Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 854 (2007) 183–191
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Partial within-animal calibration: A new calibration approach in lead optimisation high-throughput bioanalysis

Walthère Dewé [a,*], Michaël De Smet [b], Nathalie Evrard [b], Benoit Culot [b], Marc Lastelle [b], Gaëlle Saucez [b], Dominique Ingels [b], Xavier Taillieu [b]

[a] *Lilly Development Center, Therapeutic Discovery Statistics, Eli Lilly and Company, Rue Granbonpré 11, B-1348 Mont-Saint-Guibert, Belgium*
[b] *Lilly Development Center, Drug Disposition, Eli Lilly and Company, Rue Granbonpré 11, B-1348 Mont-Saint-Guibert, Belgium*

Received 20 November 2006; accepted 15 April 2007
Available online 25 April 2007

Abstract

The lead optimization phase of drug discovery requires high-throughput analyses for quantification in biological matrices and in plasma in particular. Over the last decade, some technical innovations allowed the pharmaceutical industry to improve the quality of the results. However, there is room for improvement. In this context, a new calibration strategy is proposed in this paper. Experiments were performed on dog plasma samples and it was showed that a within-animal calibration strategy can reduce the bias up to 20% and improve the precision up to 20%. However, a partial within-animal calibration is preferred to the full approach in order to avoid to many sample preparations.

Keywords: Bioanalysis; High-throughput; Calibration; Validation; Accuracy; Bias; Precision

1. Introduction

Bioanalysis in the lead optimization (LO) phase of drug discovery is a high-throughput process for quantification of compound concentration in plasma in the context of a pharmacokinetic study. The technology that is frequently used is liquid chromatography coupled with tandem mass spectrometry (LC/MS–MS) [1–13].

Due to the compound attrition rate in the LO phase, which means that most of the compounds are analyzed only once, it is not reasonable developing a specific analytical for each compound, as done in Good Laboratory Practice (GLP) studies. As a consequence, a generic method has been developed and is routinely used for analyzing most of the compounds, no matter of the chemical diversity.

The calibration required in such analyses consists of spiking blank plasma samples with different determined amounts of analyte and then in assessing the concentration–response relationship. Usually, the blank plasma samples are taken from animals that do not enter in the pharmacokinetic study.

This calibration process could be a source of bias and variation. Indeed, calibrating with plasma that is potentially not the same as the one collected during the study may induce a matrix effect [14].

A bias or a large variability in the analytical results may induce a bias and a large variability in the derived pharmacokinetic parameter estimates, like C_{max}, half-life, bioavailability . . .. Ultimately, it may lead to take inappropriate decisions about the compound of interest.

This paper describes an experiment, performed with dog plasma, for which the results demonstrate that such matrix effects do exist, in terms of bias and precision. As a consequence, this paper also includes a proposal to eliminate this matrix effect. This proposal is not based on an optimization of the analytical method. On purpose, the analytical method stayed unchanged during the whole experiment. The proposal is actually based on a generation of little information, used for calibration, from each study animal and on an appropriate calibration strategy. The results will demonstrate an almost complete elimination of the bias as well as a significant improvement of the precision.

Three calibration strategies have been considered to generate results: the 'out-of-dog' (OOD) calibration, the 'within-dog' (WD) calibration and the 'pooled' calibration. They, respectively, consist of calibrating using spiked plasma samples

* Corresponding author.
E-mail address: walthere.dewe@scarlet.be (W. Dewé).

1570-0232/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2007.04.018

coming from a non treated dog, of calibrating in each treated dog for which samples have to be analyzed, that will be treated, and of calibrating using a pool of plasma samples. The OOD calibration is the one currently used in routine. The WD calibration would allow us to calculate the concentrations of unknown samples based on spiked samples having the same matrix. The pooled calibration is usually the one used in low-throughput bioanalysis, like in a GLP study.

Experiments were designed in order to assess the impact of these different calibration strategies on the quality of the results. The dog was the lonely species considered in our experiments.

2. Material and methods

2.1. Experimental design

The different factors that were considered in the experiments are the following:

- compound (amitryptiline, sertraline and verapamil),
- animal (#8, the 8th animal being the pool of the first 7),
- operators (#2).

Each operator had to perform the same preparations. Samples were prepared for calibration purpose only (preparation 1) while others were prepared independently for performance evaluation purpose (preparation 2). These two preparations allowed us to have a comparable number of measures to estimate the parameters of interest (bias and variance). Within each preparation, plasma of the seven dogs and the pool was spiked in duplicates at 10 different concentrations of each compound (1, 2, 5, 10, 20, 50, 100, 200, 500 and 1000 ng/ml).

In total, $4 \times (8 \times 10) \times 2 = 640$ samples were prepared and $640 \times 3 = 1920$ peak areas were obtained.

2.2. Calibration approaches

What ever the calibration strategy, either the peak area or the ratio to the selected internal standard was used as response and the same weighted quadratic regression model was performed to determine the calibration curve:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$$

where Y is the response (peak area or ratio), X is the analyte concentration and ε is the error term following a normal distribution $N(0,\sigma^2)$. The parameters β_0, β_1 and β_2 are the intercept, the slope and the quadratic coefficient, respectively. A $1/X$ weighting factor was chosen.

2.2.1. Out-of-dog calibration

The presence of seven dogs in the study design allowed us to create seven different assessments of the performance of the calibration strategy. The first assessment is the following: calibration is performed with dog 1 (preparation 1) and the concentrations of dogs 2,. . ., 7 (preparation 2) are calculated based on dog 1 calibration curves. As second assessment, dog 2 is used to calibrate and concentrations of the other dogs are calculated. And so on. This is summarized in Table 1. Because of the duplicates, the number of results available in each of the 7 cases is 24, i.e. 2 analysts × 6 animals × 2 replicates.

Table 1
Description of the different out-of-dog calibration cases

Dog used to calibrate	Analyst	Dogs to get results
1	1	2, 3, 4, 5, 6, 7
	2	2, 3, 4, 5, 6, 7
2	1	1, 3, 4, 5, 6, 7
	2	1, 3, 4, 5, 6, 7
3	1	1, 2, 4, 5, 6, 7
	2	1, 2, 4, 5, 6, 7
4	1	1, 2, 3, 5, 6, 7
	2	1, 2, 3, 5, 6, 7
5	1	1, 2, 3, 4, 6, 7
	2	1, 2, 3, 4, 6, 7
6	1	1, 2, 3, 4, 5, 7
	2	1, 2, 3, 4, 5, 7
7	1	1, 2, 3, 4, 5, 6
	2	1, 2, 3, 4, 5, 6

2.2.2. Within-dog calibration

As each analyst has performed two independent preparations in each dog, the first one was used to obtain the calibration curve necessary to calculate the concentrations of the second preparation. This was performed twice as duplicates are available in each preparation. The number of results available in this (unique) case is 28, i.e. 2 analysts × 7 animals × 2 replicates.

2.2.3. Pooled calibration

The plasma of the seven dogs were also pooled and then spiked. These samples were used for calibration and allowed us to calculate the concentrations of the seven dogs (preparation 2). The number of results available in this case is also 28, i.e. 2 analysts × 7 animals × 2 replicates.

2.3. Statistical assessment

The following assessment was performed by concentration level for each calibration strategy and for the "out-of-dog" in particular, this was applied seven times as there are seven calibration possibilities.

First of all, once the parameters β_0, β_1 and β_2 of the calibration curve are estimated by REML [15], the concentrations were calculated and the percent recoveries were derived.

Then, by level of concentration, an analysis of variance, including the dog and the analyst by dog interaction as random effects, was fitted in order to estimate the mean recovery (noted m), the dog-to-dog variance, the analyst-to-analyst variance and the residual variance. The total variance was estimated by summing the three variance components and is noted s^2 [15].

Tolerance intervals, $m \pm 2\,s$, were then computed by concentration level. These intervals are supposed to contain 95% of the

individual recoveries [16–18]. For consistency, the lower bounds and the upper bounds were smoothed over the concentration range.

The method was considered to have good performances at a concentration level if the corresponding tolerance interval is within acceptance limits, i.e. between 70% and 130%. In such case, we would have guarantees that 95% of the recoveries are between 70% and 130% [19]. Consequently, the lower (upper) limit of quantification was estimated as the smallest (largest) concentration for which both the lower and upper tolerance limits are between 70% and 130%.

2.4. Sample preparation

From each pool of concentration, 50 μl of plasma were taken and diluted in 400 μl of water. Fifty microliters of internal standard (gallopamil) were added in all samples, except blank placed after each 1000 ng/ml concentrations. These blanks were used for estimating carry-over between injections. No relative carry-over higher than 0.5% was observed (data not shown).

2.5. Chemicals and reagents

Verapamil (94837), gallopamil (M5694), amitriptyline (A8404) and sertraline (S63-19) were purchased from Sigma–Aldrich, St. Louis. Their chemical structures are shown in Fig. 1. Methanol and acetonitrile (HPLC grade) were purchased from Acros.

Ammonia (25%) was purchased from VWR Internal. Sodiumbicarbonate was purchased from Merck. Deionized water was produced in house by a Milli-Q gradient system (Millipore).

2.6. Extraction conditions with symbiosis pharma

The extraction procedure used is a "generic" method used daily in the laboratory. No adaptation was made to this method for optimizing extraction recovery. Sorbent used for extraction was Hysphere C18 high density (Spark, Holland). Extraction cartridge placed in a left clamp was activated with a 1 ml of acetonitrile/methanol (70/30) at 5 ml/min. The cartridge was then equilibrated with 1 ml of ammonia 0.1% at 5 ml/min.

The plasma sample was loaded on the cartridge with 1 ml of ammonia 0.1% via the sample loop of the autosampler at 2 ml/min.

A single backwash step was done with 1 ml of ammonia 0.1% at 5 ml/min. Cartridge was transferred automatically in the right clamp before elution takes place with the mobile phase (see 2.7 Chromatographic conditions). During the elution of the cartridge, another cartridge is prepared in the left clamp allowing parallel SPE processing.

Samples were kept at 12 °C in the autosampler during analysis time. The autosampler loop and needle were washed with two different solvents: mobile phase A (see chromatographic conditions) and methanol/water/trifluoroacetic acid (50/50/0.1) (v/v).

2.7. Chromatographic conditions

The analytical column was a XTerra C_{18} 50 mm × 4.6 mm, particle size 5 μm (Waters) with a Javelin filter (Thermo). Again, the chromatographic method was generic and no further optimization was brought to it.

The flow rate applied was 1 ml/min during gradient conditions using two mobile phases:

- Mobile phase A: methanol/isopropanol/NH_4HCO_3 50 mM (5/5/90) (v/v).
- Mobile phase B: methanol/isopropanol/NH_4HCO_3 50 mM (90/5/5) (v/v).

The proportion of A in the gradient was the following: 70% during the first 2 min, 10% during the next minute and the 70% again.

2.8. Mass spectrometry conditions

A tandem mass spectrometer API 3000 (AB-Sciex, Toronto, Canada) was used in ESI mode. Mass spectrometer parameters are shown in Table 2.

Despite the use of a basic mobile phase, the positive ionization mode was used and provided an excellent signal [20,21]. The automated tuning of Analyst 1.4 (Quantitative Optimization) software was used for creating the multiple reaction monitoring mode method. No further manual tuning was made to optimize the method. No significant cross talk was observed between analytes.

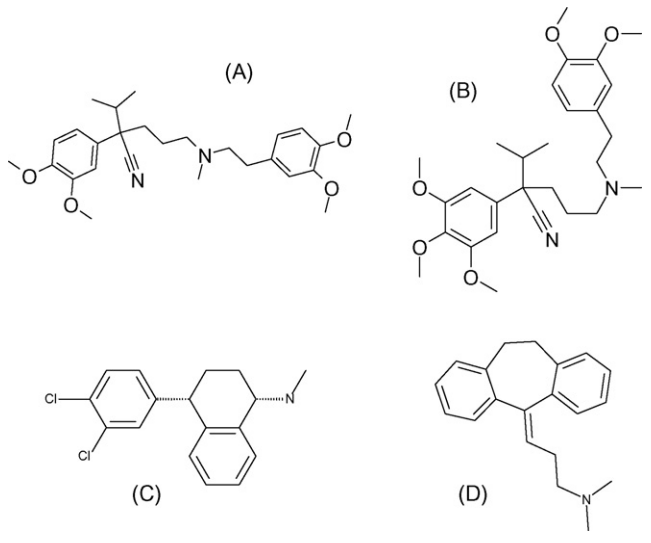
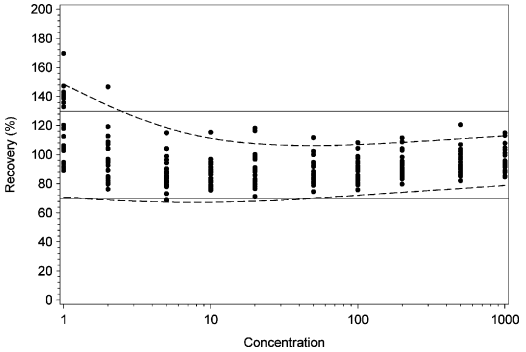
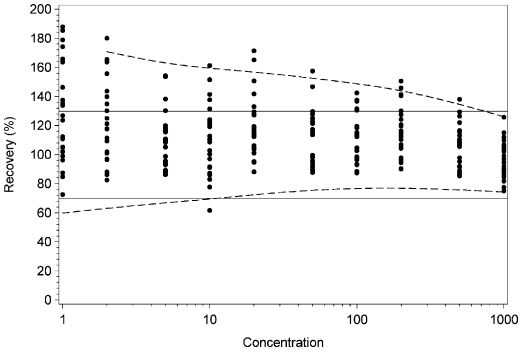
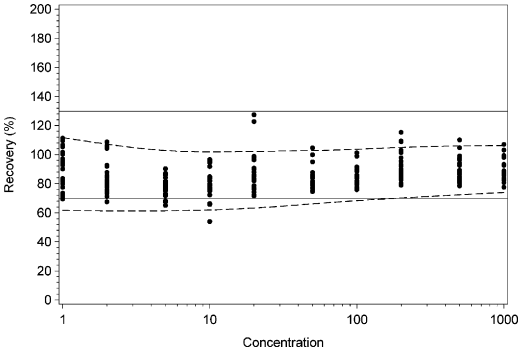
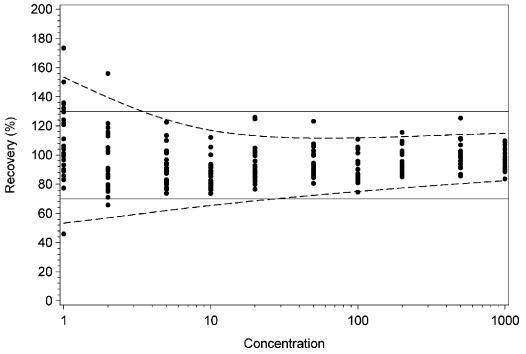
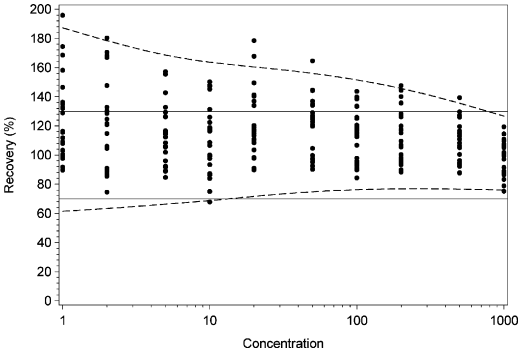
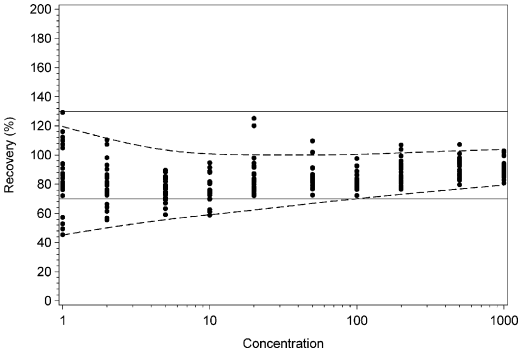
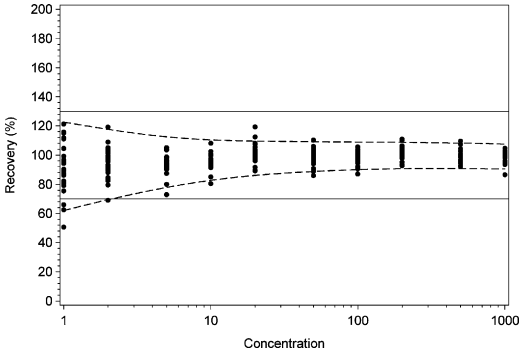
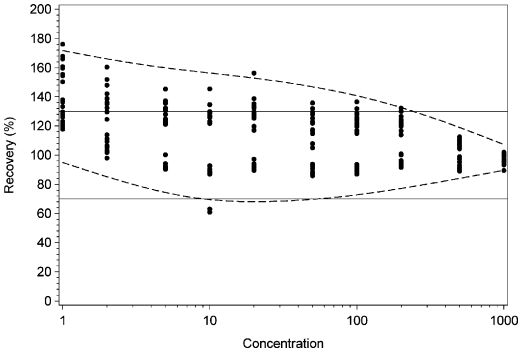
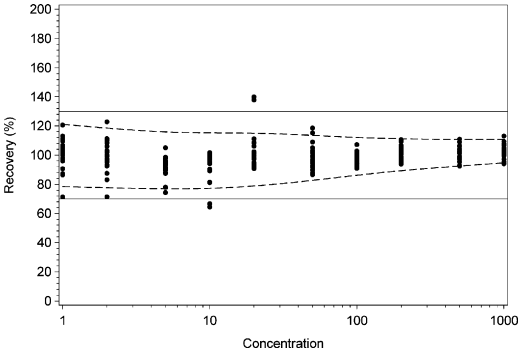


Fig. 1. Chemical structures of compounds used as analytes and internal standard.

Table 2
Details about the mass spectrometry conditions that are compound dependent

	Amitryptiline	Sertarline	Verapamil	Gallopamil
Collision Energy	23	17	39	43
Collision cell exit	14	18	12	10
MRM transition	278/233	306/275	455/165	485/165

Table 3
Distributions of percent recoveries for each analyte using the OOD calibration

Compound	Dog number used for calibration		
	1	2	6
Amitryptiline			
Sertraline			
Verapamil			

The dashed curves correspond to the tolerance interval

2.9. Software and data analysis

Analyst 1.4 Software (MDS Sciex, Toronto, Canada) was used to control the mass spectrometer, acquire and calculate data.

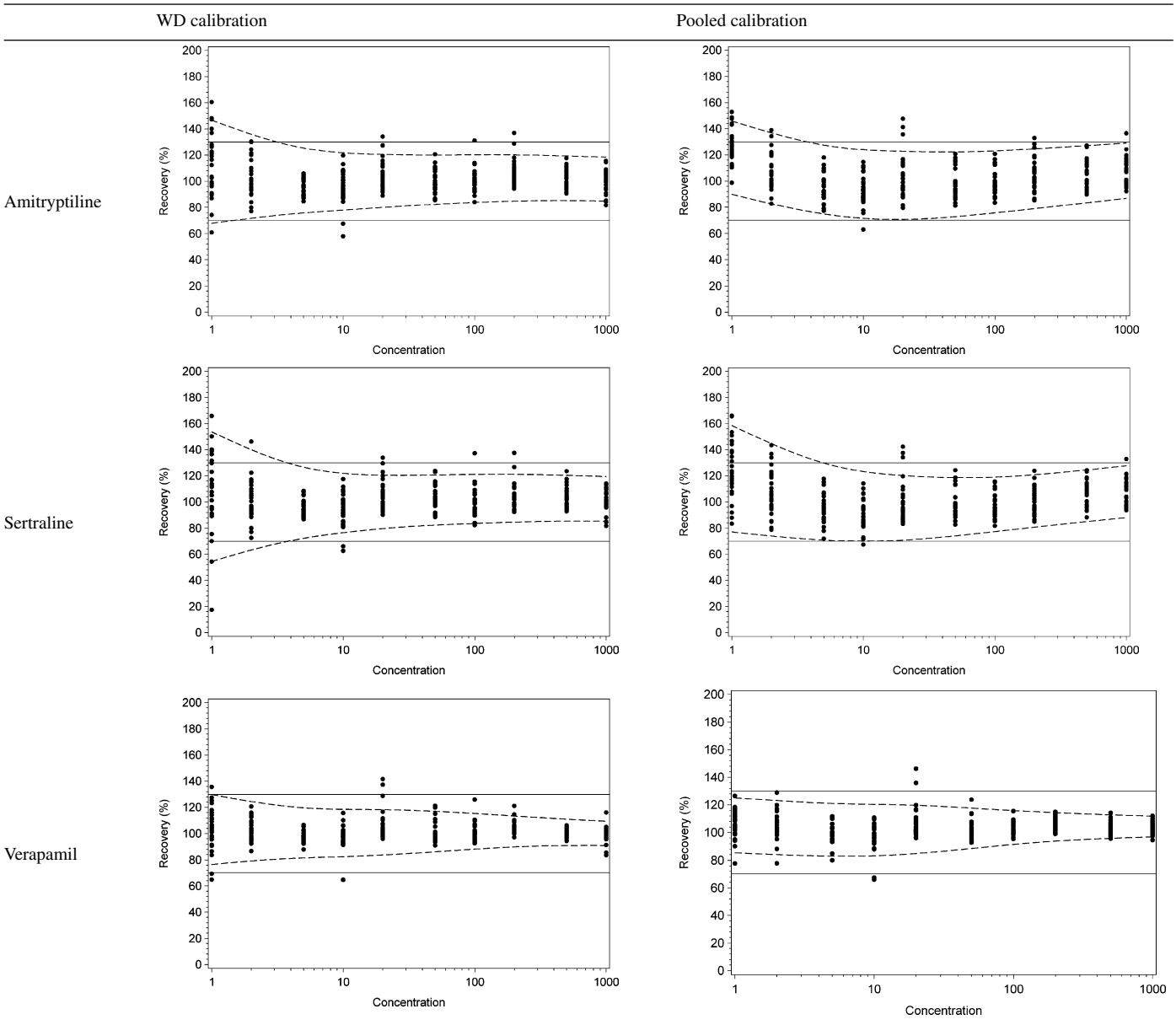
The SAS system (Statistical Analysis System, SAS Institute Inc., Cary, NC, USA), version 9.1.3, was used to perform the data management as well as the statistical computations.

3. Results

The results expressed as percent recoveries are partially (to save space) and graphically presented in Table 3 (OOD-calibration based results) and in Table 4 (WD- and pooled-calibration based results). Each graph represents a plot of the percent recoveries as a function of the introduced concentration. The solid horizontal lines represent the target, which means that we would like to have recoveries between 70% and 130%. In other words, we would the results to be within 30% of the true value. In addition, the dashed curves represent a tolerance interval in which we expect to find 95% of the future recoveries. The method is considered as valid in the concentration range for which both lower and upper curves are within the acceptance limits. The smallest concentration of such a range defines the lower limit of quantification.

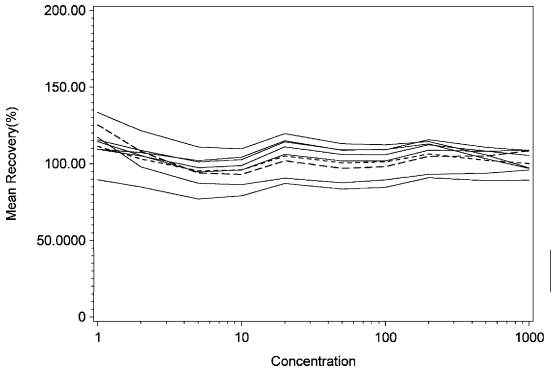
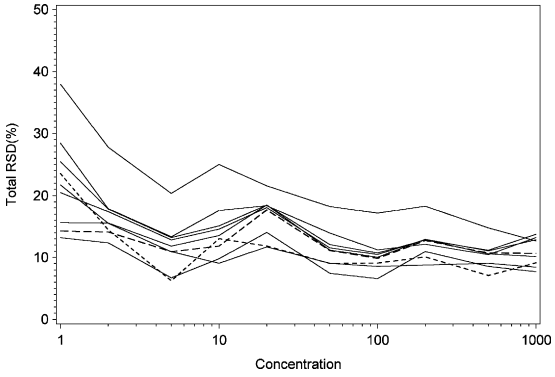
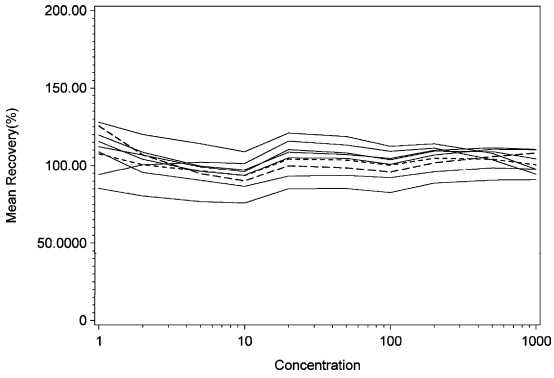
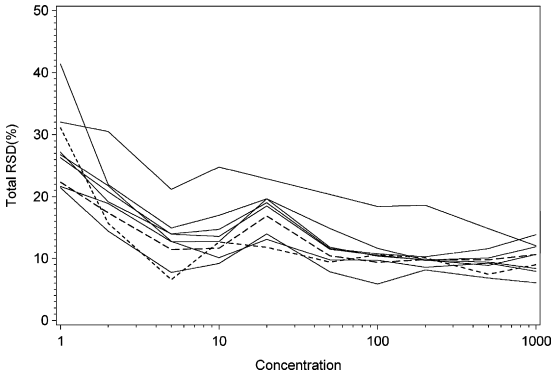
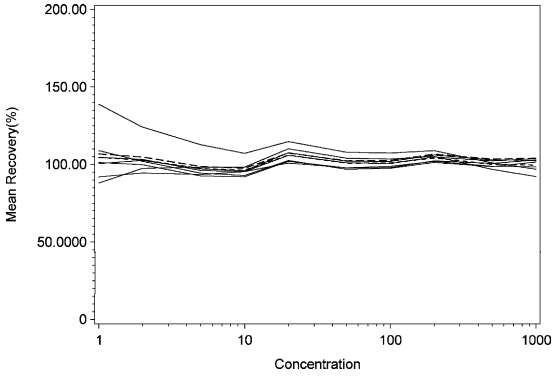
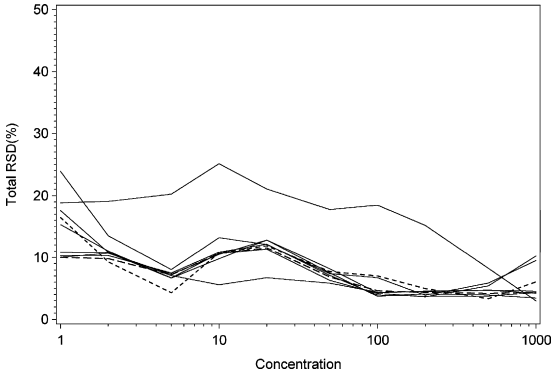
We can observe that the tolerance limits obtained with the out-of-dog calibration are rarely within the acceptance limits. In other words, the lower limit of quantification is quite high. The

Table 4
Distributions of percent recoveries for each analyte using the WD and the pooled calibration



The dashed curves correspond to the tolerance interval.

Table 5
Percent bias and percent precision estimated for each analyte

	Bias	Precision
Amitryptiline		
Sertraline		
Verapamil		

The black curves correspond to the OOD calibration, the dashed curve to the pooled calibration and the dotted curve to the WD calibration strategy.

reason is either the bias (quantifying amitryptiline or sertraline with calibration in dog 6), or the precision (quantifying amitryptiline or sertraline with calibration in dog 1), or both (quantifying with calibration in dog 2)

However, the within-dog calibration does have pretty good performances with lower limits of quantification smaller than 5 ng/ml for each compound.

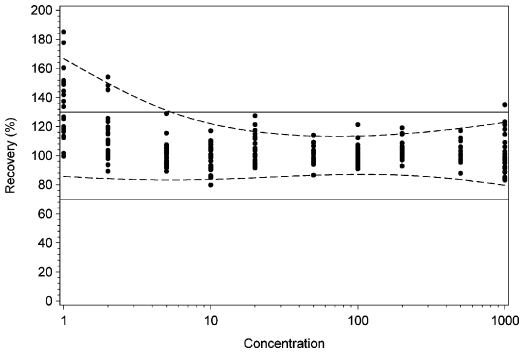
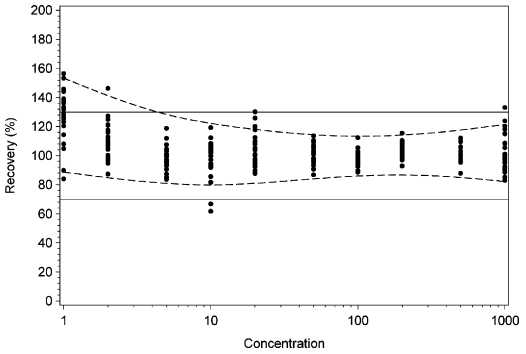
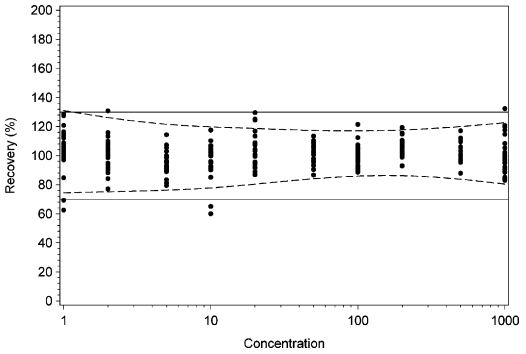
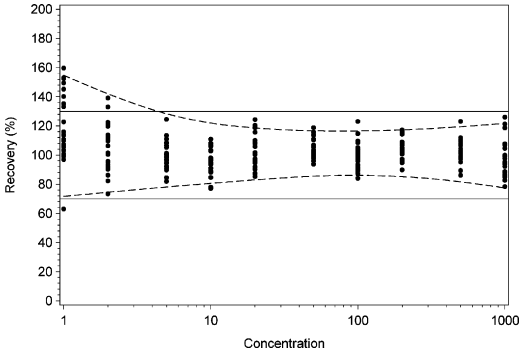
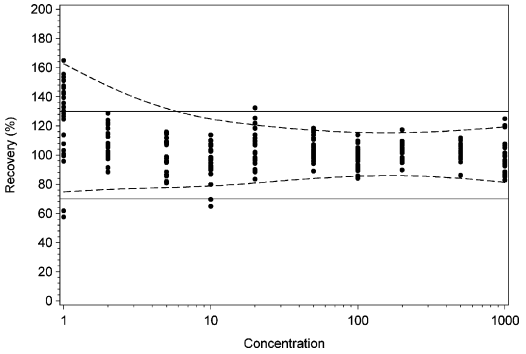
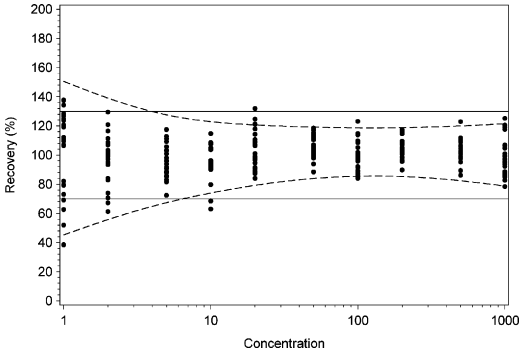
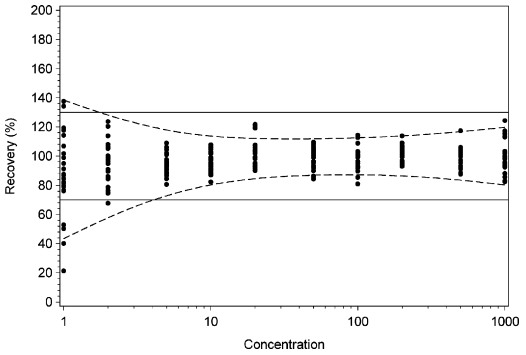
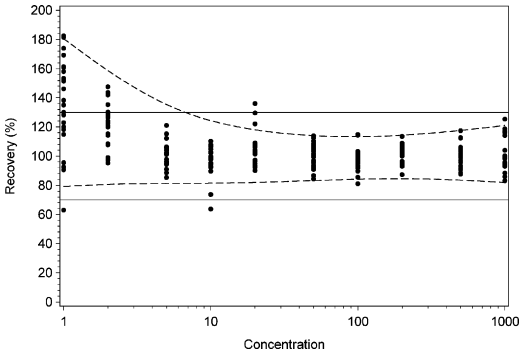
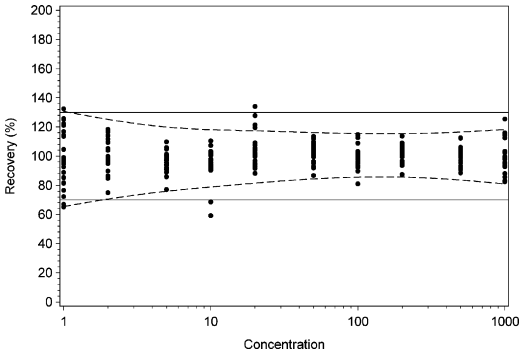
The pooled calibration has a level of performance that is between the out-of-dog calibration and the within-dog calibration. The bias is controlled but the precision is less good than the one estimated in the within-dog calibration.

All the bias and precision estimates are graphically summarized in Table 5. It can be seen that the biases of both the within-dog and the pooled calibration are close to zero as the mean recoveries are close to 100% and they are close from each other. However, depending on the dog selected to calibration, the out-of-dog calibration demonstrates significant biases, going up to 20%. The same comments can be made in terms of precision, i.e. the precision of the out-of-dog calibration can be very poor and is never better than the within-dog one.

4. Discussion

When the matrix used to prepare the calibration curve is not the same of the unknown samples, then checking the quality of a run using quality control samples, which are prepared in the same matrix as the calibration samples, does not make sense.

Table 6
Distributions of percent recoveries obtained with the WD calibration strategies

Compound	Dog number used for calibration		
	1	2	6
Amitryptiline			
Sertraline			
Verapamil			

The dashed curves correspond to the tolerance interval.

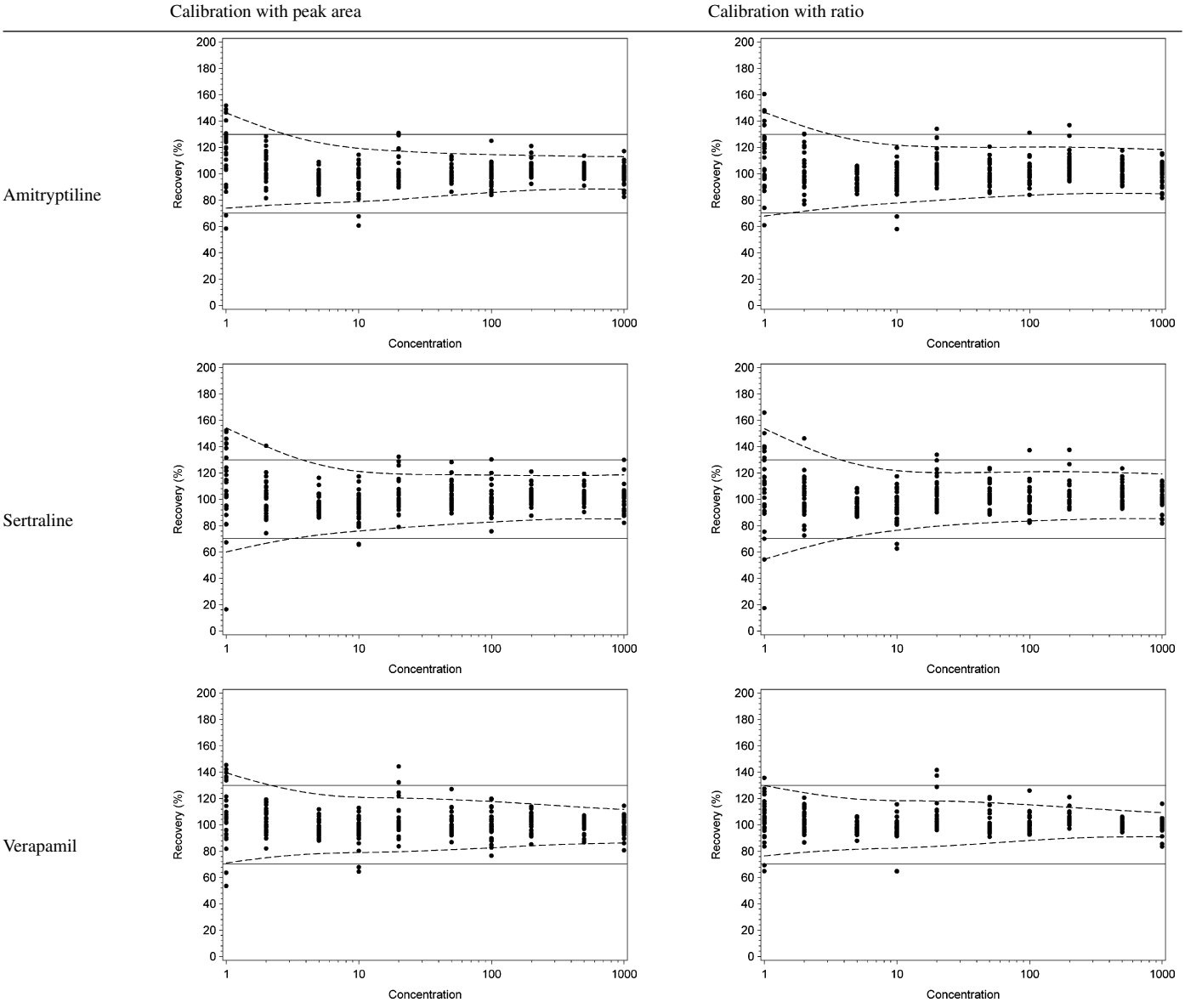
The reason is that the "within-dog" calibration does not induce any bias and its variability is smaller than the "out-of-dog" one. So despite this type of quality control, the results may be biased or the variability may be too large.

To avoid any bias and to improve the precision, the calibration should ideally be performed within each dog. However, it is not realistic to prepare a full calibration curve in each single dog that is enrolled in the study.

To solve that problem, in addition to the usual calibration curve built with dog plasma collected in animals that do not enter in the study, it could be envisaged, to analyze one or two samples from each animal, collected before the compound administration and spiked with well-chosen concentrations. By doing so, the calibration curve could be corrected for each animal and as a consequence, the bias and the precision could be improved.

For prove-of-concept, this proposal, that we could call the partial within-dog calibration, was applied using the data generated in our study. Once again, the presence of seven dogs in the study design allowed us to perform seven assessments of this proposal. The first assessment is the following: calibration is performed with dog 1 and the concentrations of dogs 2,. . ., 7 are calculated as in the first assessment of the out-of-dog calibration approach. However, instead of calibrating with dog 1 only, we added, for dogs 2,. . ., 7, in the calibration regression two samples of the first preparation, those spiked at concentrations of 100 and 500 ng/ml. Then, the second preparation samples of dogs 2,. . ., 7 were calculated based on the regression analysis. As second assessment, dog 2 is used to calibrate and concentrations of the other dogs are calculated. And so on. Two samples were necessary in this case as there were some evidence that

Table 7
Distributions of percent recoveries obtained with the WD calibration strategy after calibrating either with the peak area or with the peak ratio



The dashed curves correspond to the tolerance interval obtained at each level of concentration.

the slopes and the quadratic term coefficient were significantly different from one dog to another.

The results of these computations are summarized in Table 6. The recoveries have to be compared to those obtained with the OOD strategy (Table 3). As expected, this partial WD calibration strategy would allow us, on one hand, to significantly reduce the bias and even maybe to make it negligible in our example (dog #6), and on the other hand, to reduce significantly the variability in some cases (dog #2). If the calibration is performed with a function having a slope smaller than expected, it is going to induce a bias and an increase of the variability. If the slope is larger than it should be, then only the bias will be affected.

Using an internal standard in bioanalysis is frequent in order to improve the performances of the analytical method. An appropriate internal standard can correct a potential drift of the signal or decrease the variations. The example of the verapamil in our study illustrates the latter case (Table 7). However, if there is no drift or if the run-to-run variability is negligible or if the internal standard signals are independent of the compound signals, then the effect could be at the opposite of what expected. Indeed, taking the ratio of independent peak areas could lead to more variability than using the analyte peak area.

5. Conclusion

According to the experiments we performed in dog plasma, the bioanalytical results obtained in the lead optimization phase of drug discovery could be either biased or not precise enough if they are obtained using calibration samples prepared with plasma coming from dogs that do not enter in the study.

However, the results of the experiment described in this paper indicate that a partial within-dog calibration would allow to almost eliminate the bias (up to 15% bias reduction in our experiment), and to improve significantly the precision (up to 20% in our experiment).

As a consequence, the quality of the pharmacokinetic parameter estimates would be also improved and the risks to take a wrong decision in the pharmacokinetic study are reduced.

References

[1] M.J. Berna, B.L. Ackermann, A.T. Murphy, Anal. Chem. Acta 509 (2004) 1.
[2] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Anal. Chem. 70 (1998) 882.
[3] H. Mei, Y. Hsieh, C. Nardo, X. Xu, S. Wang, K. Ng, W.A. Korfmacher, Rapid Commun. Mass Spectrom. 17 (2003) 97.
[4] M.J. Avery, Rapid Commun. Mass Spectrom. 17 (2003) 197.
[5] M. Jemal, A. Schuster, D.B. Whigan, Rapid Commun. Mass Spectrom. 17 (2003) 1723.
[6] Y. Hsieh, M. Chintala, H. Mei, J. Agans, J.M. Brisson, K. Ng, W.A. Korfmacher, Rapid Commun. Mass Spectrom. 15 (2001) 2481.
[7] M.D. Nelson, J.W. Dolan, Ion Suppression in LC–MS-MS. LC–GC Europe, February 2002.
[8] W. Shou, W. Naidong, Rapid Commun. Mass Spectrom. 17 (2003) 589.
[9] D. Buhrman, P. Price, P. Rudewicz, J. Am. Soc. Mass Spectrom. 7 (1999) 1099.
[10] R. Bonfiglio, R.C. King, T.V. Olah, K. Merkle, Rapid Commun. Mass Spectrom. 13 (1999) 1175.
[11] L.A. Marchese, Ch. McHugh, J. Kehler, H. Bi, J. Mass Spectrom. 33 (1998) 1071.
[12] D. McLoughlin, T. Olah, J. Gilbert, J. Pharm. Biomed. Anal. 15 (1997) 1893.
[13] A. Schellen, B. Ooms, D. van de Lagemaart, B. Vreeken, W. Dongen, J. Chromatogr. B 788 (2003) 251.
[14] B.K. Matuszewski, J. Chromatogr. B 830 (2006) 293.
[15] G. Verbeke, G. Molenberghs, Linear Mixed Models in Practice: A SAS-Oriented Approach, Springer, New York, 1997.
[16] R.W. Mee, Technometrics 26 (1984) 251.
[17] R.W. Mee, Commun. Stat.—Theory Methods 17 (1988) 1465.
[18] D. Hoffman, R. Kringle, J. Biopharm. Stat. 15 (2005) 283.
[19] P. Hubert, J.J. Nguyen, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, J. Pharm. Biomed. Anal. 36 (2004) 579.
[20] B. Mansoori, D. Volmer, R. Boyd, Rapid Commun. Mass Spectrom. 11 (1997) 1120.
[21] C. Delatour, L. Leclercq, Rapid Commun. Mass Spectrom. 19 (2005) 1359.