ELSEVIER

Journal of Chromatography A, 845 (1999) 145–154

JOURNAL OF CHROMATOGRAPHY A

Determination of system suitability limits with a robustness test

Y. Vander Heyden[a], M. Jimidar[b,*], E. Hund[a], N. Niemeijer[b], R. Peeters[b], J. Smeyers-Verbeke[a], D.L. Massart[a], J. Hoogmartens[b]

[a]*Vrije Universiteit Brussel, Pharmaceutical Institute, Pharmaceutical and Biomedical Analysis, Laarbeeklaan 103, 1090 Brussels, Belgium*

[b]*Janssen Research Foundation, Analytical Development, Pharmaceutical Quality Control, Turnhoutseweg 30, 2340 Beerse, Belgium*

Abstract

A robustness test was performed on a chromatographic method to identify and assay an active substance and two related compounds in film-coated tablet. For a number of responses the originally applied system suitability criteria were evaluated based on the results of the robustness test. Ambiguous situations can occur in situations where a method is found to be robust to assay the substances, as was the case here, but when system suitability criteria for some responses are violated. To avoid this, a proposal is made to define or re-define system suitability limits based on the results of the robustness test. From the effects found in the robustness test, the experimental conditions giving the worst result that still is acceptable and probable to occur are predicted and the system suitability limits are defined from replicated experiments in these conditions.

Keywords: Robustness test; System suitability test; Validation; Ruggedness; Pharmaceutical analysis; Ridogrel

1. Introduction

Robustness is defined by the International Conference on Harmonisation (ICH) [1] as the capacity of an analytical procedure to remain unaffected by small, but deliberately introduced variations in the method variables. The term ruggedness is frequently used as a synonym [2–5]. Actually, only in Ref. [6] a distinction is made and ruggedness is defined there as the degree of reproducibility of the test results obtained under a variety of normal test conditions.

Robustness is evaluated by a robustness test. In a robustness test a statistical experimental design is applied to examine simultaneously the effect of the variation in different method variables (factors), e.g., the flow of the mobile phase, the temperature, type of column, slope of the gradient, the buffer pH, concentration of salts in the mobile phase, detector wavelength, concentration of additives, etc., on the outcome (response) of a method [2,3,7]. The factors are investigated at different levels, usually two, namely a low (−1) and a high (+1) level which are situated around the nominal (0) one [4]. The nominal levels are the conditions stated in the assay procedure. According to the ICH guidelines [1] the evaluation of robustness should be considered during the development phase of a method (or at the beginning of the validation), and not at the end of method validation as was originally the case [5].

Different types of designs can be used in robustness testing e.g., fractional factorial designs [7–9] and Plackett–Burman designs [7,10,11]. The choice of a design depends on the purpose of the test and on the number of factors to be examined. In general, the purpose of a robustness test is to indicate the factors

*Corresponding author. Fax: +32-14-605-838.
E-mail address: ijimidar@janbe.jnj.com (M. Jimidar)

0021-9673/99/$ – see front matter
PII: S0021-9673(99)00328-3

that can significantly influence the outcome of the studied responses. This gives an idea of the potential problems that might occur when the method is repeated at different conditions or transferred to, for instance, another laboratory. These problems then can be anticipated by controlling the significant factors, for example, by including a ''precautionary statement'' [1] in the method description.

It is possible to evaluate the effects of the examined factors on different responses. The parameters, evaluated in a system suitability test (SST) such as resolution, peak tailing, column efficiency, capacity factor, etc., can also be considered as responses in a robustness test. A system suitability test is an integral part of many analytical methods [1] and it ascertains the suitability and effectiveness of the operating system [6]. The SST limits for the different parameters usually are established based on the experimental results obtained during the optimisation of a method and on the experience of the analyst. However the ICH guidelines recommend that ''one consequence of the evaluation of robustness should be that a series of system suitability parameters (e.g., resolution tests) is established to ensure that the validity of the analytical procedure is maintained whenever used''. In this work, the system suitability limits for different parameters, established with formerly applied procedures, were evaluated by means of the results from a robustness test and also a proposal is made to define system suitability limits based on the evaluation of the robustness.

The case study described concerns the robustness testing of the high-performance liquid chromatographic method for identification and assay of ridogrel and for the detection of the related compounds in ridogrel oral film-coated tablets.

2. Experimental

2.1. Chemicals

The main compound: ridogrel (MC), related compound 1 (RC1), related compound 2 (RC2) and placebo formulation tablets were obtained from Janssen Research Foundation (Beerse, Belgium). Methanol was purchased from J.T. Baker (Deventer, Netherlands), ammonium acetate from Fluka (Buchs, Switzerland), acetonitrile LC-grade from Acros (Geel, Belgium). The water used was from Milli-Q equipment (Millipore, Milford, MA, USA). Chemically resistant Acrodisc filters were obtained from Schleicher and Schuell (Dassel, Germany).

2.2. Solutions

All solutions were prepared in dark amber glassware. The method of analysis uses an external standard without placebo.

2.2.1. Standard solution related compounds

Approximately 12.5 mg of related compound 1 (RC1) and of related compound 2 (RC2) are accurately weighed into a 100-ml volumetric flask, dissolved in and diluted to volume with methanol.

2.2.2. Reference solution

Approximately 25 mg of ridogrel reference material are accurately weighed into a 100-ml volumetric flask, 1.0 ml of the ''standard solution related compounds'' and 50 ml of methanol–0.25% (m/v) solution of ammonium acetate in water (9:1, v/v) are added. The mixture is mechanically shaken for 30 min and diluted to volume with the same mixture.

2.2.3. Sample solution

Approximately 25 mg of ridogrel reference material are accurately weighed into a 100-ml volumetric flask, 1.0 ml of the ''standard solution related compounds'', 10 placebo tablets and 50 ml of methanol–0.25% (m/v) solution of ammonium acetate in water (9:1, v/v) are added. The mixture is mechanically shaken for 30 min, diluted to volume with the same mixture and filtered through a 0.45-μm chemically resistant Acrodisc-filter.

2.2.4. Blank solution

The blank solution was methanol–0.25% (m/v) solution of ammonium acetate in water (9:1, v/v).

2.3. Chromatographic conditions

The method prescribes a 10 cm×4.6 mm I.D. column, packed with Hypersil BDS-C_{18}, 3 μm particle size. The substances are eluted in a gradient elution mode at a flow-rate of 1.5 ml/min and at

Table 1
Composition of the mobile phase during the solvent gradient (% volume fractions); A = 0.25% (m/v) ammonium acetate in water, B = acetonitrile, C = water

Solvent	Time (min)				
	0	13	15	17	22
A	50	50	50	50	50
B	25	43	43	25	25
C	25	7	7	25	25

ambient temperature. The solvent gradient used is shown in Table 1. UV detection is at 265 nm. The injection volume is 10 μl. The columns used in this study were (i) Alltech Hypersil 3 μm BDS-C_{18} (Laarne, Belgium) and (ii) Phenomenex Prodigy 3 μm ODS (3) 100 Å C_{18} (Macclesfield, UK). The chromatograph consisted of a Waters Alliance 2690 separation module and a Waters 486 tunable absorbance detector (Milford, MA, USA).

2.4. Calculations

Chromatographic parameters are calculated as prescribed by the USP [6]. The content of the main compound and of each related compound in the sample solution is calculated using the results for the corresponding peak obtained with the reference solution.

2.5. Software

The choice of the experimental design and of the experimental sequence was done with the software package Statgraphics Plus 2.1 for Windows (Manugistics, Rockville, USA). The calculation of effects and their statistical interpretation was also performed with the same software.

3. Results and discussion

The factors investigated in the robustness evaluation of the high-performance liquid chromatography (HPLC) method for identification and assay of ridogrel and its related compounds in ridogrel oral film-coated tablet simulations are summarised in Table 2, while the studied responses are given in Table 3. This latter table also shows the expected values under nominal conditions and the SST limits

Table 2
Factors investigated in the design

Factor	Units	Limits	Level (−1)	Level (+1)	Nominal
(1) Flow of the mobile phase	ml/min	± 0.1	1.4	1.6	1.5
(2) pH of the buffer	–	± 0.3	6.5	7.1	6.8
(3) Column temperature	°C	± 5	23	33	28°C
(4) Column manufacturer			Alltech	Prodigy	Alltech
(5) Percentage organic solvent (% B) in the mobile phase at the start of the gradient	%	±1	24	26	25
(6) % B in the mobile phase at the end of the gradient	%	±2	41	45	43
(7) Concentration of the buffer	%, m/v	±10%	0.225	0.275	0.25
(8) Detection wavelength	nm	±5	260	270	265

Table 3
Responses studied (MC = main compound, RC1 = related compound 1, RC2 = related compound 2)

Response	Substances considered	Expected value at nominal levels	SST limits
(1) Critical resolution (R_s)	MC-RC1	5.7	4.1
(2) Capacity factor (k')	MC	3.6	3.4
(3) Tailing factor (Asf)	MC	1.5	1.7
(4) Analysis results	MC, RC1, RC2	100%	No limit set
(5) Analysis time (t_R)	RC2	9.9 min	No limit set

that were established before the robustness test was applied.

The low and high levels of the factors in Table 2 were selected according to a standard procedure applied in the Janssen Research Foundation. Some were chosen as a constant percentage above (+) and below (−) the nominal level. In general, the extreme factor levels can be selected based on the uncertainty with which a factor level can be set [4]. Some factors are quantitative (continuous) while others are qualitative (discontinuous), e.g., the column manufacturer. The eight factors were examined in a Plackett–Burman design for 11 factors requiring 12 experiments (Table 4). In the three spare columns (randomly selected) dummy factors are entered. These are imaginary variables of which the change from one level to the other does not cause a physical change in the method. The effects estimated from these dummies can be used as a measure for the experimental error on an effect and therefore also in the statistical evaluation of the effects of the tested factors [7].

For each of the 12 experimental design runs, four injections were performed: a blank injection, two injections of the reference solution and an injection of the sample solution. From the second injection of the reference solution, the system suitability test parameters were determined. The two reference solution injections were used as calibration runs in order to estimate the recovery of the main and related compounds in the sample solution.

A typical chromatogram for a reference solution obtained at nominal conditions, is shown in Fig. 1. Table 5 shows the experimentally obtained design values for the responses that are studied. As can be observed the recoveries of the main peak (% MC) range from 98.4% to 102.3%, those of the first related compound (% RC1) from 97.1% to 103.0% and those of the second related compound (% RC2) from 98.7% to 103.1%, the resolution (R_s) between MC and RC1 from 4.96 to 7.48, the capacity factors [k'(MC)] from 3.17 to 5.82, the tailing factors [Asf(MC)] from 0.81 to 1.55 and the analysis times [t_R(RC2)] from 8.42 to 13.80 min.

The effects of the factors on a response are calculated as

$$E_X = \frac{\sum Y(+1)}{n} - \frac{\sum Y(-1)}{n} \quad (1)$$

where E_X is the effect of factor X; $\Sigma Y(+1)$ and $\Sigma Y(-1)$ are the sums of the responses where factor X was at level (+1) and at level (−1), respectively and n is the number of runs from the design where the factor was at level (+1) or at level (−1), usually equal to $N/2$ with N the number of design experiments.

Table 4
The Plackett–Burman design (−1 = low factor level, 1 = high factor level)

Experiment No.	Factors[a]										
	A pH	B Column	C Dum1	D Temp.	E % B begin	F % B end	G Dum2	H Flow	I Wavelength	J Buffer conc.	K Dum3
1	1	1	1	−1	1	1	−1	1	−1	−1	−1
2	1	1	−1	1	−1	−1	−1	1	1	1	−1
3	1	−1	1	1	−1	1	−1	−1	−1	1	1
4	1	−1	−1	−1	1	1	1	−1	1	1	−1
5	1	−1	1	−1	−1	−1	1	1	1	−1	1
6	−1	1	1	1	−1	1	1	−1	1	−1	−1
7	−1	1	−1	−1	−1	1	1	1	−1	1	1
8	−1	−1	−1	1	1	1	−1	1	1	−1	1
9	−1	−1	1	1	1	−1	1	1	−1	1	−1
10	−1	1	1	−1	1	−1	−1	−1	1	1	1
11	1	1	−1	1	1	−1	1	−1	−1	−1	1
12	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1

[a] Abbreviations: pH, pH of the buffer; Column, column manufacturer; Dum1, Dum2, Dum3, dummy variables; Temp., column temperature; % B begin, percentage B in the mobile phase at the start of the gradient; % B end, percentage B in the mobile phase at the end of the gradient; Flow, flow of the mobile phase; Wavelength, wavelength of the detector; Buffer conc., concentration of the buffer.

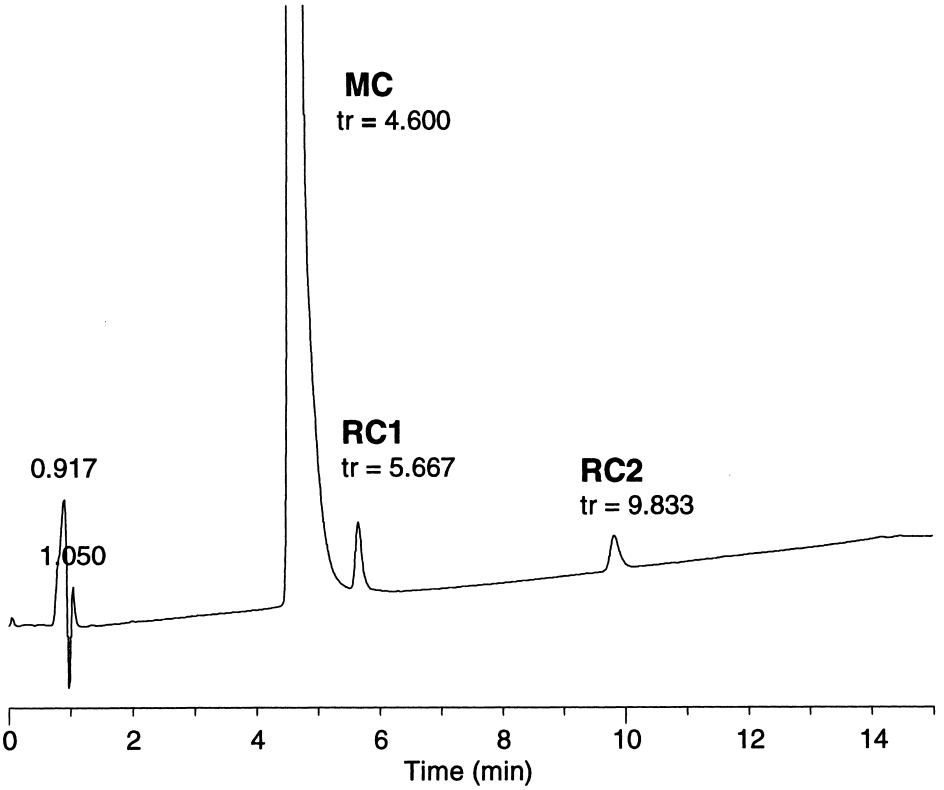

Fig. 1. Chromatogram of the reference solution containing 100% main compound and 0.5% related compounds under nominal method conditions (Alltech column). MC = Main compound, RC1 = related compound 1, RC2 = related compound 2. tr = retention time in min.

Table 5
Results of the experiments of Table 4

Experiment	Responses						
	% MC	% RC1	% RC2	R_s(MC-RC1)	k'(MC)	Asf(MC)	t_R(RC2)
1	101.6	100.9	101.4	5.691	3.800	0.813	11.500
2	101.7	101.2	102.7	7.484	5.083	1.031	13.000
3	101.6	101.7	101.3	5.770	4.000	1.453	9.833
4	101.9	103.0	102.9	5.025	3.167	1.549	9.483
5	101.8	99.3	99.1	5.440	3.800	1.458	10.317
6	101.1	99.9	101.7	5.711	5.817	0.861	12.567
7	101.1	100.8	101.4	5.932	5.250	0.836	12.083
8	101.6	100.2	98.8	4.962	3.200	1.059	8.417
9	98.4	97.1	101.8	5.427	3.367	0.977	9.200
10	99.7	100.5	99.3	6.344	5.350	0.853	13.800
11	99.7	98.6	98.7	6.715	4.783	0.920	13.317
12	102.3	101.1	103.1	5.186	4.933	1.412	11.150
Mean	101.0	100.4	101.0				
RSD (%)	1.15	1.52	1.61				

Table 6
Analysis of variance (ANOVA) table for the interpretation of the significance of effects on the response ''capacity factor of MC''

Source	Effect	Sum of squares	Df	Mean square	*F* ratio	*P* value
(A) pH	−0.547	0.8987	1	0.8987	203	0.0007
(B) Column	1.269	4.834	1	4.834	1092	0.0001
(D) Temp.	−0.008	0.0002	1	0.0002	0.05	0.8421
(E) % B begin	−0.869	2.267	1	2.267	512	0.0002
(F) % B end	−0.347	0.3612	1	0.3612	81	0.0029
(H) Flow	−0.592	1.050	1	1.050	237	0.0006
(I) Wavelength	0.047	0.0067	1	0.0067	1.52	0.3056
(J) Buffer concentration	−0.019	0.0011	1	0.0011	0.25	0.6493
Total error		0.0132	3	0.0044		

To identify significant effects an analysis of variance (ANOVA) table is created. An example is shown in Table 6. The principles of this approach can be found in Refs. [7,8]. The sum of squares $(SS)_X$ for a factor can be calculated as $[E_X N/2]^2/N$ and the one for total error is the sum of the sums of squares from the dummies. The mean square for a factor, $(MS)_X$, is the $(SS)_X$ divided by the number of degrees of freedom (df). The *F* ratio is obtained by dividing $(MS)_X$ by $(MS)_{\text{total error}}$ and the *P* value gives an indication of the significance of an effect. It can be considered as the probability of taking the wrong decision when accepting that an effect is significant. Therefore when the *P* value is below the considered level of confidence α, an effect is considered to be statistically significant. For example, when $P<0.01$ then an effect is significant at $\alpha=0.01$.

The standard error on an effect ($SE_e=0.0384$ with three degrees of freedom in this example) is calculated as $\sqrt{\frac{\Sigma E_{d_i}^2}{n_{d_i}}}$ where is the effect of a dummy and the number of dummies used.

The effects of the different factors on the considered responses and the corresponding *P* values, extracted from the ANOVA tables, are shown in Table 7. The factors having a statistically significant effect on a response, at significance levels of 5% ($P<0.05$) and of 10% ($P<0.1$), were indicated in Table 7, part b. From this Table it can be observed that none of the factors has a significant effect on the determination of the recovery of the main and related compounds. From Table 5 it was already observed that the ranges of the analysis results were narrow and the percent relative standard deviations were small (1.2%, 1.5% and 1.6% for MC, RC1 and RC2, respectively). Based on these facts the method for assay of ridogrel and its related compounds can be considered robust.

However, looking at the effects on the other responses, which describe the performance of the method under the different design conditions, it can be seen that several effects are significant (Table 7). The column for instance has a significant effect on all these responses. The responses, capacity factor and analysis time are affected by several of the tested factors.

As already stated elsewhere [10,12], a statistical significant effect on those responses is not always chromatographically relevant. To evaluate this relevance, one could firstly look at the most extreme results from the design experiments and compare them with the existing SST limits. The most extreme resolution (4.96) and tailing factor (1.55) from the design results are within the SST specifications, namely above 4.1 and below 1.7, respectively, while the capacity factor (3.17) is not since it is below 3.4 (see Tables 3 and 5). However, the most extreme design results are not necessarily the worst results that can be obtained in the examined experimental domain since the combination of factor levels that gives this worst result is not necessarily executed in the design. For instance, the worst-case situation for resolution is the factor combination giving the lowest resolution. For the capacity factor it is the one causing the smallest capacity factor, while for the tailing factor it is the situation resulting in the highest value. To decide on the conditions of this

Table 7
(a) Effects of the factors on the different responses, (b) P values obtained for these effects

Factors	(a) Effects on						
	% MC	% RC1	% RC2	R_s(MC-RC1)	k'(MC)	Asf(MC)	t_R(RC2)
pH	0.683	0.850	0.000	0.427	−0.547	0.204	0.039
Column	−0.450	−0.083	−0.300	1.011	1.269	−0.432	2.978
Dum1	−0.683	−0.917	−0.500	−0.154	−0.047	−0.065	−0.039
Temp.	−0.717	−1.150	−0.367	0.408	−0.008	−0.103	−0.333
% B begin	−1.117	−0.617	−1.067	−0.226	−0.869	−0.147	−0.539
% B end	0.883	1.450	0.467	−0.584	−0.347	−0.013	−1.150
Dum2	−0.750	−1.150	−0.167	−0.198	−0.030	−0.003	−0.122
Flow	−0.017	−0.883	−0.300	0.031	−0.592	−0.146	−0.939
Wavelength	0.517	0.650	−0.533	0.041	0.047	0.067	0.084
Buffer concentration	−0.617	0.717	1.100	0.380	−0.019	0.029	0.022
Dum3	−0.250	−0.350	−2.500	0.106	0.036	−0.011	0.144
	(b) P values						
pH	0.340	0.402	1.000	**0.073	*0.0007	*0.013	0.751
Column	0.510	0.930	0.852	*0.008	*0.0001	*0.002	*0.0001
Temp.	0.320	0.279	0.820	**0.080	0.842	**0.074	**0.058
% B begin	0.161	0.531	0.522	0.245	*0.0002	*0.031	*0.017
% B end	0.239	0.195	0.772	*0.034	*0.0029	0.751	*0.002
Flow	0.980	0.386	0.852	0.857	*0.0006	*0.032	*0.004
Wavelength	0.455	0.510	0.742	0.812	0.306	0.180	0.508
Buffer concentration	0.382	0.472	0.510	**0.094	0.649	0.499	0.857

**=Significance at $\alpha=0.10$ level, *=significance at $\alpha=0.05$ level.

worst-case experiment only the statistically significant factors (at $\alpha=0.05$) and the ones that come close to it ($P<0.1$, significant at $\alpha=0.1$) were considered. These factors were included since they were thought able to cause a systematic change in a response when changed from one level to the other. The factors with a $P>0.1$ were considered as negligible and their effects were considered to originate only from experimental error. As the executed experimental design is a saturated two-level design, only linear effects for the maintained factors can be considered in the prediction of the worst-case situation. This can be justified by the fact that in robustness testing one is working in such a restricted domain of the response surface that only linear effects are important. The factor level combination leading to the worst result for a response Y is predicted by the equation:

$$Y = E_{F_1}F_1 + E_{F_2}F_2 + ... + E_{F_k}F_k \quad (2)$$

In Eq. (2) E_{F_i} represents the effect of the factor considered for the worst-case experiment and F_i the level of this factor. Non-important factors ($P>0.1$) are kept at nominal value ($F_k=0$). The worst-case factor-level combinations for the different responses when applying the above rules are shown in Table 8. The worst-case experiment is carried out in three independent replicates and the mean result was then compared with the system suitability limit by a one-sided t-test (the one-sample case [13]) to investigate whether the system suitability limit is statistically violated. The results of the worst-case experiments

Table 8
Predicted worst-case factor-level combinations for the different responses

Factors	Responses		
	R_s(MC-RC1)	k'(MC)	Asf(MC)
pH	−1	+1	+1
Column	−1	−1	−1
Temp.	−1	0	−1
% B begin	0	+1	−1
% B end	+1	+1	0
Flow	0	+1	−1
Wavelength	0	0	0
Buffer concentration	−1	0	0

Table 9
Results of the worst-case experiments for the different responses and of the *t*-tests

Run	R_s(MC-RC1)	k'(MC)	Asf(MC)
1	4.870	2.800	1.453
2	4.819	2.800	1.483
3	4.702	2.817	1.429
Mean	4.797	2.806	1.455
SD (s)	0.0861	$9.81 \cdot 10^{-3}$	0.0271
n	3	3	3
SST limit	4.1	3.4	1.7
t value	14.02	−104.88	−15.685
$t_{critical}$ ($\alpha = 0.05$, 2 df)	2.92	2.92	2.92
	SST limits from worst case results		
	$4.797 - 2.92\frac{0.0861}{\sqrt{3}} = 4.65$	$2.806 - 2.92\frac{0.0098}{\sqrt{3}} = 2.79$	$1.455 + 2.92\frac{0.0271}{\sqrt{3}} = 1.59$

for the different responses and of the *t*-tests are shown in Table 9. For the resolution it is found that the worst case result is not significantly smaller than the SST limit (hypotheses tested: H_0: $R_s >= 4.1$; H_1: $R_s < 4.1$). The capacity factor, however, is found to be significantly smaller than the SST limit (hypotheses tested: H_0: $k' \geq 3.4$; H_1: $k' < 3.4$). The tailing factor is not found to be significantly larger than the limit (hypotheses: H_0: Asf ≤ 1.7; H_1: $R_s > 1.7$).

These results indicate that when the method is transferred to another laboratory, it is possible that some SST criteria will be violated. In our example, it could be the case for the capacity factor. However, from the results of the robustness test it is known that the quantitative results of the method are robust. This shows that a more or less arbitrary selection of system suitability test parameter limits can lead to problems, which are not relevant to quality and therefore highly undesirable.

A better approach is to derive the system suitability limits from the results of the experimental design. This has already been proposed by Mulholland et al. [3,14]. They define the system suitability limits for a response as the extreme results from the design. Since these extreme results are not necessarily the worst ones possible we propose to use the results of the worst-case situations to define the SST limits. An advantage of this approach, besides avoiding ambiguous situations, is that the system suitability limits are established from the robustness test results as is recommended by the ICH guidelines. A possible procedure is to define the SST limit as the upper or lower limit from the one-sided 95% confidence interval [15] around the worst-case mean. For the resolution and the capacity factor, for instance, the lower limit would be chosen, while for the tailing factor it would be the upper one. The confidence interval is defined as $\left[\bar{X}_{\text{worst-case}} - t_{\alpha,n} \cdot (\frac{s}{\sqrt{n}}),\ \infty\right]$ when the lower limit needs to be considered and as $\left[\bar{X}_{\text{worst-case}} + t_{\alpha,n} \cdot (\frac{s}{\sqrt{n}})\right]$ when it is the upper one. This would lead to system suitability limits of 4.65 for the resolution, of 2.79 for the capacity factor and of 1.59 for the tailing factor (see Table 9). It can be seen that with this approach some SST limits are defined stricter than the previously used ones (resolution, tailing factor) while others are less (capacity factor).

The idea behind applying this approach is the following. In the experimental domain of the design it was found that the quantitative response was not significantly affected by any factor. Therefore it can be expected that in none of the points of the experimental domain there would be a problem with the quantitative response, including those at which certain responses have their worst result. In cases where one does not a priori accept the hypothesis that the worst case conditions do not affect the

quantitative results they can easily be verified in practice. The SST limit for a response like the resolution is then defined as the confidence limit above which one has to be situated in 95% of the experiments executed at these worst case conditions. This means that considering all possible acceptable experimental conditions (since the worst-case conditions will only be met rarely in a laboratory), the response (resolution) is situated above this limit with a probability that is far above 95%.

Considering the latter fact, a less strict and easier alternative, which in this situation also could lead to acceptable SST limits would be to choose the worst case result, occasionally even determined without replicates, as the SST limit.

Beside the recommendation of the ICH guidelines also practical reasons were handled for defining SST limits based on the results of a robustness test. From experience with robustness testing it was observed that when a separation method was properly optimised, the quantitative results did not change significantly, although some of the SST limits, selected rather arbitrarily and independently from the results of a robustness test, were frequently violated. The reason for it is that the limits were set too strictly during method optimisation. On the other hand, it is not considered desirable either to choose as SST limit the most extreme situation that still allows a quantitative determination. For instance, when the operational conditions after method development gave a resolution of about six, a resolution of two is not considered acceptable, even if quantitation still is possible. The reason for this is that at all times one likes to maintain the method close to the conditions at which it was validated. Therefore it was thought preferable to derive the system suitability limits from the robustness test, since it explores the most extreme variations in the factors that are likely to occur.

4. Conclusions

The conclusion for the robustness test of the chromatographic method for the analysis of ridogrel and its related compounds in film-coated tablets, is that the method is robust concerning the analysis results of the main and related compounds.

Defining system suitability limits based on the worst-case results for which the conditions were predicted from the robustness test, allows to avoid an undesirable situation where a method is found to be robust for its quantitative aspect while some externally defined system suitability criteria are violated.

The column is rather crucial as it affects the resolution, the capacity factor, the tailing factor and the analysis time. While the use of an Alltech column is acceptable for this method, a Prodigy column could be a good alternative. Inclusion of only two columns in the study does not allow drawing conclusions about the population of columns, i.e., about the robustness of the method towards the particular type of columns to which the two selected ones belong. Only conclusions about the robustness of the method towards the two examined columns can be made and any extrapolation to whatever other column is not allowed.

Acknowledgements

The authors thank the Fund for Scientific Research (FWO)–Vlaanderen, for financial support.

References

[1] ICH Harmonised Tripartite Guideline, Validation of Analytical Procedures – Methodology, The International Conference on Harmonisation of Technical Requirements for registration of Pharmaceuticals for Human Use, 6 November 1996.
[2] J.A. Van Leeuwen, L.M.C. Buydens, B.G.M. Vandeginste, G. Kateman, P.J. Schoenmakers, M. Mulholland, Chemomet. Intell. Lab. Systems 10 (1991) 337–347.
[3] M. Mulholland, Trends Anal. Chem. 7 (1988) 383–389.
[4] Y. Vander Heyden, F. Questier, D.L. Massart, J. Pharm. Biomed. Anal., in press
[5] W.J. Youden, E.H. Steiner, Statistical Manual of the Association of Official Analytical Chemists, The Association of Official Analytical Chemists, Arlington, VA, 1975, pp. 33–36, 70–71, 82–83.
[6] The United States Pharmacopeia, 23rd ed, National Formulary 18, United States Pharmacopeial Convention, Rockville, 1995.
[7] Y. Vander Heyden, D.L. Massart, in: A. Smilde, J. de Boer, M. Hendriks (Eds.), Robustness of Analytical Methods and Pharmaceutical Technological Products, Elsevier, Amsterdam, 1996, pp. 79–147.

[8] E. Morgan, Chemometrics – Experimental Design, Analytical Chemistry by Open Learning, Wiley, Chichester, 1991, pp. 118–188.
[9] M. Jimidar, M.S. Khots, T.P. Hamoir, D.L. Massart, Quím. Anal. 12 (1993) 63–68.
[10] M. Jimidar, N. Niemeijer, R. Peeters, J. Hoogmartens, J. Pharm. Biomed. Anal., in press
[11] R.L. Plackett, J.P. Burman, Biometrika 33 (1946) 305–325.
[12] Y. Vander Heyden, C. Hartmann, D.L. Massart, L. Michel, P. Kiechle, F. Erni, Anal. Chim. Acta 316 (1995) 15–26.
[13] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics – A Textbook, Elsevier, Amsterdam, 1988, pp. 41–48.
[14] M. Mulholland, J. Waterhouse, J. Chromatogr. 395 (1987) 539–551.
[15] 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. 73–79.