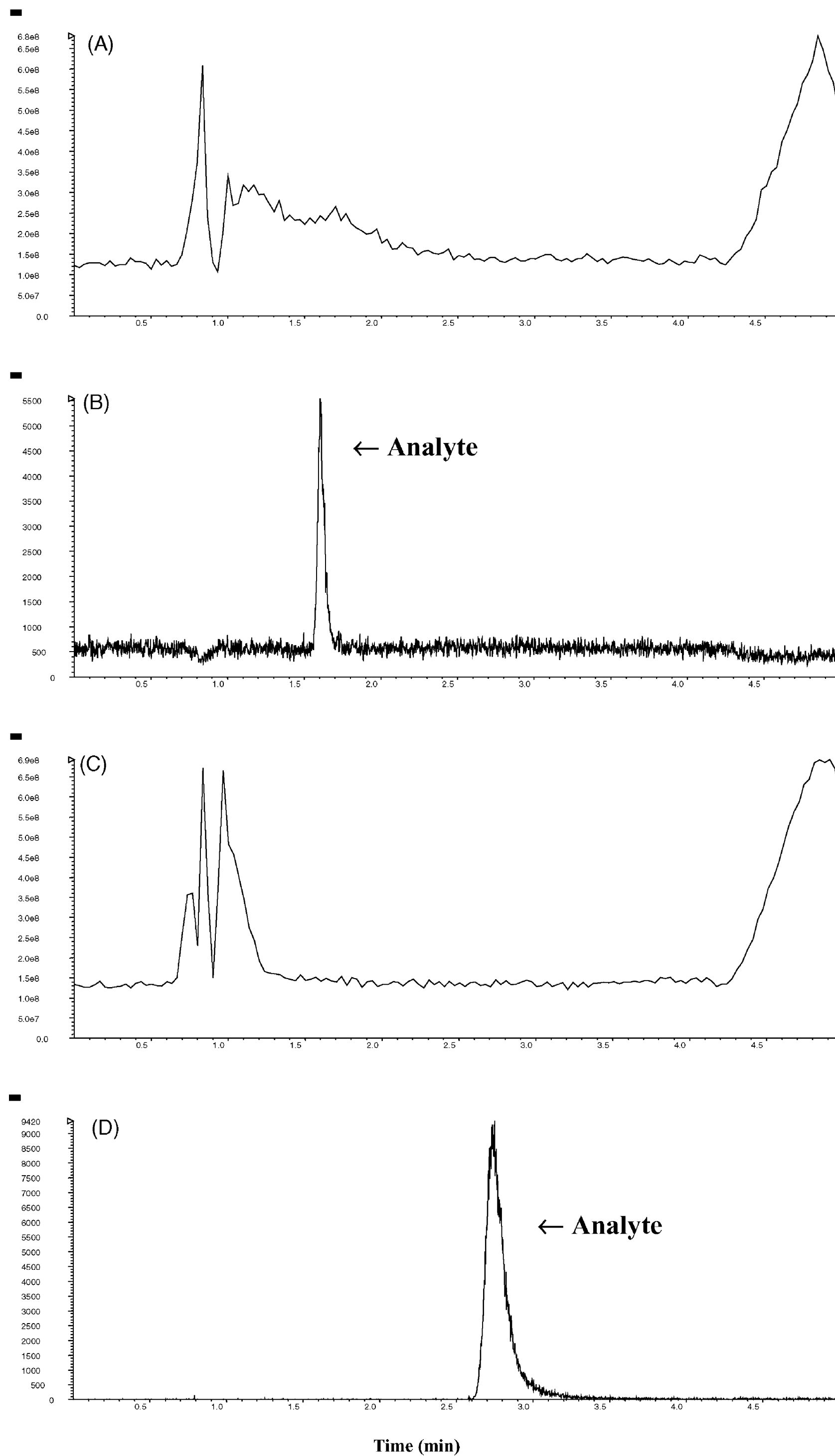

Fig. 2. Time–intensity plots corresponding to methods used for the example in Fig. 1; (A) the TIC plot and (B) the SRM plot for the gradient elution method; (C) the TIC plot and (D) the plot SRM for the isocratic method.

the TIC plot (Fig. 2, plot A) appears to show elution of matrix components in the region in which the analyte peak elutes. In addition to this example, literature reports [10–12] also indicate the possibility of relatively large errors due to ion-suppression effects. With regards to these examples at least, the magnitude of potential error appears to be far greater than the levels of imprecision that cause an assay to fail according to standard QC criteria (e.g. >15–20% of target). Whilst validation guidelines [2] indicate "matrix effects should be investigated" for LC–MS or LC–MS–MS methods, no detail is given of how this should be done. In the literature two approaches for the identification of matrix effects are generally reported, involving either post-column infusion of the analyte [10,19–21] or comparison of responses from various analyte solutions and spiked samples [8,11,12]. With the infusion approach, the analyte is continuously infused into the LC–MS–MS system and injections of blank matrix extract are made, allowing the identification of areas of the chromatogram where elution of endogenous materials might significantly alter the response from the analyte. Whilst this can be useful to avoid potential areas of significant interference in the chromatogram, it does not yield quantitative information under the exact conditions of the assay. In our laboratories, we have therefore adopted the approach described in detail by Matuszewski et al. [8] of comparing the response of analyte solutions with spiked extracted biofluid samples, and extracts of control biofluids spiked with analyte after extraction. This allows the calculation of matrix effect, extraction recovery, and overall "process efficiency". A simplified check is also suggested comparing responses of five different sources of biofluids spiked with analyte, recognizing that ion enhancement or suppression may be present but not affect the assay if relatively constant in all samples, or if fully compensated for by an internal standard. This later check is also useful for online extraction systems where the absence of a separate "extract" precludes experiments requiring the post-extraction spiking of samples. In his discussion on how many sources of control biofluid to examine for matrix effect, Matuszewski et al. [8] in effect describes a risk management approach. Examination of five sources of biofluids is recommended even though this number could be seen as inadequate for clinical methods that will be applied to samples from hundreds of individuals. However, this recommendation is practical, and compared with examining only one source of biofluid "the likelihood of providing more accurate bioanalytical and PK data may dramatically increase" [8].

3. Risk management tool

Fig. 3 is a risk management tool adapted for the bioanalytical situation. The diagram attempts to compare the risk of certain problems occurring with the impact on the analysis if that problem has occurred. It is assumed that methods are already validated according to FDA guidelines and standard batch acceptance criteria are being applied in subsequent application of the method. The probability of occurrence (frequency) was judged on how often an event might occur across a series of different methods. It must be emphasized that the diagram is constructed on a subjective basis from experiences within our laboratories and discussions with other bioanalytical scientists. For example, we have judged that problems with linearity (1) are not common with validated methods (risk of occurrence is low), particularly as issues with linearity will generally be discovered and resolved during method development. The impact is also judged as low because, if problems do occur during analysis, they will readily be detected by standard calibration and QC procedures.

In terms of risk assessment, the items falling in bottom left quartile (low-mid risk of occurrence/low-mid impact) of the diagram (Fig. 3), can be regarded as relatively "safe" and not needing improvement in terms of risk management. Items occurring in the mid-high risk or impact areas, or mid-high segments for both, need attention either to reduce the risk of occurrence, or to mitigate their impact if they do occur.

Most of the items grouped in the "low risk" quartile are the "classic" validation parameters, including linearity (1), sensitivity (2), precision and accuracy (3), stability (4), and selectivity (visible interference) (5). One of the main reasons for placing these items as low risk is that spiked samples were judged as effective in evaluating and controlling for these parameters in validation and routine analysis. Also in this quartile were robustness (reproducibility) (6), which was judged as low impact as problems should be readily detected

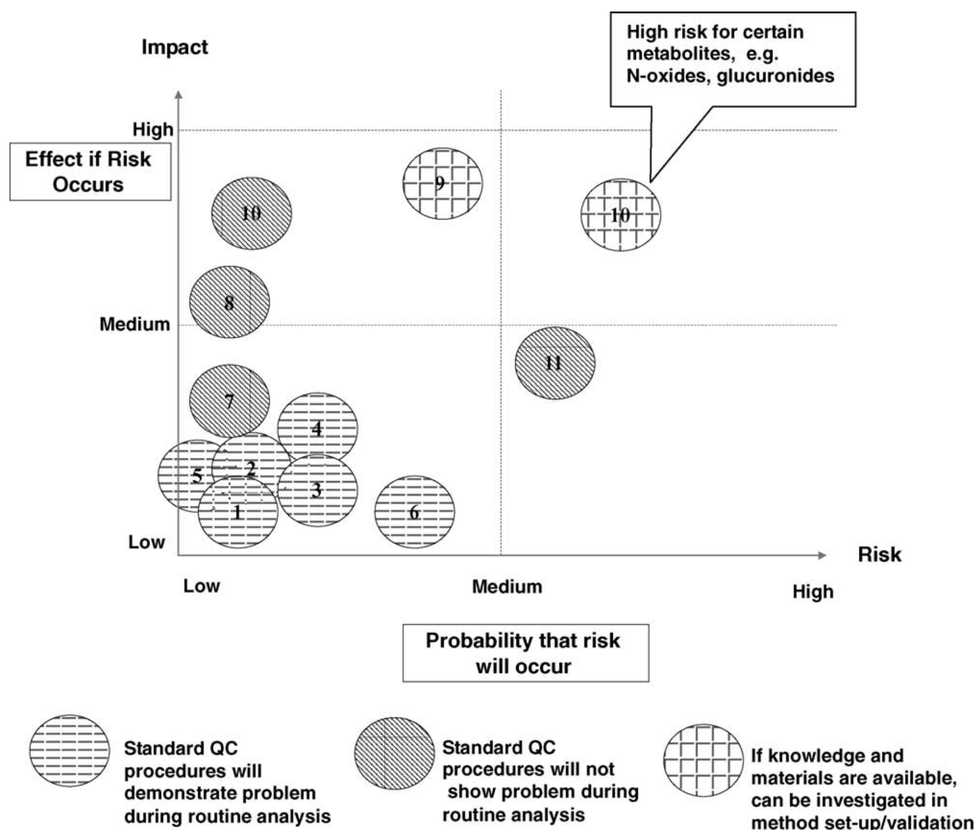

Fig. 3. Risk assessment diagram. Key: (1) linearity, (2) sensitivity, (3) precision and accuracy, (4) stability, (5) selectivity (visible peak), (6) robustness, (7) stability: different in study samples, (8) extraction recovery different in real samples, (9) ion-suppression, (10) stability: metabolite → parent drug, and (11) inter-laboratory cross-validation.

(robustness is in any case not commonly evaluated [22]), and stability: different in real samples (7), which was judged to occur relatively infrequently.

In the higher risk areas of the diagram were: extraction recovery different in real samples (8) and unstable metabolites that revert to another analyte (10); these were judged to occur, in general, fairly infrequently, but the impact of occurrence was judged as relatively high, as the analytical errors in study samples could be large but standard QC procedures would not detect any problem in the analysis. Whilst we judged the risk of occurrence for item (10) to be generally low, when certain types of labile metabolite (e.g. N-oxides, glucuronides) are present [23,24] the risk becomes higher and should be checked for as soon as the availability of materials allows. We have indicated this by showing item (10) twice on the diagram.

Based on our own experiences and the increasing number of recent publications (see Section 1) we judged the probability of occurrence to be somewhat higher for ion-suppression (9). Note the terms ion-suppression and matrix effect are used somewhat inter-changeability in the literature, here we have used the term ion-suppression for convenience to refer to the effect of any co-eluting material that might increase or decrease the ionization efficiency of the analyte in an LC–MS system. A variety of possible causes for ion-suppression have been reported, e.g. different sources of biomatrix, anti-coagulants, metabolites and co-medications, even different suppliers of blood collection tubes [11]. We judged the impact to be potentially high as the magnitude of the effect on the analysis can be large and often calibration and QC procedures will not indicate that such problems are

occurring in the study samples. Of course if the necessary information and materials exist, many of these potential problems listed above can be checked for in method development or validation, and some items on the diagram (Fig. 3) have been coded to indicate this possibility. For example, if available, potential unstable metabolites can be tested for reversion to parent drug, or intravenous formulation excipients can be checked for ion-suppression effects. However, particularly in early drug development, both knowledge and availability of relevant materials may be limited.

The final item in the risk assessment diagram (Fig. 3) is inter-laboratory cross-validation [11]. Whilst not a standard part of every validation exercise, anecdotal evidence suggests that it is relatively common to find differences between the same method run in different laboratories, or between different methodologies, even though the both methods were fully validated. The recent guidance [2] has therefore been helpful in defining exactly when cross-validation exercises are essential.

4. Discussion

Taken overall, the greatest "risks" in bioanalytical methods seem to be those associated with problems where spiked QC samples are a poor model for real samples. As exemplified by recent concerns about ion-suppression effects, approaches to dealing with these potential problems include: extending or re-designing validation experiments to better detect such problems [10,12], seeking instrumental and method conditions that are in general more robust against such effects [8,11] or making QC samples a more relevant model for study samples [27]. Matuszewski et al. [8] indicates a number of approaches to "eliminate" matrix effects, including having adequate chromatographic retention, good sample clean up, choice of LC–MS–MS interface, and use of a stable labeled internal standard. Whilst these comments were written as recommendations to eliminate matrix effects after they have been discovered, they could equally be applied as ways to improve robustness against problems that might occur in future application of the method.

It is often stated that validation should be a continuous process even if most validation exercises are conducted as discrete experiments. Certainly it is important that bioanalysts remain vigilant, particularly as some of the more unusual bioanalytical problems have been identified by chance observation rather than during validation experiments. It is important also to remain updated on emerging information about the analyte (e.g. metabolism) and study conditions (e.g. co-medications) as a method continues to be used, so that additional validation experiments can be run if required. One interesting approach to the on going monitoring of methods has been outlined by Jemal et al. [27]. He suggests using pooled study samples from the first clinical study to monitor long-term performance and stability in subsequent studies to try and identify "real samples" issues relatively early in application of the method.

An implicit assumption in almost all validation experiments is that spiked QC samples are an adequate model for "real" study samples. Clearly this assumption is sometimes incorrect. This possibility can be interpreted in two ways: (a) at worst spiked samples may be an "easier" (less complex) matrix for than study samples; if a method performs badly in validation with spiked samples it is unlikely to show good performance with "more complex" study samples; (b) if spiked samples behave differently to study samples they are an irrelevant model for validation, and any validation data generated with them is worthless.

Even if we accept spiked QCs as a good model, a number of authors have questioned the appropriateness and consistency [25,26,28–30] of the recommended [2,3] procedures and acceptance limits applied to validation and routine analysis batches. One important detail of the recommended batch acceptance limits was "tightened" from 20 to 15% (i.e. for batch acceptance 66% of QC samples need to be within 15% of their target values) following the workshop held in Washington in 2000 [1–3]. In attempting to assess risks faced in application of bioanalytical methods, albeit in a subjective fashion, our attention has been drawn not to the detail of these various acceptance procedures, but to the risks of quite major unknown errors occurring during analysis of study samples due to factors such as ion-suppression or labile metabolites. These concerns match those of other authors in recent publications [8,27].

Our proposal is that we need to acknowledge some aspect of risk will exist in application of

bioanalytical methods however many validation experiments are performed. A science-based risk management approach could therefore help to focus experimental work and possibly to rethink our validation strategies. Perhaps we need to place greater emphasis, especially in early application of a method, on gaining experience with analysis of incurred study samples, and monitoring on going performance of the method in actual application. The scope of the extensive pre-study validation experiments, now usually performed with spiked samples, could be reconsidered together with the acceptance limits that are applied.

5. Conclusions

Even with the best validation approaches, diversity and unknown factors in study samples pose a threat of unseen, and undetected, errors in bioanalysis and there is no simple way to guard against or check for every possibility. The objective of this article was therefore to suggest that a science-based risk management approach should be considered for bioanalytical validation, with the hope that this will stimulate further discussion on the idea of more formally addressing "risk" in validation approaches and guidelines.

References

[1] 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 (2000) 1551–1557.

[2] Guidance for Industry: Bioanalytical Method Validation, US Department of Health and Human Services, Food and Drug Administration, CDER, Rockville, USA, May 2001.

[3] V.P. Shah, K.K. Midha, S. Dighe, I.J. McGilveray, J.P. Skelly, A. Yacobi, T. Layloff, C.T. Viswanathan, C.E. Cook, R.D. McDowall, K.A. Pittman, S. Spector, Pharm. Res. 9 (1992) 588–592.

[4] J.W.A. Findlay, W.C. Smith, J.W. Lee, G.D. Nordblom, I. Das, B.S. DeSilva, M.N. Khan, R.R. Bowsher, J. Pharm. Biomed. Anal. 21 (2000) 1249–1293.

[5] P. Kebarle, L. Tang, Anal. Chem. 65 (1993) 972A–986A.

[6] D.L. Buhrman, P.I. Price, P.J. Rudewicz, J Am. Soc. Mass Spectrom. 7 (1996) 1099–1105.

[7] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Anal. Chem. 70 (1998) 882–889.

[8] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Anal. Chem. 75 (2003) 3019–3030.

[9] P.R. Tiller, L.A. Romanyshyn, Rapid Commun. Mass Spectrom. 16 (2002) 92–98.

[10] W.Z. Shou, W. Naidong, Rapid Commun. Mass Spectorm. 17 (2003) 589–597.

[11] H. Mei, Y. Hsieh, C. Nardo, X. Xu, S. Wang, K. Ng, W.A. Korfmacher, Rapid Commun. Mass Spectrom. 17 (2003) 97–103.

[12] M.J. Avery, Rapid Commun. Mass Spectrom. 17 (2003) 197–201.

[13] J. Chamberlain, Special problems with biological fluids, The Analysis of Drugs in Biological Fluids, second ed., CRC Press Inc., USA, 1995, pp. 35–66.

[14] S. Braggio, R.J. Barnaby, P. Grossi, M. Cugola, J. Pharm. Biomed. Anal. 14 (1996) 375–388.

[15] FDA Strategic Plan, The Food and Drug Administration's Strategic Action Plan Protecting and Advancing America's Health: Responding to new challenges and opportunities, US Department of Health and Human Sciences, Food and Drug Administration, Rockville, USA, August 2003.

[16] FDC Reports, The Gold Sheet, 36 (2002) 10–32.

[17] L. Romanyshyn, P.R. Tiller, R. Alvaro, A. Pereira, E.C. Cornelis, Rapid Commun. Mass Spectrom. 15 (2001) 313–319.

[18] G. Hopfgartner, E. Bourgogne, Mass Spectrom. Rev. 22 (2003) 195–214.

[19] R. Bonfiglio, R.C. King, T.V. Olah, K. Merkle, Rapid Commun. Mass Spectrom. 13 (1999) 1175–1185.

[20] Y. Hsieh, M. Chintala, H. Mei, J. Agans, J.-M. Brisson, K. Ng, W.A. Korfmacher, Rapid Commun. Mass Spectrom. 15 (2001) 2481–2487.

[21] C. Müller, P. Schäfer, M. Störtzel, S. Vogt, W. Weinmann, J. Chromatogr. B 733 (2002) 47–52.

[22] H. Rosing, W.Y. Man, E. Doyle, A. Bult, J.H. Beijnen, J. Liq. Chromatogr. Rel. Technol. 23 (2000) 329–354.

[23] R. Kostiainen, T. Kotiaho, T. Kuuranne, S. Auriola, J. Mass Spectrom. 38 (2003) 357–372.

[24] W. Tong, S.K. Chowdhury, J-C. Chen, R. Zhong, K.B. Alton, J.E. Patrick, Rapid Commun. Mass Spectrom. 15 (2001) 2085–2090.

[25] R.O. Kringle, Pharm. Res. 11 (1994) 556–560.

[26] K.A. Selinger, J. Pharm. Biomed. Anal. 13 (1995) 1427–1436.

[27] M. Jemal, Z. Ouyang, M.L. Powell, Rapid Commun. Mass Spectrom. 16 (2002) 1538–1547.

[28] C. Hartmann, J. Smeyers-Verbeke, D.L. Massart, R.D. McDowall, J. Pharm. Biomed. Anal. 17 (1998) 193–218.

[29] Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, Anal. Chim. Acta 391 (1999) 135–148.

[30] B. Boulanger, P. Chiap, W. Dewe, J. Crommen, Ph. Hubert, J. Pharm. Biomed. Anal. 32 (2003) 753–765.