ELSEVIER

Improving method capability of a drug substance HPLC assay

B. Dejaegher [a], M. Jimidar [b,*], M. De Smet [b], P. Cockaerts [b],
J. Smeyers-Verbeke [a], Y. Vander Heyden [a]

[a] *Department of Pharmaceutical and Biomedical Analysis, Pharmaceutical Institute, Vrije Universiteit Brussel (VUB),
Laarbeeklaan 103, 1090 Brussels, Belgium*
[b] *Global Analytical Development, Johnson & Johnson Pharmaceutical Research and Development,
A Division of Janssen Pharmaceutica N.V., Turnhoutseweg 30, 2340 Beerse, Belgium*

Abstract

The assay of a drug substance (DS) is one of the tests required to confirm the active pharmaceutical ingredient (API) quality at release. In the past, usually volumetric titration methods were performed, that were precise, but often non-specific. Nowadays specific chromatographic assay procedures are preferred. However, high performance liquid chromatographic (HPLC) methods, the way they are usually executed, tend to be less precise and have a larger total method variation compared to titration methods. The capabilities of fully validated titration and HPLC assay methods were determined and compared. It was studied which factors had the largest effects on the capability of chromatographic HPLC methods in order to improve their precision and precision-to-tolerance ratio. This was done using multiple Gage R&R (repeatability & reproducibility) studies and an experimental design approach. The investigations showed that it was feasible to define an HPLC method with a similar capability as the titration method. The most important factor determining the precision was demonstrated to be higher sample and reference material weights. When low weights are to be used, increasing the number of sample preparations and the number of reference solutions may enhance the method capability.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Method capability; Six Sigma; Gage R&R; Experimental design; *n*-Way ANOVA; HPLC assay

1. Introduction

Regulatory authorities demand release testing of drug substances (DS) and drug products (DP) in order to confirm the quality of the products at release. Guidelines and notes for guidance made by the International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), by the Food and Drug Administration (FDA) and by the Committee for Proprietary Medicinal Products (CPMP), a committee of The European Agency for the Evaluation of Medicinal Products (EMEA), can be found in the literature [1–6]. For an existing DS described in the European or in an EU-member-state Pharmacopoeia, each batch must comply with the current requirements of that Pharmacopoeia. When releasing a new DS or an existing DS not described in a Pharmacopoeia, the requirements are defined in notes for guidance [4,5].

Recommended tests and acceptance criteria are: (1) a description of the DS, (2) an identification test, (3) an assay of the DS, and (4) a determination test for impurities. Other possible tests are the description of physicochemical properties, particle size of the solid drug substance, tests for chiral substances, polymorphic forms, water content, inorganic impurities and microbial limits. The latter tests are not always required, but should be evaluated on a case-to-case assessment [6].

This study is related to DS assay methods for release testing. From the regulatory perspective, the specification limits for the assay of an active substance at release are usually set at 98.0–102.0%. The total method variance is included in these specification limits. Therefore, there will be more unnecessary and incorrect out-of-specification (OOS) cases when the applied method is less precise.

Up to now, assay methods for release testing of drug substance production batches are usually performed using volumetric titration methods. Previously performed studies (Table 1a) [7–11] show that the precision of such methods is very high (range 0.1–0.5% R.S.D.) and therefore this technique is sup-

* Corresponding author. Tel.: +32 14 60 33 87; fax: +32 14 60 55 61.
E-mail address: IJIMIDAR@PRDBE.jnj.com (M. Jimidar).

Table 1
(a) Repeatability of several titration methods, and (b) Repeatability and intermediate precision of several recently published HPLC methods, to assay a drug substance (DS) and/or drug product (DP)

DS/DP	Assay of	Repeatability (% R.S.D.)	Reference
(a) Repeatability of several titration methods			
DS	Methimazole (0.4–4.0–10.0–20.0 μmol/ml)	0.82–0.73–0.31–0.19	[7]
DS	Acetaminophen (75.60–100–500 mg)	0.13–0.65–1.01	[8]
DS	Chinine hydrochloride	0.39	[9]
DS	Chlorpromazine hydrochloride	0.37	[9]
DS	Cocaine hydrochloride	0.04	[9]
DS	Codeine hydrochloride dihydrate	0.23	[9]
DS	Ephedrine hydrochloride	0.27	[9]
DS	Homatropine hydrochloride	0.38	[9]
DS	Lidocaine hydrochloride	0.60	[9]
DS	Atropine	0.27	[10]
DS	Mefenamic acid	0.44	[11]
DS	Fenbufen	0.33	[11]
DS	Ibuprofen	0.69	[11]
DS	Diclofenac sodium	0.61	[11]
DP	Mefenamic acid tablets	0.78	[11]
DP	Fenbufen capsules	0.76	[11]
DP	Ibuprofen tablets	0.58	[11]
DP	Diclofenac sodium tablets	0.62	[11]

DS/DP	Assay of	Repeatability (% R.S.D.)	Intermediate precision (% R.S.D.)	Reference
(b) Repeatability and intermediate precision of several recently published HPLC methods				
DS	Atropine	0.54	–	[10]
DS	Budenoside	1.1	1.6	[16]
DS	Glimepiride	0.36	0.66	[17]
DS	Prazosin (50–200–500 μg/ml)	0.69–0.26–0.68	1.53–1.57–0.99	[18]
DS	Terazosin (50–200–500 μg/ml)	0.54–0.39–0.12	1.79–1.58–0.12	[18]
DS	Doxazosin (50–200–500 μg/ml)	0.42–1.32–0.25	1.71–1.91–0.53	[19]
DS	Nicotine (2–4–8–16–32–40 μg/ml)	2.00–1.74–0.95–0.69–0.64–1.24	1.99–2.21–1.63–0.51–1.21–0.48	[20]
DS	Sertraline	0.45	0.54	[20]
DP	Sertraline tablets	0.48	0.61	[20]
DP	Sertraline capsules	0.50	0.67	[20]
DP	LAS 34475[a] tablets	0.51	–	[21]
DP	Roxithromycine and	1.8	1.9	[22]
	Ambroxol.HCl tablets	1.7	1.7	

[a] A novel highly selective COX-2 inhibitor [21].

posed to have the better capability to meet specification limits. Other advantages are the simplicity and speed of analysis. Disadvantages of titration methods are the lack of specificity [12,13] and the need to perform two separate release tests (for assay and purity) [6]. Nowadays there is a general trend in industry to prefer chromatographic assay procedures for release testing of DS materials because they are more specific, and both assay and purity tests can be performed in one simultaneous procedure [6]. However, HPLC method development is difficult and very time-consuming compared to volumetric methods, since many parameters, e.g. stationary phase properties, composition and pH of the mobile phase, must be selected or optimized [14], and the HPLC system requires time to equilibrate after changing conditions, e.g. between gradient runs [15]. Moreover, it is known from previous studies (Table 1b) [10,16–22] that the precision of HPLC is generally poor (>0.5% R.S.D.) compared to that of titration methods, and therefore these methods are supposed to be less capable to meet specification limits [12,23]. For a small molecular weight drug substance the specification limits may be 98.0–102.0%, while for

drug products less tight limits, such as 95.0–105.0%, are often applied.

Important in this context are the concepts Six Sigma and Process Sigma [24–27]. Basically Six Sigma is a quality goal, where sigma is a statistical measure of the variability in a measurement system. When a method results in data of Six Sigma quality, it has a Process Sigma of at least 6, i.e. the measurement system or the total method standard deviation is at the most one-twelfth of the total allowable spread or tolerance [24]. This implies that only precise methods result in a Six Sigma quality. More information about the Six Sigma concept, the applied tools and methods can be found in Section 2.

When applying specification limits of 98.0–102.0% for the assay of an active substance, titration methods have a Process Sigma larger than 6 and apply to Six Sigma quality. However, when applying an HPLC method, usually a Process Sigma of about 3 is obtained due to the inherently lower precision, and no compliance to Six Sigma quality is seen, which will result in a higher number of OOS cases [23]. To verify and to improve the Process Sigma's, the capabilities of a fully validated HPLC and

(1) Purify DS batches ➔ $\overline{X} > 99.5\%$

$\overline{X} = 99.5\%$

s

LSL = 98.0 %   USL = 102.0 %

$\overline{X} = 98.5\%$

OOS

s

LSL = 98.0 %   USL = 102.0 %

(2) Reduce method variance

$\overline{X} = 98.5\%$

s

LSL = 98.0 %   USL = 102.0 %

(3) Change release specifications

$\overline{X} = 98.5\%$
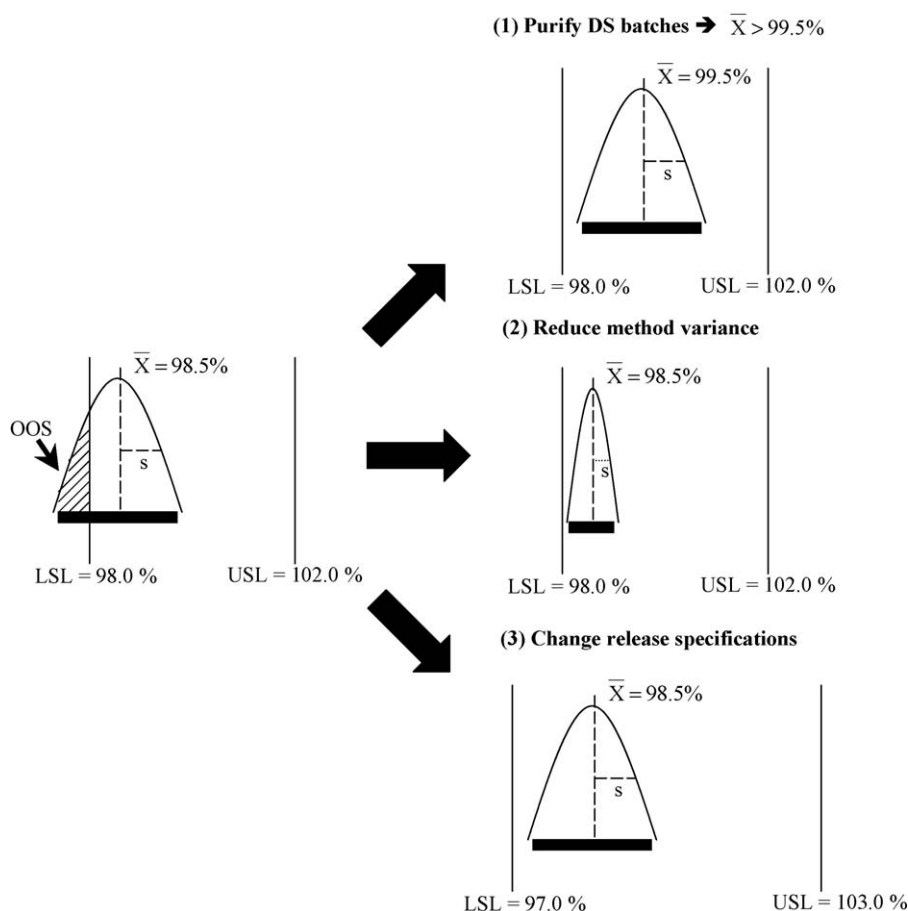
s

LSL = 97.0 %   USL = 103.0 %

Fig. 1. Example of incorrectly considered out-of-specification (OOS) cases for an HPLC method with three possible solutions to reduce these OOS cases: (1) purify DS batches, (2) reduce method variance, and (3) change release specifications. LSL = lower specification limit and USL = upper specification limit.

a titration assay method were determined and compared. Both methods concerned the assay of galantamine hydrobromide drug substance.

In order to meet the HPLC method capability towards the specification levels, theoretically, three potential approaches may be followed: (1) purify DS batches, (2) reduce method variance, and (3) change release specifications (Fig. 1). Broadening the tolerance to e.g. 97.0–103.0% would create more suitable specification limits for an HPLC method, but this is not preferred by regulatory policies and therefore not further considered. Changing the purity of the samples is an option, but may result in a significant increase of costs for the production process (costs of goods). Therefore, it was studied how the precision of the HPLC method could be improved in order to increase the Process Sigma of the assay, while maintaining the specification limits (98.0–102.0%). The purpose was to determine which factors had the major effects/impacts on the capability of chromatographic HPLC assay methods for process improvement, process control and acceptance testing of DS materials. This was done using multiple Gage R&R (repeatability & reproducibility) studies [24,26,28–30] and the results were compared to that of a titration. The critical input variables (with possible major effects) were investigated using an experimental design approach and evaluated statistically [31–33], in order to decide which have the largest influence on the method capability.

2. Theory

As stated in Section 1, important concepts related to method capability are Six Sigma and Process Sigma. The latter is an expression of process yield based on the number of defects per million opportunities (DPMO). In Ref. [25], Six Sigma is defined as "*an organized and systematic method for strategic process improvement, and new product and service development that relies on statistical methods, and the scientific method to make dramatic reductions in customer defined defect rates.*" Six Sigma, initially adopted by Motorola, is thus a way to express the quality goal of 3.4 DPMO (see below), where a defect opportunity is a process failure that is critical to the customer.

The Process Sigma of a method can be determined as follows. With the average response and the total method standard deviation of a response, a normal distribution can be defined. The part of this distribution within the specification limits (98.0–102.0%) is the yield. First the data are standardized ($z_1 = (\text{LSL} - \mu)/\sigma$ and $z_2 = (\text{USL} - \mu)/\sigma$) and the % yield is calculated with Eq. (1). The area under the curve in Eq. (1) can be found in a z-table:

$$\% \text{ yield} = 100 - (\text{area under the curve } (z < z_1))$$
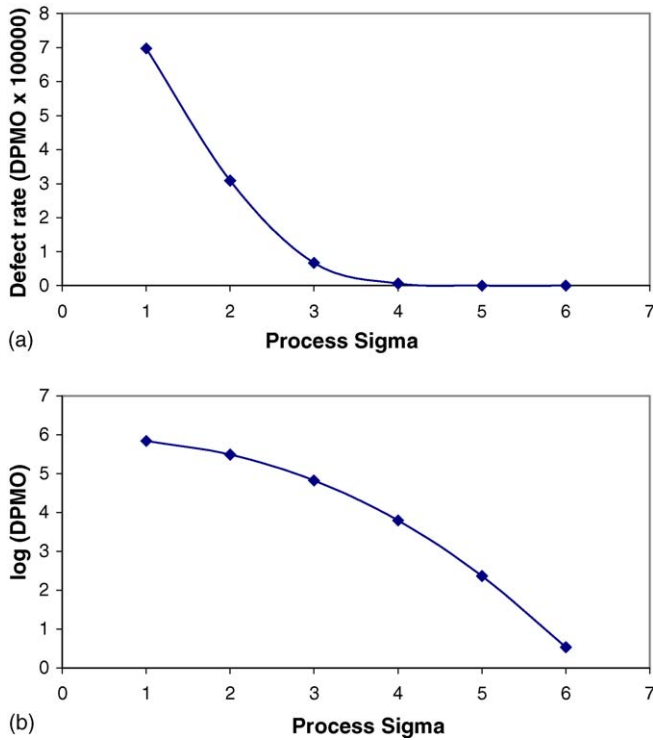$$- (\text{area under the curve } (z > z_2)) \qquad (1)$$

Fig. 2. Defect rate, expressed as (a) DPMO (defects per million opportunities), and (b) log (DPMO), as a function of Process Sigma.

Then the Process Sigma of the method can be derived from a Yield-to-Sigma conversion table (Table 2).

With a Process Sigma of 6, the nearest specification limit is at least Six Sigma, i.e. 6 standard deviations (S.D.), from the mean. It is also assumed that the process can be subject to disturbances that cause a shift of the process mean by as much as 1.5 S.D. in either direction [24–26]. A Process Sigma of 6 means a process yield of 99.9997% and 3.4 DPMO, while a Process Sigma of 3 results in 93.3200% yield and 66800 DPMO. Fig. 2 shows the relationship between the defect rate, expressed as DPMO in (a) or as log (DPMO) in (b), and Process Sigma, assuming a normal distribution of the data [25].

Six Sigma uses unique metrics, structured methods and tools [25]. A popular approach uses a five-step problem-solving process, called DMAIC. The DMAIC toolbox is an acronym for define, measure, analyze, improve and control [25–27]. Each step consists of a number of objectives accompanied by a toolbox of statistical and quality techniques, to help achieving the goal. The approach is not rigid and can be adapted to any situation, such as pharmaceutical processes [27]. The funnelling effect caused by the DMAIC approach is described in Fig. 3. Frequently applied tools are process mapping, flow charts, matrix diagrams, prioritization matrices, failure mode and effect analysis (FMEA) and Gage or Gauge R&R (repeatability & reproducibility) studies [24]. A *process map* gives a graphical representation of a process, showing the sequence of steps and alternative possibilities in order to facilitate effective planning. A process *flow chart* graphically displays the inputs, actions and outputs of a system in order to understand the process. The systematic analysis of the correlations between two

Table 2
Yield-to-Sigma conversion table (extracted from http://www.isixsigma.com)

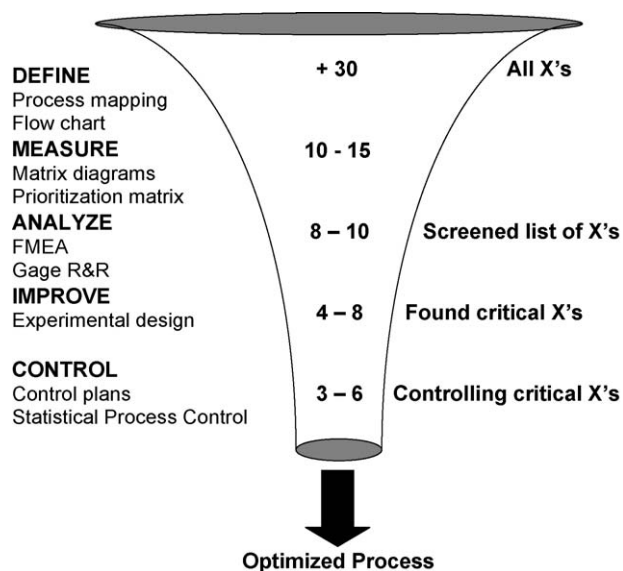| % Yield	Process Sigma
99.9997	6.00
99.9995	5.92
99.9992	5.81
99.9990	5.76
99.9980	5.61
99.9970	5.51
99.9960	5.44
99.9930	5.31
99.9900	5.22
99.9850	5.12
99.9770	5.00
99.9670	4.91
99.9520	4.80
99.9320	4.70
99.9040	4.60
99.8650	4.50
99.8140	4.40
99.7450	4.30
99.6540	4.20
99.5340	4.10
99.3790	4.00
99.1810	3.90
98.9300	3.80
98.6100	3.70
98.2200	3.60
97.7300	3.50
97.1300	3.40
96.4100	3.30
95.5400	3.20
94.5200	3.10
93.3200	3.00
91.9200	2.90
90.3200	2.80
88.5000	2.70
86.5000	2.60
84.2000	2.50
81.6000	2.40
78.8000	2.30
75.8000	2.20
72.6000	2.10
69.2000	2.00
65.6000	1.90
61.8000	1.80
58.0000	1.70
54.0000	1.60
50.0000	1.50
46.0000	1.40
43.0000	1.32
39.0000	1.22
35.0000	1.11
31.0000	1.00
28.0000	0.92
25.0000	0.83
22.0000	0.73
19.0000	0.62
16.0000	0.51
14.0000	0.42
12.0000	0.33
10.0000	0.22
8.0000	0.09

DEFINE
Process mapping
Flow chart
+ 30 | **All X's**

MEASURE
Matrix diagrams
Prioritization matrix
10 - 15

ANALYZE
FMEA
Gage R&R
8 – 10 | **Screened list of X's**

IMPROVE
Experimental design
4 – 8 | **Found critical X's**

CONTROL
Control plans
Statistical Process Control
3 – 6 | **Controlling critical X's**

Optimized Process

Fig. 3. The funnelling effect caused by the DMAIC approach, X = input variable.

groups or factors is the goal of *matrix diagrams*. A *prioritization matrix* combines a tree diagram with a matrix chart and can be developed in several ways. It is applied to determine the key factors or variables that are most important for the process. *FMEA* classifies failures according to their effect on the system (severity and probability of failure) in order to select the critical variables to concentrate capability improvement efforts on.

A Gage or Gauge R&R (repeatability & reproducibility) study is a set of trials conducted to assess the repeatability, reproducibility, stability, part-to-part variation, bias (accuracy) and linearity of the measurement system [24,28]. A Gage R&R study will indicate whether the measurement system is capable and, in our context, can be used for process improvement, control and acceptance testing. The contributions to the variation in the study results, which originate from operators of different laboratories, preparations or batches, are determined [24,28–30]. Therefore, the study is conducted by selecting batches that represent the full range of variation typically seen in the process. A statistical approach is then able to estimate the total variance and the components of the variance due to, for instance, the process, the preparation, and the operators (laboratories) [28]. If the variability of the measurement process is excessive and consequently the precision insufficient, action should be taken to improve

the measurement system in general, because repeatability and reproducibility are basic requirements for the measurement [26]. Control charts, such as R-charts, X-bar charts and s-charts, are often used to provide a graphical display of the measurement process [24].

The output of a Gage R&R is % R&R, which is defined as the fraction of the total process variation, s_{total}, due to the measurement system variation, s_M (Eq. (2)) [24]. The measurement system or total method variation, s_M, and the total process variation, s_{total}, are both expressed as standard deviations. The % R&R is an important metric when starting process improvement: it will define whether a method is capable to monitor process improvements. It is also an important metric for statistical process control (SPC):

$$\% \, R\&R = \frac{s_M}{s_{total}} \times 100 \tag{2}$$

Another output is % P/T (precision-to-tolerance ratio), which is the fraction of the specification that is due to the measurement system variation, s_M (Eq. (3)) [24,30]. The upper specification limit (USL) and the lower specification limit (LSL) determine the specification or tolerance (USL − LSL):

$$\% \, P/T = \frac{5.15 s_M}{USL - LSL} \times 100 \tag{3}$$

The variance components are obtained by means of ANOVA. They are computed using the expected mean squares (EMS) equations (Table 3) [31] and Eqs. (4)–(7):

$$\sigma^2(\text{repeatability}) = \sigma^2(\text{error}) \tag{4}$$

$$\sigma^2(\text{reproducibility}) = \sigma^2(\text{analyst}) + \sigma^2(\text{batch} \times \text{analyst}) \tag{5}$$

$$\sigma^2(M) = \sigma^2(R\&R) = \sigma^2(\text{repeatability}) + \sigma^2(\text{reproducibility}) \tag{6}$$

$$\sigma^2(\text{total}) = \sigma^2(\text{batch}) + \sigma^2(M) \tag{7}$$

The aim of improving a DS HPLC assay is to achieve a method with a % P/T smaller than 40% and a Process Sigma larger than 6. More information about Six Sigma metrics, methods and tools can be found in Refs. [24–31].

Table 3
ANOVA-table for the determination of the variance components: source of variation, number of degrees of freedom (d.f.), sum of squares (SS), mean squares (MS) and expected mean squares (EMS)

Source of variation	d.f.	SS	MS = SS/d.f.	EMS
Batch (a)	$a - 1$	SS_{batch}	MS_{batch}	$EMS_{batch} = \sigma^2_{error} + n\sigma^2_{B \times A} + nb\sigma^2_{batch}$
Analyst (laboratory) (b)	$b - 1$	$SS_{analyst}$	$MS_{analyst}$	$EMS_{analyst} = \sigma^2_{error} + n\sigma^2_{B \times A} + na\sigma^2_{analyst}$
Batch × analyst	$(a - 1)(b - 1)$	$SS_{B \times A}$	$MS_{B \times A}$	$EMS_{B \times A} = \sigma^2_{error} + n\sigma^2_{B \times A}$
Error	$d.f._{error} = d.f._{total} - d.f._{batch} - d.f._{analyst} - d.f._{B \times A}$	SS_{error}	MS_{error}	$EMS_{error} = \sigma^2_{error}$
Total	$abn - 1$			

For the HPLC method: $a = 5$, $b = 3$ and n = the number of results = 2; while for the titration method: $a = 6$, $b = 2$, and $n = 3$.

3. Experimental

3.1. HPLC method

3.1.1. Nominal conditions

The *chromatographic method* prescribes a 100 mm × 4.6 mm i.d. C18 column. The substances are eluted in a linear gradient elution mode at a flow rate of 1.5 ml/min. At the beginning of the gradient, the composition of the mobile phase is 100% disodium hydrogen phosphate dihydrate in water/methanol (95:5, v/v) (solvent A) and 0% acetonitrile/methanol (95:5, v/v) (solvent B). After 40 and 45 min the % compositions A/B are 75/25 and 60/40, respectively. Agilent and Waters Alliance (Waldbronn, Germany) ternary gradient pumps with autosamplers as injection systems were used. The injection volume was 20 µl and UV-detection (Agilent Technologies, Waters Alliance) at 230 nm is applied. The column was kept at a constant temperature in a thermostatted compartment with solvent heat exchanger (Agilent Technologies, Waters Alliance). All chemicals are from Merck (Darmstadt, Germany).

The *solutions prepared* in the operating procedure are: (1) a *sample solution*, containing 32.00 mg/50.0 ml of sample analyte, (2) a *reference solution* 1, containing 32.00 mg/50.0 ml of galantamine hydrobromide reference substance (100% of main compound), (3) a *reference solution* 2, prepared by diluting reference solution 1 2000 times in two steps (0.05% of main compound), and (4) a *selectivity solution*, containing 16.00 mg/25.0 ml of a selectivity batch with known impurity profile. The solutions are freshly prepared with water/methanol (95:5, v/v) as solvent.

3.1.2. The set-up of the Gage R&R experiments

To have a correct representation of the method performance, the samples in the Gage R&R study should cover the full, but normal variation of the process. For this purpose five batches (*a* = 5 in Table 3) were selected. Three different analysts in different laboratories (*b* = 3 in Table 3), using their proper solvents and different qualified instruments, each prepared three reference solutions of 32.00 mg (A1, A2, A3) and three of 160.00 mg (B1, B2, B3) per 50.0 ml of solution (Table 4). From each batch (*i* = 1–5), four sample solutions of 32.00 mg (sequence A) and four of 160.00 mg (sequence B) per 50.0 ml were also prepared (*j* = 1–4 in Table 4). All solutions were injected twice, except the first reference solutions (Ref A1 and B1) that were injected five times. A random order of analysis for the samples was chosen, while the standards were injected according to a fixed schedule, as is shown in Table 4. For the samples of low weight (sequence A), the precisions of injection, expressed as relative standard deviations (% R.S.D.), were 0.037%, 0.120% and 0.051% for analysts 1, 2, and 3, respectively, while for the samples of high weight (sequence B) 0.195%, 0.078% and 0.154% were found.

For each experimental condition each analyst calculated the percentage recoveries of the samples:

$$\% \text{ recovery} = \frac{A_\text{sample}}{m_\text{sample}(\text{mg})} \times \frac{m_\text{reference}(\text{mg})}{\bar{A}_\text{reference}} \times P \times 100\% \quad (8)$$

where A_sample is the peak area of the sample, m_sample and $m_\text{reference}$ the sample and reference weights, respectively,

Table 4
Work plan for each analyst, with sequence and number of injections for each reference (Ref) or sample (Samp) solution

Injected solution	No. of injections	Injected solution	No. of injections
Ref A1	5×	Ref B1	5×
Ref A2	2×	Ref B2	2×
Ref A3	2×	Ref B3	2×
Samp A43	2×	Samp B43	2×
Samp A54	2×	Samp B54	2×
Samp A14	2×	Samp B14	2×
Samp A44	2×	Samp B44	2×
Samp A13	2×	Samp B13	2×
Ref A1	2×	Ref B1	2×
Samp A22	2×	Samp B22	2×
Samp A32	2×	Samp B32	2×
Samp A53	2×	Samp B53	2×
Samp A34	2×	Samp B34	2×
Samp A24	2×	Samp B24	2×
Ref A1	2×	Ref B1	2×
Samp A33	2×	Samp B33	2×
Samp A31	2×	Samp B31	2×
Samp A41	2×	Samp B41	2×
Samp A52	2×	Samp B52	2×
Samp A11	2×	Samp B11	2×
Ref A1	2×	Ref B1	2×
Samp A42	2×	Samp B42	2×
Samp A51	2×	Samp B51	2×
Samp A21	2×	Samp B21	2×
Samp A12	2×	Samp B12	2×
Samp A23	2×	Samp B23	2×
Ref A1	2×	Ref B1	2×

Sequence of sample injections was randomly selected. Samp A*ij* represents a sample from sequence A: batch *i*, *j*th preparation; Samp B*ij* represents a sample from sequence B: batch *i*, *j*th preparation.

$\bar{A}_\text{reference}$ the (fixed or updated) average peak area of the reference (calculated depending on the calibration procedure, see further (Section 4.1)) and P is the purity of the reference substance.

3.2. Titration method

3.2.1. Description of the method

The potentiometric titration method uses a Titroprocessor 670, a Dosimat 665 burette (both from Metrohm, Herisau, Switzerland) and a glass electrode (Mettler-Toledo, Greifensee, Switzerland).

Neutralized acetic acid is prepared by dissolving 80 mg α-naphtolbenzeine indicator in 2.5 l glacial acetic acid, and then neutralizing this solution by titration with perchloric acid until the colour changes from orange to green. For the preparation of a *neutralized mixture of methyl ethyl ketone* (MEK)/*glacial acetic acid* (7:1, v/v), 80 mg α-naphtolbenzeine indicator is dissolved in a mixture of 0.3 l glacial acetic acid and 2.1 l MEK and then neutralized with perchloric acid (orange to green). A 3% *mercury acetate solution* is made by dissolving 15 g mercury acetate in 500 ml neutralized acetic acid and further neutralizing with perchloric acid (orange to green). All chemicals are from Merck (Darmstadt, Germany), except MEK which is from Baker (Deventer, The Netherlands).

The *sample solution* is prepared by accurately weighing and dissolving approximately 275 mg of drug substance in 70 ml neutralized mixture of MEK/glacial acetic acid (7:1, v/v). After adding 6 ml 3% mercury acetate, the solution is titrated potentiometrically with 0.1 M perchloric acid.

3.2.2. *The set-up of the Gage R&R experiments*

For the Gage R&R study, six different drug substance batches ($a = 6$ in Table 3) were selected. Two different analysts in different laboratories ($b = 2$ in Table 3) prepared three sample solutions from each batch. All solutions were analysed randomly.

After performing the experiments, each analyst calculated the percentage recoveries of the samples as follows

$$\% \text{ recovery} = \frac{m_{\text{titrated}}(\text{gal.HBr})}{m_{\text{weighed}}(\text{gal.HBr})} \times 100\% \qquad (9)$$

$$m_{\text{titrated}}(\text{gal.HBr}) = M_{\text{HClO}_4}(\text{mmol/ml}) \times V_{\text{HClO}_4}(\text{ml})$$
$$\times MW_{\text{gal.HBr}}(\text{mg/mmol}) \qquad (10)$$

where m_{weighed} and m_{titrated} are the weighed sample weight and that derived from the titration, respectively, M_{HClO_4} and V_{HClO_4} the molarity and the titrated volume, respectively, of perchloric acid, and $MW_{\text{gal.HBr}}$ is the molecular weight (368.27 g/mol) of galantamine HBr. All titrations were performed three times.

4. Results and discussion

4.1. *Experimental design and Gage R&R studies for the HPLC method*

In order to enhance the Process Sigma or the precision of the HPLC assay method, first the factors most worthwhile for further investigation were identified. For this purpose a sequence of Six Sigma tools, as described in the Theory, was applied. The process map and the flow chart both describe the process and all major input variables. The process map is given in Fig. 4, where supplier and customer refer to the laboratories where the method was developed and used or implemented, respectively. Table 5 shows the results of the prioritization matrix, where the score (Eq. (11)) is calculated for all process inputs:

$$\text{score} = \sum_{\text{responses}} \text{weight} \times \text{rating} \qquad (11)$$

In Eq. (11), weight is the weight that is assigned for each response (% R&R, % P/T and defect rate) to each process input and rating is the rating factor for each response based on the importance to the customer: 1 and 10 are given for the lowest and highest effect, respectively. The values for weight and rating are awarded based on expertise. The prioritization matrix identifies nine key input variables.

These nine variables were further examined using FMEA (Table 6). The risk priority number (RPN) (Eq. (12)) is calcu-
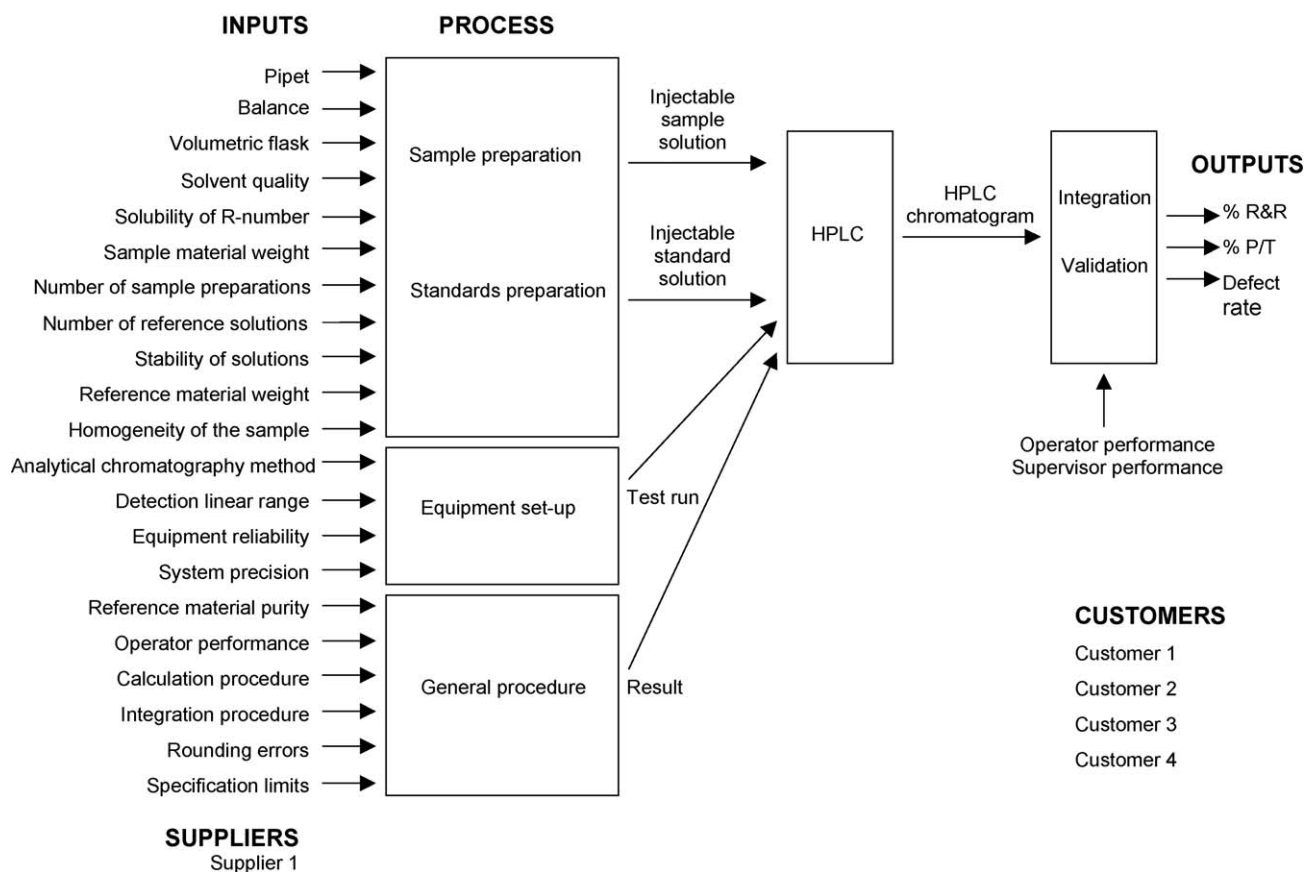

Fig. 4. Process map of the HPLC assay method.

Table 5
Prioritization matrix: score for all process inputs

No.	Input variables		Weight (rating = 10 for each response)			Score
	Process step	Process input	% R&R (response = 1)	% P/T (response = 2)	Defect rate (response = 3)	
1	Sample preparation	Balance	9	9	9	270
2	Sample preparation	Sample material weight	9	9	9	270
3	Sample preparation	Number of sample preparations	9	9	9	270
4	Sample preparation	Number of Reference solutions	9	9	9	270
5	Sample preparation	Reference material weight	9	9	9	270
6	Equipment set-up	System precision	9	9	9	270
7	General procedure	Operator performance	9	9	9	270
8	General procedure	Calculation procedure	9	9	9	270
9	General procedure	Specification limits	9	9	9	270
10	Sample preparation	Homogeneity of the sample	5	5	5	150
11	Equipment set-up	Detection linear range	5	5	5	150
12	Equipment set-up	Equipment reliability	5	5	5	150
13	General procedure	Reference material purity	5	5	5	150
14	Sample preparation	Volumetric flask	1	1	1	30
15	Sample preparation	Solubility of R-number	1	1	1	30
16	Sample preparation	Stability of solutions	1	1	1	30
17	Equipment set-up	Analytical chromatography method	1	1	1	30
18	General procedure	Integration procedure	1	1	1	30
19	General procedure	Rounding errors	1	1	1	30
20	Sample preparation	Pipet	0	0	0	0
21	Sample preparation	Solvent quality	0	0	0	0

lated for the nine key input variables by multiplying the severity, occurrence and detection ranking factors (each between 1 and 10) for each key input variable. The values for severity, occurrence and detection ranking factors are determined based on expertise. The variables with the highest RPN's are then retained to concentrate capability improvement efforts on. For more theoretical information of the above we refer to [34]:

$$RPN = severity \times occurrence \times detectability \qquad (12)$$

Six input variables were found to be potentially critical: the calculation procedure (i.e. fixed or updated average, see further), the sample and reference material weights, the number of sample preparations, the number of reference solutions, and the specification limits. Changing the release specification is not acceptable from a regulatory perspective and therefore not further considered.

The influence of the calculation procedure (A) (see further) was not fully known. Facts and figures were needed to describe its effect. It was investigated whether the updated average approach could reduce the short-term variance. Increasing the weighed amount reduces the relative standard deviation (R.S.D.) due to weighing errors and here lies the rationale to consider sample and reference material weights as critical variable (B). These two weights were seen as one factor to investigate in the experimental design. The selection of the number of sample preparations (C) and the number of reference solutions (D) could be understood from a statistical point of view (central limit theorem) [24,32] and average results are expected to have less variation than single measurement results. The four variables (A–D) identified as potentially critical input variables, were studied using Gage R&R studies and an experimental design approach. A two-level full factorial design was selected to investigate the influence of the four factors (Table 7) and to determine

Table 6
Failure mode and effect analysis (FMEA): risk priority number (RPN) for the nine key input variables

No.	Key input variables (process input)	Ranking			RPN
		Severity	Occurrence	Detectability	
1	Balance	9	7	7	441
2	Sample material weight	9	9	9	729
3	No. of sample preparations	9	9	9	729
4	Reference material weight	9	9	9	729
5	No. of Reference solutions	9	9	9	729
6	Operator performance	9	4	7	252
7	System precision	9	5	8	360
8	Calculation procedure	9	9	9	729
9	Specification limits	9	8	9	648

Table 7
Two-level full factorial design for four factors with the results of the responses examined: total method variance, total method standard deviation, % R&R and % P/T

Standard order	Run order	Design conditions				Responses			
		(A) Calculation procedure	(B) Sample/Ref material weight (mg)	(C) No. of sample preparations	(D) Number of Reference solutions	Total method variance	Total method standard deviation	% R&R	% P/T
1	11	1	32	1	1	0.871	0.933	94.7	120.2
2	15	1	32	1	3	0.846	0.920	94.2	118.4
3	8	1	32	2	1	0.658	0.811	87.4	104.4
4	9	1	32	2	3	0.349	0.591	79.7	76.0
5	12	1	160	1	1	0.046	0.215	59.1	27.7
6	7	1	160	1	3	0.039	0.198	56.1	25.5
7	13	1	160	2	1	0.035	0.188	57.0	24.2
8	2	1	160	2	3	0.055	0.235	66.0	30.2
9	6	2	32	1	1	0.831	0.912	93.3	117.4
10	4	2	32	1	3	0.829	0.910	93.9	117.2
11	16	2	32	2	1	0.598	0.773	86.5	99.5
12	14	2	32	2	3	0.378	0.615	81.9	79.2
13	10	2	160	1	1	0.033	0.183	52.4	23.6
14	1	2	160	1	3	0.041	0.201	55.0	25.9
15	3	2	160	2	1	0.029	0.169	53.6	21.7
16	5	2	160	2	3	0.053	0.231	66.9	29.7

which had the largest influence on the method capability of the HPLC method.

The minimal level of the factors in the design was the nominal level described in the assay procedure, i.e. in the currently applied HPLC procedure. The maximal levels for each factor were chosen in such a way that an improvement of the method capability could be expected (Table 7). For the sample and reference material weights (B), 32.00 and 160.00 mg were chosen as minimal and maximal levels, respectively. Concerning the number of sample preparations (C), one or two preparations were considered. For one sample preparation, the two results taken from the Gage R&R studies are the first two reported percentages, i.e. the averages of injections of the first two sample solutions (Table 8). On the other hand, for two sample preparations, the average values of first/second and third/fourth reported percentages (Table 8) were used ($n = 2$ in Table 3). Four sample solutions were therefore prepared. The practical implications of this factor for future experiments are to decide on the preparation of one or two sample solutions. When one sample preparation will be considered best to improve the method precision, in the future only one sample solution will be prepared and injected twice. The average of these two injections will then be used for further calculations. On the other hand, if two sample preparations will improve the method precision most, two sample solutions will be prepared and each injected twice. The mean of the two average values of the duplicated injections will then be used.

The possible calculation procedures (A) are summarized in Fig. 5. Two options were possible for the number of reference solutions (D): (1) single-level (one reference solution) and (2) multi-level (three reference solutions) calibration. When looking at the single-level procedure, again two possibilities were considered. The assay procedure described a calibration using a *single point* (*fixed average*) *approach* (calculation procedure 1 in Table 7). The fixed average from the first injections of the

reference solution (five times Ref A1 or B1 in Table 4) was used to determine the percentage recovery in all sample solutions. The reference solution was re-injected after 10 consecutive sample injections solely to ensure the absence of system drift (Evaluation 1 and 2 in Fig. 5). In this study the *updated average approach* was also investigated (calculation procedure 2 in Table 7). The percentage recovery was then calculated using the updated averages, which were determined by each previously injected reference solution, i.e. taking into account both the initially and the later injected reference solutions (see Fig. 5). The latter approach can circumvent small, but significant drifts (1–2%) of the system, while the first considers these drifts to be within the system suitability. For example, when considering one reference solution (Ref A1), the updated average used to compute the % recovery in samples A22, A32, A53, A34, and A24 (Table 4), is based on the first five injections of Ref A1 and the two injections of Ref A1 between samples A13 and A22.

When a multi-level calculation procedure was applied, the peak areas of the references, $A_{\text{reference}}$, were normalized, $A_{\text{normalized}}$ (Eq. (13)), in order to calculate the (fixed or updated) average normalized peak area of the references, $\bar{A}_{\text{normalized}}$, expressed in mg^{-1}. The reference material weight, expressed in mg, is represented by $m_{\text{reference}}$:

$$A_{\text{normalized}} = \frac{A_{\text{reference}}}{m_{\text{reference}}(\text{mg})} \tag{13}$$

To estimate the % recovery in the samples, the ratio $m_{\text{reference}}(\text{mg})/\bar{A}_{\text{reference}}$ in (Eq. (8)) is then replaced by $1/\bar{A}_{\text{normalized}}$, i.e. the reciprocal of the (fixed or updated) average normalized peak area of the references.

The percent recoveries of the samples from the HPLC method at the different conditions are shown in Table 8. More information on the reported values is already given above. Each column of Table 8 will be used to calculate the responses for two runs of

Table 8
Gage R&R for the HPLC method at the different conditions: percent recoveries of the analyses of five batches by three analysts in different laboratories, for each run of the experimental design from Table 7 (see numbers between parentheses)

| Analyst	Batch	Sample solution	Percentage recoveries (%) (standard order)							
			(1) (3)	(9) (11)	(5) (7)	(13) (15)	(2) (4)	(10) (12)	(6) (8)	(14) (16)
1	1	1	99.3	99.4	99.6	99.6	99.5	99.5	100.0	99.8
1	1	2	99.6	99.7	99.7	99.8	99.8	99.7	100.1	99.9
1	1	3	99.6	99.6	100.0	100.0	99.8	99.8	100.4	100.4
1	1	4	99.5	99.5	99.7	99.7	99.7	99.7	100.1	100.1
1	2	1	98.7	98.7	99.6	99.7	98.9	98.8	100.0	99.9
1	2	2	100.1	100.1	99.6	99.7	100.3	100.2	100.0	99.9
1	2	3	99.6	99.6	99.7	99.8	99.8	99.7	100.1	100.0
1	2	4	99.0	99.0	99.7	99.7	99.2	99.2	100.1	100.0
1	3	1	98.8	98.9	99.7	99.7	99.0	99.0	100.0	99.9
1	3	2	99.9	99.9	99.7	99.8	100.1	100.1	100.1	100.1
1	3	3	99.3	99.3	99.6	99.6	99.5	99.4	99.9	99.8
1	3	4	98.9	98.9	99.6	99.7	99.1	99.0	100.0	99.9
1	4	1	98.6	98.6	98.9	99.0	98.8	98.7	99.3	99.2
1	4	2	98.9	98.9	98.5	98.5	99.1	99.0	98.9	98.7
1	4	3	98.6	98.6	99.1	99.1	98.8	98.8	99.5	99.5
1	4	4	98.2	98.2	99.1	99.1	98.4	98.4	99.4	99.4
1	5	1	99.1	99.1	99.1	99.2	99.3	99.2	99.5	99.4
1	5	2	99.0	99.0	99.2	99.3	99.2	99.1	99.6	99.5
1	5	3	98.1	98.1	99.2	99.2	98.2	98.2	99.6	99.5
1	5	4	98.3	98.3	99.7	99.7	98.5	98.5	100.1	100.1
2	1	1	99.7	99.6	99.9	99.8	99.7	99.6	99.9	99.8
2	1	2	100.4	100.3	99.7	99.6	100.4	100.3	99.7	99.6
2	1	3	99.8	99.8	99.9	99.9	99.8	99.8	100.0	100.0
2	1	4	99.7	99.7	99.8	99.8	99.7	99.7	99.9	99.9
2	2	1	100.1	100.0	99.8	99.7	100.1	100.0	99.9	99.8
2	2	2	100.2	100.1	99.8	99.7	100.1	100.1	99.8	99.8
2	2	3	100.2	100.1	99.7	99.6	100.2	100.1	99.8	99.7
2	2	4	100.2	100.1	99.8	99.7	100.1	100.1	99.8	99.7
2	3	1	100.1	100.0	99.8	99.7	100.1	100.0	99.8	99.7
2	3	2	99.0	99.0	99.8	99.7	99.0	99.0	99.8	99.8
2	3	3	98.9	98.8	99.7	99.6	98.8	98.8	99.7	99.6
2	3	4	100.4	100.4	99.7	99.6	100.4	100.4	99.7	99.6
2	4	1	99.2	99.1	99.2	99.1	99.2	99.1	99.2	99.1
2	4	2	99.7	99.6	99.2	99.1	99.7	99.6	99.2	99.1
2	4	3	99.7	99.7	99.3	99.3	99.7	99.7	99.3	99.3
2	4	4	99.3	99.3	99.2	99.2	99.3	99.3	99.2	99.2
2	5	1	95.5	95.4	99.4	99.3	95.5	95.4	99.4	99.3
2	5	2	99.8	99.7	99.4	99.3	99.8	99.7	99.4	99.3
2	5	3	99.5	99.5	99.4	99.3	99.5	99.5	99.4	99.3
2	5	4	99.6	99.6	99.3	99.3	99.6	99.6	99.3	99.3
3	1	1	99.5	99.2	99.6	99.6	98.9	98.9	99.7	99.6
3	1	2	100.3	100.0	99.8	99.7	99.8	99.7	99.9	99.8
3	1	3	100.7	100.7	99.5	99.5	100.2	100.2	99.6	99.6
3	1	4	101.7	101.7	99.0	99.0	101.2	101.2	99.1	99.1
3	2	1	99.8	99.4	99.1	99.1	99.2	99.2	99.2	99.1
3	2	2	99.8	99.7	99.8	99.7	99.3	99.3	99.9	99.8
3	2	3	100.8	100.5	99.6	99.6	100.3	100.2	99.7	99.6
3	2	4	101.9	101.7	100.1	100.0	101.4	101.3	100.2	100.1
3	3	1	99.8	99.5	99.7	99.7	99.3	99.2	99.8	99.8
3	3	2	100.0	99.9	99.7	99.7	99.5	99.5	99.8	99.8
3	3	3	101.1	100.9	99.1	99.1	100.6	100.6	99.2	99.2
3	3	4	100.1	99.9	99.7	99.7	99.5	99.5	99.8	99.8
3	4	1	99.3	99.1	99.0	99.0	98.8	98.8	99.1	99.0
3	4	2	100.1	99.8	99.3	99.2	99.6	99.5	99.4	99.3
3	4	3	100.2	100.2	98.9	98.9	99.6	99.6	99.0	99.0
3	4	4	100.4	100.4	99.2	99.2	99.8	99.8	99.3	99.3
3	5	1	99.4	99.0	99.7	99.6	98.8	98.8	99.8	99.7
3	5	2	99.7	99.4	99.3	99.3	99.1	99.1	99.4	99.4
3	5	3	99.4	99.2	99.0	99.0	98.9	98.8	99.1	99.1
3	5	4	100.1	100.1	99.6	99.6	99.5	99.5	99.7	99.7
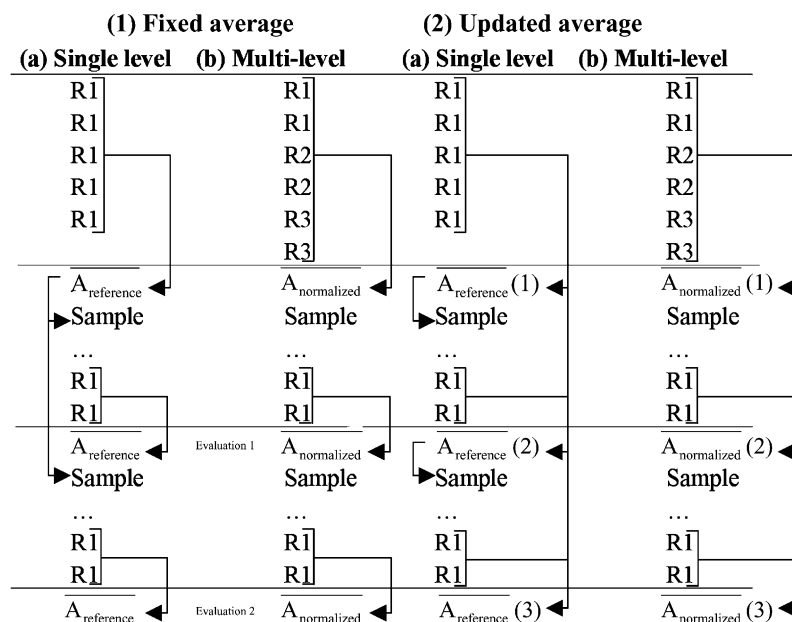
Fig. 5. Calculation procedures: (1a) single-level and (1b) multi-level fixed average approach, (2a) single-level and (2b) multi-level updated average approach. R1, R2, R3 = reference solutions 1, 2 and 3 (Table 4), $\bar{A}_{\text{reference}}$ = mean peak area, $\bar{A}_{\text{normalized}}$ = normalized mean peak area, and . . . = 10 sample injections.

the experimental design (Table 7). The studied design responses were the total method variance (and standard deviation), % R&R and % P/T. All responses were calculated for each run of the experimental design (Table 7). The total process variance was calculated and divided into parts using the EMS [31] (Table 3) and Eqs. (4)–(7).

Both graphical and statistical methods are used to analyze the results of the Gage R&R study shown in Table 8. In order to visually observe which factor (A–D) effects are critical for the method precision, main effect plots [33] and two-factor interaction plots are made for the design response total method variance. The main effect plot is shown in Fig. 6(1). The *X*-axis on this plot represents the minimal (1) and maximal (2) level of the investigated factors and the *Y*-axis the response. The two-factor interaction effect plots are given in Fig. 6(2). These plots allow deciding which main and interaction effects are important and which levels are to be preferred in order to achieve the smallest total method variance, i.e. the best method precision. Interaction between two factors exists if the effect of one factor depends on the level of the second, i.e. graphically the two lines in the interaction effect plot will not be parallel. If the effect of one factor is the same at both levels of the other, no interaction exists and the two lines are parallel.

Analysis of variance (ANOVA) was applied to examine statistically the significance of each factor. Since the two-factor interactions with factor (A) are less important (see Fig. 6(2)), only the two-factor interactions between the three remaining factors (B–D) were considered. Three- and four-factor interactions also were neglected and not included in the ANOVA model. The main effects of the four factors (A–D) and three two-factor interaction (BC, BD, CD) effects were estimated from the experimental design and investigated using multi-way ANOVA

[31–33]. The model employed was of the form:

$$
\begin{aligned}
Y = {} & b_0 + b_A X_A + b_B X_B + b_C X_C + b_D X_D + b_{BC} X_B X_C \\
& + b_{BD} X_B X_D + b_{CD} X_C X_D + \varepsilon
\end{aligned}
\tag{14}
$$

where Y is the response value for a specific combination of factor settings, b_0 the overall average or mean value of the design results, b_i the main coefficient of a factor i, b_{ij} the two-factor interaction coefficient between factors i and j, X_i the design level of factor i, and ε is the error term or random residual variation. The results of the ANOVA are summarized in Table 9.

From observing Fig. 6(1) and Table 9, it can be concluded that the influence of the calculation procedure (A) is the smallest and the updated average calculation procedure is not statistically different ($p = 0.617$) from the currently applied approach, i.e. fixed average. The other three main effects (B–D) are all statistically significant ($p < 0.0005$ for (B and C), and $p = 0.038$ for (D)). Indeed, applying higher sample and reference weights (B), more sample preparations (C) or a higher number of reference solutions (D), leads to a smaller total method variance, which is also observed in Fig. 6(1). This figure and Table 9 also showed that the precision of the method is by far most improved using higher weights (≥ 160 mg).

When evaluating Fig. 6(2), the three two-factor interactions with factor (A) are practically not existing since the two lines are (almost) parallel. From the remaining two-factor interactions, the lowest interaction was observed between the number of sample preparations and the number of reference solutions (CD). Indeed, the interaction (CD) is not statistically significant ($p = 0.056$) (Table 9). On the other hand, the other two-factor interactions considered (BC and BD) are statistically significant ($p < 0.0005$ for (BC) and $p = 0.019$ for (BD)), which can also be
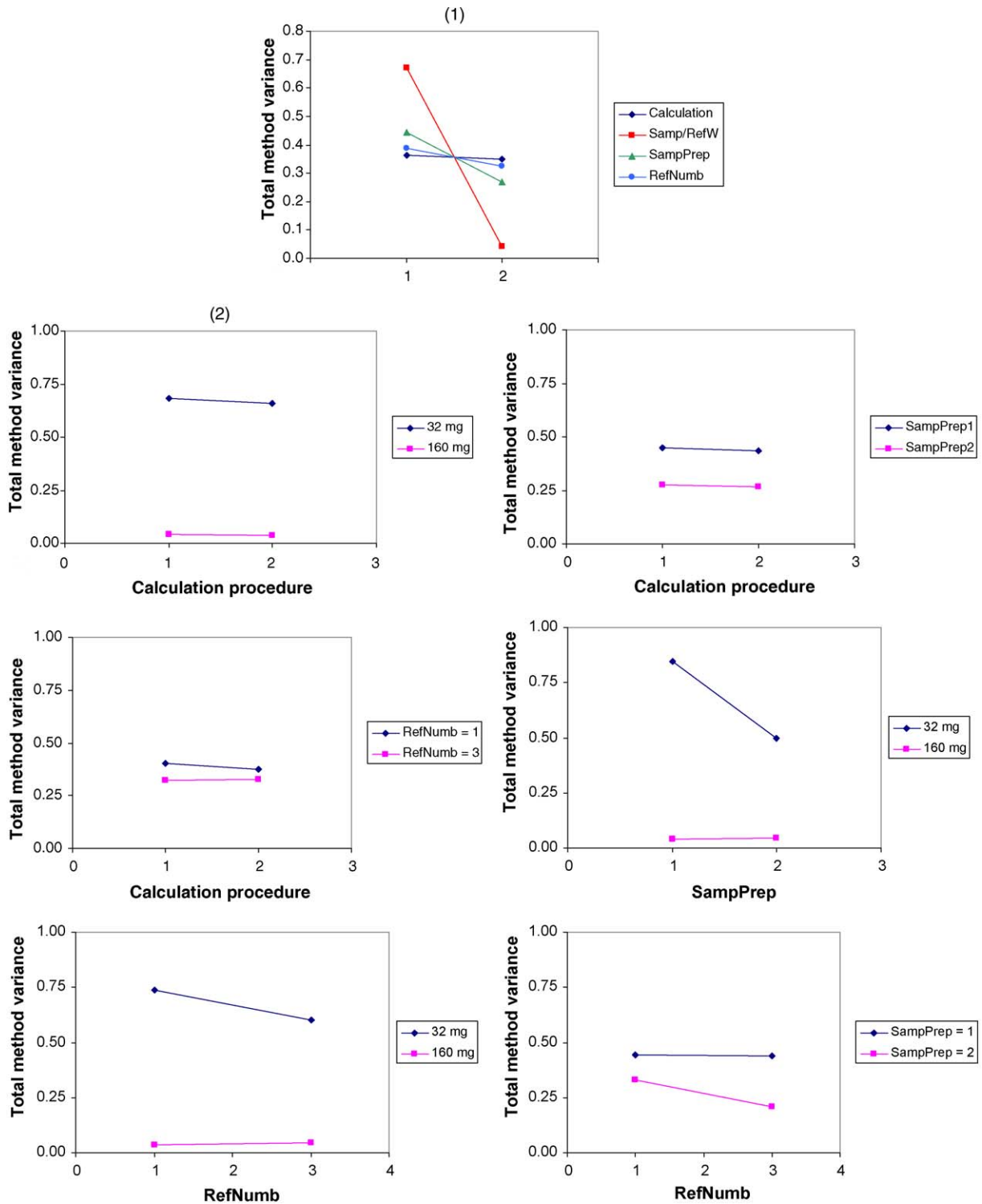
Fig. 6. (1) Main effect plot and (2) two-factor interaction effect plots on the response total method variance. Calculation procedure (Calculation), sample and reference material weights (Samp/RefW), number of sample preparations (SampPrep) and number of reference solutions (RefNumb), 1 on *X*-axis = low factor level and 2 on *X*-axis = high factor level.

seen in Fig. 6(2). The lowest total method variance is obtained when high sample and reference material weight is used, and then the number of sample preparations or the number of reference solutions does not have much influence on the variance. On the other hand, when applying a low weight, the latter two factors become important. More sample preparations or a higher number of reference solutions will then lead to a smaller total method variance, i.e. a better precision of the method. However, in both cases it is still worse than that obtained at high sample and reference weights.

Table 9
Analysis of variance (ANOVA) for the statistical evaluation of the significance of each factor on the response total method variance: investigation of four main effects and six two-factor interaction effects from a two-level four-factor full factorial design

Source of variation	d.f.	SS	MS = SS/d.f.	$F = MS_X/MS_{error}$	P
(A) Calculation ($a = 2$)	$(a-1) = 1$	$bcd\sum_a(\bar{X}_A - \bar{X})^2 = 0.0007$	0.0007	0.27	0.617
(B) Samp/RefW ($b = 2$)	$(b-1) = 1$	$acd\sum_b(\bar{X}_B - \bar{X})^2 = 1.5807$	1.5807	598.74	<0.0005
(C) SampPrep ($c = 2$)	$(c-1) = 1$	$abd\sum_c(\bar{X}_C - \bar{X})^2 = 0.1192$	0.1192	45.15	<0.0005
(D) RefNumb ($d = 2$)	$(d-1) = 1$	$abc\sum_d(\bar{X}_D - \bar{X})^2 = 0.0163$	0.0163	6.18	0.038
(BC) Samp/RefW × SampPrep	$(b-1)(c-1) = 1$	$ad\sum_b\sum_c(\bar{X}_{BC} - \bar{X}_B - \bar{X}_C + \bar{X})^2 = 0.1237$	0.1237	46.87	<0.0005
(BD) Samp/RefW × RefNumb	$(b-1)(d-1) = 1$	$ac\sum_b\sum_d(\bar{X}_{BD} - \bar{X}_B - \bar{X}_D + \bar{X})^2 = 0.0226$	0.0226	8.55	0.019
(CD) SampPrep × RefNumb	$(c-1)(d-1) = 1$	$ab\sum_c\sum_d(\bar{X}_{CD} - \bar{X}_C - \bar{X}_D + \bar{X})^2 = 0.0132$	0.0132	4.99	0.056
Error	d.f.$_{total}$ − d.f.$_{main+interaction}$ = 8	SS$_{total}$ − $\sum$ SS$_{main+interaction}$ = 0.0211	0.0026		
Total	$(abcd - 1) = 15$	$\sum_a\sum_b\sum_c\sum_d(X_{ABCD} - \bar{X})^2 = 1.8975$			

Calculation procedure (Calculation), sample and reference material weight (Samp/RefW), number of sample preparations (SampPrep) and number of reference solutions (RefNumb), each with two levels (respectively $a = b = c = d = 2$).

4.2. Gage R&R study for the titration method

The set-up and results of the Gage R&R study for the titration method are summarized in Table 10. All titrations were

Table 10
Gage R&R for the titration method: results (percentage recoveries) of the analyses of six batches by two analysts from different laboratories

Analyst (laboratory)	Batch	Sample solution	Percentage recoveries (%)
1	1	1	99.5
1	1	2	99.6
1	1	3	99.6
1	2	1	99.6
1	2	2	99.5
1	2	3	99.6
1	3	1	99.6
1	3	2	99.7
1	3	3	99.6
1	4	1	99.6
1	4	2	99.5
1	4	3	99.6
1	5	1	99.6
1	5	2	99.5
1	5	3	99.5
1	6	1	99.8
1	6	2	99.6
1	6	3	99.5
2	1	1	100.0
2	1	2	100.1
2	1	3	100.0
2	2	1	100.0
2	2	2	99.8
2	2	3	100.0
2	3	1	100.1
2	3	2	100.1
2	3	3	99.9
2	4	1	100.0
2	4	2	100.1
2	4	3	100.0
2	5	1	99.7
2	5	2	99.8
2	5	3	99.9
2	6	1	100.0
2	6	2	100.0
2	6	3	100.0

performed three times ($n = 3$ in Table 3) and averages were reported in Table 10. The extent of the study for the titration method was limited, because of the known good precision of volumetric methods. The percentage recoveries of galantamine HBr are reported. The main effect plots are shown in Fig. 7. The X-axis represents the levels of batch, analyst (laboratory) and sample solutions, while the Y-axis represents the average percent recovery, at a given level. No major differences in the response between batches are noticed. The reason for the small variation between batches might be found in the lack of specificity of the titration method [12,13]. Certain degradation products or impurities will also be measured/quantified. The percent degradation products and impurities in a batch usually increases when the percentage of drug substance decreases, and vice versa. Titration methods often tend to determine this total sum, which explains the small difference in response, while HPLC methods, on the other hand, separate DS from degradation products and impurities and determine only the active DS, leading to higher differences in response for different batches. The different sample solutions did not affect the percent recovery much either. The analyst (laboratory) was found to have the largest influence on the response, although its influence is less compared to the variation that is observed when examining the main effect plots for the currently applied HPLC method, which are given in Fig. 8 for comparison.

The total method variance, the total method standard deviation, % R&R and % P/T were calculated for the titration method and were 0.082, 0.287, 98.5% and 37.0%, respectively. The Process Sigma was determined and was found to be larger than 6.0 (% yield > 99.9997%).

4.3. Summary of the Gage R&R studies

Table 11 summarizes the results of the Gage R&R studies obtained for both methods. For each method, the Process Sigma and the % P/T were calculated as described in the Theory section. When comparing the titration (% yield > 99.9997%, Process Sigma > 6.0, % P/T = 37.0%) and the nominal HPLC method, i.e. all factors of the experimental design at minimal level (% yield = 87.7096, Process Sigma = 2.7 and % P/T = 120.2%)
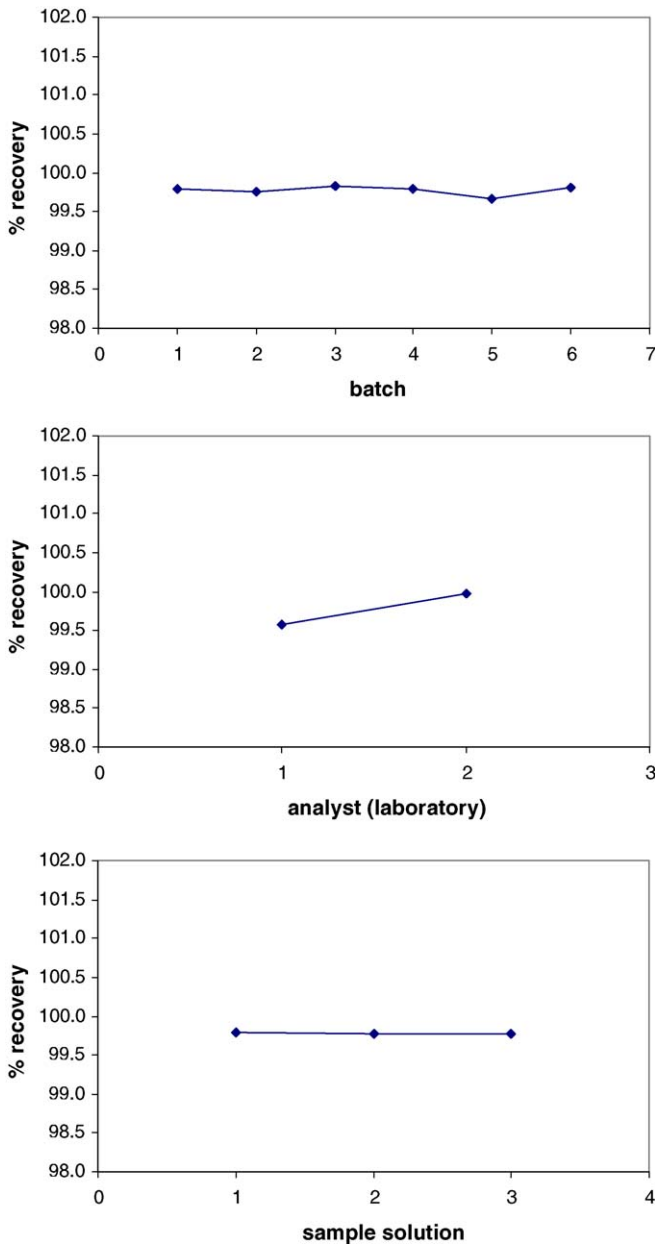
Fig. 7. Main effect plots for the titration method: influences of batch, analyst (laboratory) and sample solutions on the percent recovery.
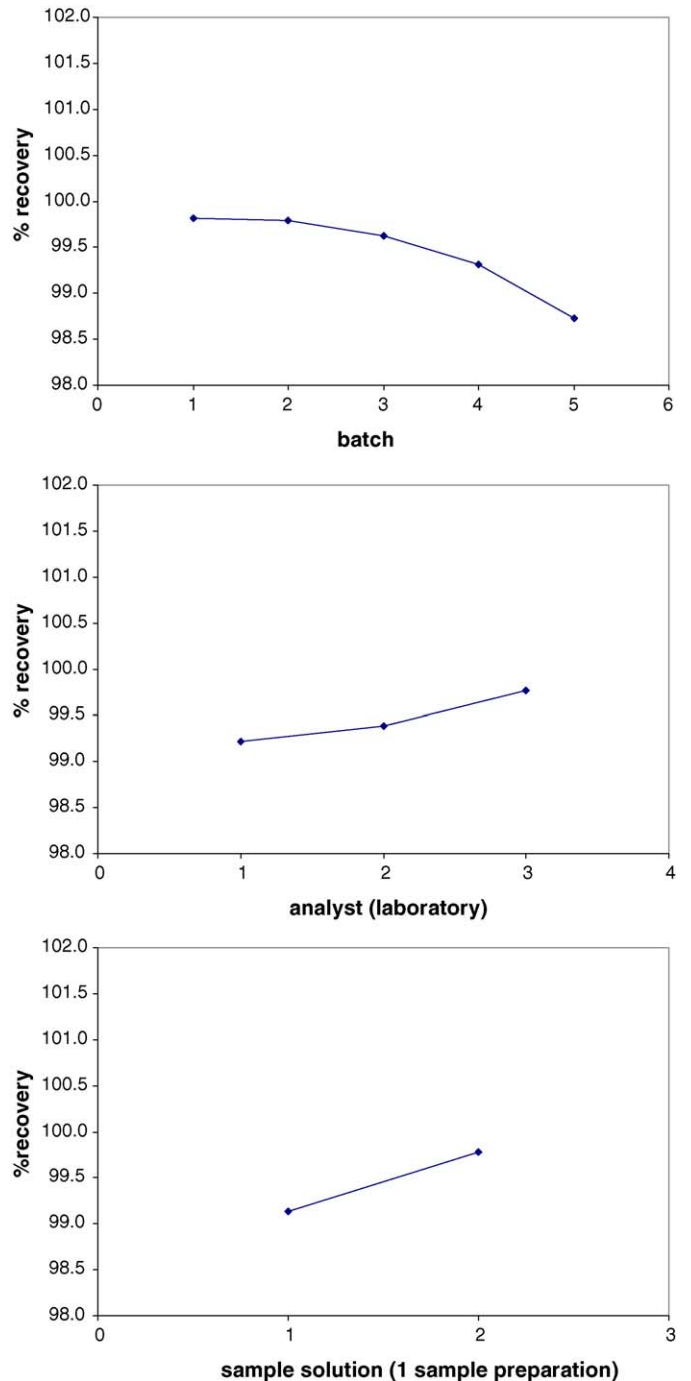

Fig. 8. Main effect plots for the currently applied HPLC method (standard order 1 in Table 7): influences of batch, analyst (laboratory) and sample solutions on the percent recovery.

(standard order (1) in Tables 7 and 11), one can conclude that the HPLC method, when specified as thus, is less precise and in practice this will result in more incorrect OOS cases.

However, the summary of the Gage R&R results (Table 11) also shows that it is possible for a HPLC method to result in a similar method capability (Process Sigma > 6.0 and % P/T < 40%) as the titration method when adjusting certain factors in the current nominal HPLC procedure, such as the sample and reference material weights, the number of sample preparations and the number of reference solutions. Since the influence of sample and reference material weights was found the largest, the most efficient way to increase the precision and consequently to improve the method capability is by using higher sample weights (≥160 mg) (standard orders (5)–(8) and (13)–(16) in

Tables 7 and 11). When using higher weights, the number of sample preparations and the number of reference solutions are not very critical since higher numbers do not improve the precision considerably anymore. If somehow high weights are impossible to achieve, for example due to the higher costs, increasing the latter two factors will also enhance the method capability (standard orders (4) and (12) in Tables 7 and 11), but less than using higher sample and reference weights. When increasing both the number of sample preparations and the number of reference solu-

Table 11
Gage R&R results for the titration method and for all HPLC interpretations, i.e. at all conditions of the experimental design: Process Sigma for (a) number of sample preparations, $n = 1$ and (b) number of sample preparations, $n = 2$

Method—parameters	Process Sigma	% P/T
(a) Number of sample preparations = 1		
Titration method	>6.0	37.0
HPLC method (standard order)		
Single level—fixed average 32 mg (1)	2.7	120.2
Single level—updated average 32 mg (9)	2.6	117.4
Multi level—fixed average 32 mg (2)	2.5	118.4
Multi level—updated average 32 mg (10)	2.5	117.2
Single level—fixed average 160 mg (5)	>6.0	27.7
Single level—updated average 160 mg (13)	>6.0	23.6
Multi level—fixed average 160 mg (6)	>6.0	25.5
Multi level—updated average 160 mg (14)	>6.0	25.9
(b) Number of sample preparations = 2		
HPLC method (standard order)		
Single level—fixed average 32 mg (3)	3.1	104.4
Single level—updated average 32 mg (11)	3.1	99.5
Multi level—fixed average 32 mg (4)	3.7	76.0
Multi level—updated average 32 mg (12)	3.6	79.2
Single level—fixed average 160 mg (7)	>6.0	24.2
Single level—updated average 160 mg (15)	>6.0	21.7
Multi level—fixed average 160 mg (8)	>6.0	30.2
Multi level—updated average 160 mg (16)	>6.0	29.7

Between parentheses: the standard order of the experiment in Table 7.

tions is not feasible, e.g. due to the higher costs, precision is best improved by using two sample preparations and the reporting of averages (standards orders (3) and (11) in Tables 7 and 11). Using three reference solutions (standards orders (2) and (10) in Tables 7 and 11) leads to a smaller improvement of precision.

5. Conclusions

The method capability of an HPLC method to assay a drug substance can be improved, in such a way that a similar capability is obtained as with a titration method. Hence, it would no longer be needed to apply a more precise volumetric method that shows lack of specificity to assay a drug substance for release testing, since HPLC methods with good specificity and precision can be defined.

The factor with the largest impact on the capability of the chromatographic HPLC method for process improvement, control and acceptance testing of drug substance (DS) materials is the sample and reference material weight. When applying low weights (here ≤ 32 mg), factors as the number of sample preparations and the number of reference solutions become important in controlling the capability since increasing either the number of sample preparations or both factors reduces the total method variance. On the other hand, when using high weights (here ≥ 160 mg) the two latter factors are less important since they do not decrease the total method variance considerably anymore. Therefore, for this assay, it is recommended to use sample and reference material weights of at least 160 mg. Although this requires a higher reference standard material consumption and thus an increase in costs, less incorrect out-of-specification cases will occur.

The main recommendation from this study is to apply large enough weights. This recommendation will be taken into account for future method description of DS HPLC assay methods.

References

[1] International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Guidelines. http://www.ich.org/.
[2] Food and Drug Administration (FDA), Department of Health and Human Services. http://www.fda.gov/.
[3] The European Agency for the Evaluation of Medicinal Products, Evaluation of Medicines for Human Use. http://www.emea.eu.int/.
[4] Guideline Requirements in Relation to Active Substances, Notice To Applicants, 3AQ6A, vol. IIIA. http://pharmacos.eudra.org/.
[5] Note for Guidance on Summary of Requirements for Active Substances in the Quality part of the dossier, CPMP/QWP/297/97 Rev 1, The European Agency for the Evaluation of Medicinal Products, Evaluation of Medicines for Human Use (EMEA), London, 2002.
[6] Note for Guidance Specifications: Test procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances, CPMP/ICH/367/96, The European Agency for the Evaluation of Medicinal Products, Evaluation of Medicines for Human Use (EMEA), London, 1999.
[7] M. Aslanoglu, N. Peker, J. Pharm. Biomed. Anal. 33 (2003) 1143–1147.
[8] G. Burgot, F. Auffret, J.L. Burgot, Anal. Chim. Acta 343 (1997) 125–128.
[9] K. Takács-Novák, G. Völgyi, Anal. Chim. Acta 507 (2004) 275–280.
[10] C. Kirchhoff, Y. Bitar, S. Ebel, U. Holzgrabe, J. Chromatogr. A 1046 (2004) 115–120.
[11] O. Cakirer, E. Kiliç, O. Atakol, A. Kenar, J. Pharm. Biomed. Anal. 20 (1999) 19–26.
[12] J. Tsau, J.W. Poole, Int. J. Pharm. 12 (1982) 185–197.
[13] S. Görög, J. Pharm. Biomed. Anal. 36 (2005) 931–937.
[14] Y. Vander Heyden, C. Perrin, D.L. Massart, in: K. Valkó (Ed.), Handbook of Analytical Separations, Separation Methods in Drug Synthesis and Purification, vol. 1, Elsevier, Amsterdam, 2000, pp. 163–212.
[15] E. Van gyseghem, M. Jimidar, R. Sneyers, D. Redlich, E. Verhoeven, D.L. Massart, Y. Vander Heyden, J. Chromatogr. A 1074 (2005) 117–131.
[16] S. Hou, M. Hindle, P.R. Byron, J. Pharm. Biomed. Anal. 24 (2001) 371–380.
[17] P. Kovariková, J. Klimes, J. Dohnal, L. Tisovská, J. Pharm. Biomed. Anal. 36 (2004) 205–209.
[18] M. Bakshi, T. Ojha, S. Singh, J. Pharm. Biomed. Anal. 34 (2004) 19–26.
[19] K.R. Tambwekar, R.B. Kakariya, S. Garg, J. Pharm. Biomed. Anal. 32 (2003) 441–450.
[20] D. Chen, S. Jiang, Y. Chen, Y. Hu, J. Pharm. Biomed. Anal. 34 (2004) 239–245.
[21] I. Toro, J.F. Dulsat, J.L. Fábregas, J. Claramunt, J. Pharm. Biomed. Anal. 36 (2004) 57–63.
[22] M. Qi, P. Wang, R. Cong, J. Yang, J. Pharm. Biomed. Anal. 35 (2004) 1287–1291.
[23] B. Renger, J. Chromatogr. B 745 (2000) 167–176.
[24] T. Pyzdek, The Six Sigma Handbook: A Complete Guide for Greenbelts, Blackbelts, and Managers at All Levels, McGraw-Hill, Quality Publishing, Tuscon, NY, 1999, pp. 140–142, 252–273, 295–300, 377–385, 452–464, 595–599.
[25] K. Linderman, R.G. Schroeder, S. Zaheer, A.S. Choo, J. Oper. Manag. 21 (2003) 193–203.
[26] H. Chonghun, L. Young-Hak, Ann. Rev. Contr. 26 (2002) 27–43.
[27] D. Johnston, Pharm. Tech. Eur. 15 (2003) 57–61.
[28] M. Ciopec, Chemometrics Intell. Lab. Syst. 21 (1993) 21–34.
[29] S.B. Vardeman, E.S. Van Valkenburg, Technometrics 41 (1999) 202–211.

[30] W.H. Johnson, W.A. Keenan, T. Wetteroth, Nucl. Instrum. Meth. Phys. Res. B55 (1991) 148–153.

[31] R.R. Sokal, F.J. Rohlf, Biometry: The Principles and Practice of Statistics in Biological Research, 2nd ed., W.H. Freeman, New York, 1981, pp. 372–393.

[32] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, 1997, pp. 56–58, 121–150, 177.

[33] A.J. Duncan, Quality Control and Industrial Statistics, 4th ed., Irwin, Illinois, 1986, pp. 702–747.

[34] Z. Bluvband, P. Grabov, O. Nakar, Expanded FMEA (EFMEA). http://www.ald.co.il/articles/Expanded_FMEA%20_EFMEA.pdf.