Journal of Chromatography, 395 (1987) 539–551
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1230

DEVELOPMENT AND EVALUATION OF AN AUTOMATED PROCEDURE FOR THE RUGGEDNESS TESTING OF CHROMATOGRAPHIC CONDITIONS IN HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

M. MULHOLLAND*
Philips Analytical, Cambridge (U.K.)
and
J. WATERHOUSE
Glaxo Group Research, Ware, Hertfordshire (U.K.)

SUMMARY

When analytical methods are developed for use in many different laboratories it is particularly important to be aware of the sensitivity of the method to variations in the specified conditions. It was the aim of this research to establish a ruggedness test which would enable the user: (1) to decide which variables (flow-rate, etc.) have a significant effect on the chromatographic results; (2) to define sensible maximum acceptable deviations from the specified variables setting; (3) to define objective system suitability parameters. The test uses a system of fractional factorial experiments, based on the Plackett-Burman design schemes. The test calculates main effects and standard errors. It has also been expanded to include the methodology to test two extreme values for each variant.

Both eight-experiment and sixteen-experiment Plackett-Burman schemes were investigated initially, using the well-understood high-performance liquid chromatographic analysis of aspirin and salicylic acid. Several variables were tested, including flow-rate, acid type, mobile phase composition, detection wavelength, detector attenuation and response time. Both qualitative (retention time, plate count, peak symmetry, etc.) and quantitative (peak area, height, concentration, etc.) parameters were checked for variations.

By testing the ruggedness of chromatographic methods in this manner, the analyst can achieve a comprehensive understanding of the limitations and stability of the methodology employed.

INTRODUCTION

High-performance liquid chromatography (HPLC) requires optimisation of several variables to arrive at an ideal method for a given application. Before this method can be accepted for extensive usage, the analyst must perform a programme of validation experiments. This usually consists of tests for accuracy. specificity, repeatability, reproducibility, linearity of detector response and homogeneity of chro-

0021-9673/87/$03.50

matographic peaks. However, if the method is to be used outside the constraints in which these tests were performed, *e.g.* different equipment and laboratories, the analyst must establish its limitations with changes in the specified conditions. This is normally referred to as "ruggedness testing". This paper describes the derivation and evaluation of a ruggedness test for HPLC methodology.

The performance of a HPLC method can be affected by instrumental parameters, quality and type of materials and changes in temperature. Instrumental parameters vary within the range specified by the manufacturer and also with age and condition of equipment. The true value of detection wavelength, monitored by uncalibrated detectors can vary by several nanometers, the accuracy of solvent composition and flow-rates depends on several instrumental features, particularly the type of solvent viscosity compensation employed. Choice of materials (*e.g.* column type and manufacture) can also significantly effect method performance. Changes in temperature also affect the chromatogram, especially if it is specified as ambient (uncontrolled). By ruggedness-testing an HPLC method, the analyst should achieve a comprehensive understanding of the limitations and stability of the method to these variations.

THEORY

Fractional factorial experiment design

To perform a full factorial design[1], the chromatographer selects a fixed number of "levels" (or "versions"), l, for each of a number of chromatographic variables, n, and then runs experiments with all possible combinations. The number of experiments, N_{exp}, required can be calculated from:

$$N_{exp} = l^n \tag{1}$$

Hence, to test six chromatographic variables at two levels, the nominal method value and one extreme value, $N_{exp} = 64$. This number of experiments is so large as to discourage the use of full factorials for routine ruggedness testing.

From such a test, a complete set of main effects, ME, and interaction effects can be calculated, where

$$\mathrm{ME} = \frac{[\Sigma(x)_1 - \Sigma(x)_2]}{N_{exp}} \tag{2}$$

$(x)_1$ is the mean value of a parameter calculated from experiments using the nominal method level and $(x)_2$ is the value of a parameter calculated from experiments using the extreme level. (For calculation of interaction effects *cf.* ref. 1.)

Many of these statistics can be considered redundant if we assume higher-order interaction effects to be so small as to be negligible. A half-fractional factorial design ignores all interaction effects higher than first-order. For this scheme

$$N_{exp} = l^{(n-1)} \tag{3}$$

hence for six variables at two levels, $N_{exp} = 32$. This is still an impractical number

of experiments, particularly when we consider that for chromatography it would be desirable to test two extreme levels. Plackett and Burman produced saturated fractional factorial designs which allow the unbiased estimation of all main effects for $N_{exp} - 1$ variables[2]. They derived designs for $N_{exp} = 8$ and $N_{exp} = 16$ and in this paper these designs were evaluated with regards to their suitability for a HPLC ruggedness test.

To design their fractional factorial schemes Plackett and Burman calculated the arrangement of levels for the first variable; for $N_{exp} = 8$, arrangement is 1110100; and for $N_{exp} = 16$, the arrangement is 111101011001000 where 0 is the nominal method level and 1 is the extreme level.

The complete design is then constructed by taking this as the first column (or row), shifting it cyclically by one place $N_{exp} - 2$ times and adding a final row of 0 signs (this represents the method performed at originally specified conditions). To test the ruggedness of a HPLC method fully, extremes at either side of the nominal method value should be examined. This can be done by performing a second identical scheme consecutively using a second extreme level. Both schemes are treated individually and two sets of independent main effects are calculated. However, the total number of experiments will be $2N_{exp} - 1$, as the experiment using nominal method conditions will be the same for both schemes. The complete schemes derived for $N_{exp} = 8$ is shown in Table I, using -1 to denote the second extreme level. The schemes are drawn for six variables.

TABLE I

EXPERIMENTAL SCHEME DERIVED FOR $N_{exp} = 8$

Experiment	*Scheme for six variables*					
1	1	0	0	1	0	1
2	1	1	0	0	1	0
3	1	1	1	0	0	1
4	0	1	1	1	0	0
5	1	0	1	1	1	0
6	0	1	0	1	1	1
7	0	0	1	0	1	1
8	0	0	0	0	0	0
9	−1	0	0	−1	0	−1
10	−1	−1	0	0	−1	0
11	−1	−1	−1	0	0	−1
12	0	−1	−1	−1	0	0
13	−1	0	−1	−1	−1	0
14	0	−1	0	−1	−1	−1
15	0	0	−1	0	−1	−1

Each scheme is treated as two individual fractional factorials where main effects and standard errors, SE, which estimate the average error between duplicate experiments, are calculated for each variable. The standard errors are calculated as follows:

$$SE = \sqrt{\frac{4}{2 \cdot N_{exp}} \cdot \frac{\Sigma(di^2)}{2_g}} \tag{4}$$

where *di* is the difference between duplicate experiments and *g* is the number of degrees of freedom = N_{exp}. Eqn. 4 can be simplified as follows:
for $N_{exp} = 8$

$$SE = \sqrt{[0.0156 \cdot \Sigma(di^2)]} \tag{5}$$

for $N_{exp} = 16$

$$SE = \sqrt{[3.91 \cdot 10^{-3} \cdot \Sigma(di^2)]} \tag{6}$$

The results obtained for all the different parameters tested (*e.g.*, retention time, peak area, resolution) are calculated as a percentage of the parameter value at nominal method conditions, *x*, hence

$$\%ME = \frac{ME}{x} \cdot 100 \tag{7}$$

and

$$\%SE = \frac{SE}{x} \cdot 100 \tag{8}$$

A relative standard deviation (R.S.D.) is calculated for each set of results as a reassurance that the main effects and standard errors accurately reflect the level of variation.

Full factorial schemes have previously been applied to HPLC for method optimisation[3,5]. Plackett-Burman fractional factorials are most often applied to the optimisation of production processes[6,7]. However, some papers have been published on their application to ruggedness testing of analytical techniques other than chromatography[8,9].

Selection of chromatographic calculations

Having decided on a suitable statistical technique, the next problem encountered is which chromatographic parameters should be studied to give a complete evaluation. The primary function of a ruggedness test is to investigate the effect of chromatographic variables on the quantitative results (*e.g.* concentration by peak area). However, further useful information can be obtained by also investigating effects on the quality of method performance (*e.g.* retention time, resolution). For instance, the qualitative results obtained throughout the test can be examined to find the range of values within which the method remains quantitative. This allows the analyst to decide objectively whether a given chromatographic system (instrumental, columns) is suitable for use of the tested method. This range of parameter values is often referred to as system suitability parameters.

The analysis selected for the initial evaluation of the ruggedness test used external standardisation.

Thus, each experiment consisted of six injections as follows: (1) calibration 1, (2) calibration 1, (3) sample 1, (4) sample 1, (5) calibration 2 and (6) calibration 2.

These are treated as two duplicate experiments, 1, 3, 5 and 2, 4, 6. The following parameters are calculated for each experiment:

(1) Concentration of each component in the sample using both peak heights and peak areas.

(2) Mean retention time for each component, t.

(3) Sample peak area, a, and sample peak height, h.

(4) The mean number of theoretical plates is N, where

$$N = \left[\frac{\sqrt{(2\pi)} \cdot h \cdot t}{a}\right]^2 \tag{9}$$

(5) The mean resolution between each peak and its nearest neighbouring peak is R_s, where

$$R_s = \tfrac{1}{2}\frac{(t_2 - t_1)}{(t_1 + t_2)} N^1 \tag{10}$$

where N^1 = mean N between both peaks.

(6) Mean peak symmetry, S_y,

$$S_y = \frac{t - t_s}{t_e - t}$$

where t_s and t_e are the end and start times, respectively.

Automation of the data handling/reduction

The flow chart for the data handling/reduction is shown in Fig. 1. The number of peaks was found for the calibration solution analysed under nominal method conditions. For each new experiment, the retention times for this number of peaks is established from the calibration solution. This information is then used to identify the sample peaks.

Selection of test analysis

The analysis of aspirin and salicylic acid was selected to evaluate the proposed ruggedness test. Many authors have published work on this assay[10–14]; thus, a substantial amount of information was available on which variables to select for investigation. The method is also sufficiently complex to enable a thorough evaluation of the ruggedness test.

EXPERIMENTAL

Materials

Reference materials for aspirin and salicylic acid, and perchloric, acetic and orthophosphoric acid were obtained from Sigma (St. Louis, MO, U.S.A.). All solvents were of HPLC grade (Rathburn Chemical, Edinburgh, U.K.). (Solvents were continually degassed by sparging with helium). Aspirin tablets were from Cox Pharmaceuticals (A. M. Cox and Co., Barnstaple, U.K.).

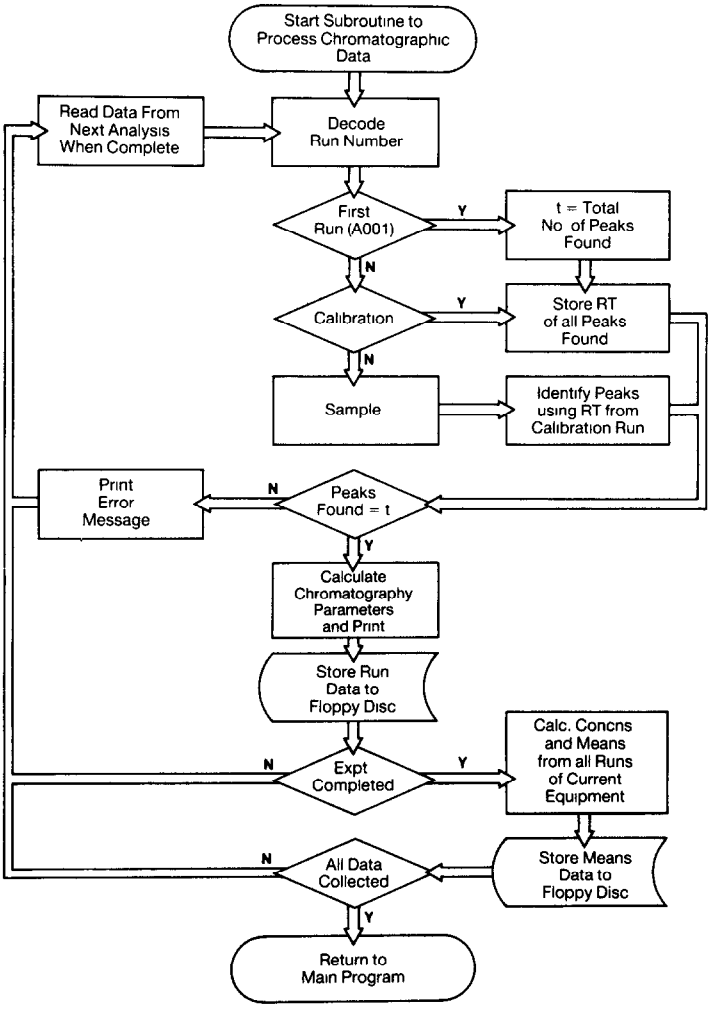

Fig. 1. Flow-chart for the data handling/data reduction software.

Instrumentation

All components of the HPLC system were from Philips Analytical (Cambridge, U.K.) a PU4100 chromatograph, fitted with a 10-μl injection loop; an oven; a PU4100 UV detector, equipped with an 8-μl flow-cell. A PU4021 multichannel UV–VIS detector, equipped with an 8-μl flow-cell was employed for the preliminary method validation. The detector outputs were connected to a PU4850 data station, for which software was written to perform the ruggedness test. The system was fitted with a 25 cm $\times$ 4 mm I.D. Lichrocart C_{18} cartridge column with a Lichrocart C_{18} guard cartridge (BDH, Dagenham, U.K.).

Chromatographic conditions

Calibration solutions were prepared by dissolving 60 mg of aspirin or 1.8 mg of salicylic in 50 ml of acetonitrile–methanol–orthophosphoric acid (92:8:0.5, v/v/v). Information was available to show that the rate of degradation in this extraction solvent is insignificant[14].

Sample solutions were prepared by placing one tablet and 9 mg of salicylic acid in a 250-ml volumetric flask, adding 150 ml of the above solvent mixture and ultrasonicating for 15 min. The solution was then made up to volume with the solvent and filtered through a 0.45-μm Alpha-450 filter (Philips Analytical).

Aspirin and salicylic acid were eluted with a mobile phase of aqueous acetonitrile, acidified to pH 2.6, using orthophosphoric acid, at a flow-rate of 1.5 ml/min. For UV detection, 295 nm with 0.1 a.u.f.s. sensitivity and a response time of 0.5 s was used. The column temperature was maintained at 40°C.

Design of the ruggedness tests

The listing of selected variables and their extreme levels for the $N_{exp} = 8$ scheme is shown in Table II. Since equilibration to 35°C was slow the lower extreme level for temperature for $N_{exp} = 16$ was changed to 38°C. This enabled the completion of that scheme within the lifetime of sample and calibration solutions. A seventh variable was not included so that effects for an imaginary variable could be calculated as an error check.

TABLE II

LIST OF INFORMATION FOR VARIABLE LEVELS TO BE TESTED

Name of variant	*Method value*	*Minimum value*	*Maximum value*
1 Acetonitrile (%)	25	23	27
2 Acid type	1	2	3
3 Flow-rate (ml/min)	1.5	1.4	1.6
4 Temperature (%)	40	38	45
5 Wavelength (nm)	295	290	30
6 Response time (s)	0.500	0.120	2

The experimental order of both schemes was sorted on acid type, such that long equilibration times required for acid changes were minimized. The $N_{exp} = 8$ scheme is shown in Table III. When the ranges were selected the fact that only three levels were being tested was considered. For instance, the UV spectra of aspirin and salicylic acid were examined and the levels chosen did not ignore a spectral maximum within their range.

RESULTS AND DISCUSSION

Preliminary method validation

The initial method validation programme included the following tests, specificity, spectral purity of chromatography peaks, repeatability and linearity of detector response. Satisfactory results were obtained from these experiments.

TABLE III

EXPERIMENTAL ORDER BASED ON FRACTIONAL FACTORIAL PATTERN

Exp. No.	*Acetonitrile (%)*	*Acid type*	*Flow-rate (ml/min)*	*Temperature (°C)*	*Wavelength (nm)*	*Response time (s)*
8	25	1	1500	40	295	0.120
1	27	1	1500	45	295	2
5	27	1	1600	45	300	0.120
7	25	1	1600	40	300	2
9	23	1	1500	35	295	0.500
13	23	1	1400	35	290	0.120
15	25	1	1400	40	290	0.500
2	27	3	1500	40	300	0.120
3	27	3	1600	40	295	2
4	25	3	1600	45	295	0.120
6	25	3	1500	45	300	2
10	23	2	1500	40	290	0.120
11	23	2	1400	40	295	0.500
12	25	2	1400	35	295	0.120
14	25	2	1500	35	290	0.500

Ruggedness test

The sequence of injections for experiment 8 from the $N_{exp} = 8$ scheme is shown in Fig. 2. The calculation for symmetry was not very robust, due mainly to large variations in peak end times. Therefore it was not possible to monitor effects on the peak symmetry.

The effect of changing the acetonitrile composition was observed as a reduction in retention for both components with increasing solvent strength, peak height increasing correspondingly. Resolution was slightly affected but not enough to change

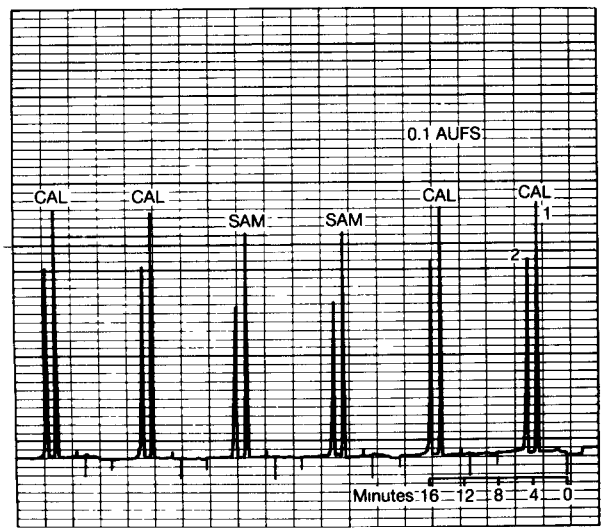

Fig. 2. Sequence of injections for experiment 8 ($N_{exp} = 8$). 1 = Aspirin, 2 = salicylic acid.

the concentration calculations. Many significant effects were observed on changing the acid used for pH control, showing both chromatographic changes (*e.g.*, retention and resolution) and detection changes (peak areas/heights). The spectra for both aspirin and salicylic acid experienced bathochromic shifts. Fig. 3 shows the changes in spectral characteristics for aspirin in the three different mobile phases.

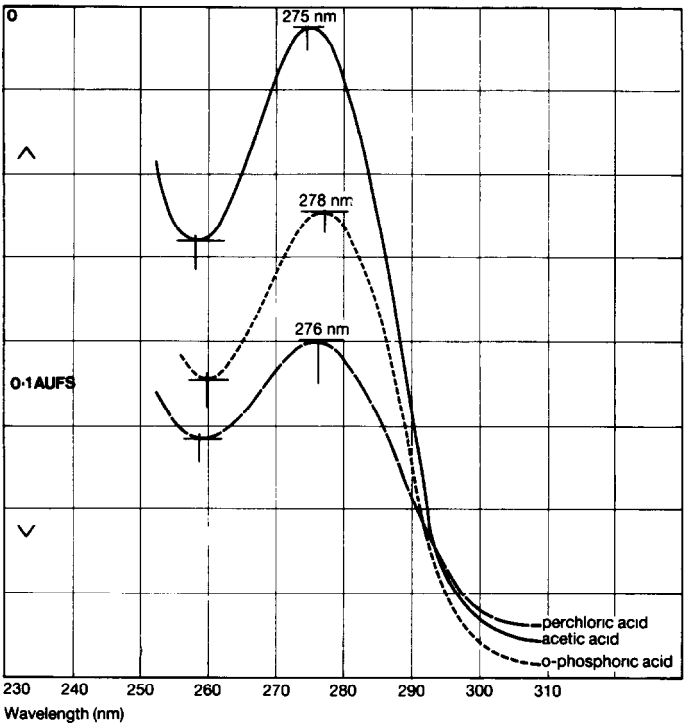

Fig. 3. Spectra of aspirin showing bathochromic shifts with different acids in the mobile phase.

Increasing the temperature was found to reduce retention and N. This caused a decrease in resolution and an increase in peak heights. An increase in flow-rate reduced retention and residence time of components in the flow cell, thus causing a decrease in peak areas.

The results obtained from both schemes, $N_{exp} = 8$ and $N_{exp} = 16$ compared well. Both data sets for wavelength effects are presented in Table IV. This indicates that the use of saturated fractional factorial designs is adequate for the ruggedness test.

The effects of wavelength changes were the most dramatic of all the variables tested, peak areas changing by up to 100%, but concentration values remained relatively unchanged. It must be noted, however, that limits of detection would be seriously affected.

The results for the effect of changes in response time for aspirin, observed for the $N_{exp} = 16$ scheme show an increase in peak heights on reducing the response

TABLE IV

EFFECT OF CHANGES IN WAVELENGTH

Wavelength (method) value = 295 nm

Chromatographic parameters		N_{exp} = 8		N_{exp} = 16	
		Aspirin	*Salicylic acid*	*Aspirin*	*Salicylic acid*
Retention time	290 nm	−2.078	−2.253	−0.424	0.923
	300 nm	−0.355	−0.274	0.032	0.697
Plates	290 nm	0.328	0.333	0.050	0.743
	300 nm	0.984	−0.010	0.238	−0.003
Resolution	290 nm	−0.021			
	300 nm	0.468			
Peak area	290 nm	−101.637	9.102	−102.234	9.863
	300 nm	35.194	−5.739	35.091	−6.574
Peak height	290 nm	−82.869	8.888	−91.610	7.335
	300 nm	39.312	−6.141	38.592	−7.254
Concentration (area)	290 nm	−1.008	−1 391	−0.004	−1.236
	300 nm	0.394	0.637	−0.555	0.297
Concentration (height)	290 nm	−0.507	−1.412	0.155	−1.875
	300 nm	0.333	0.533	0.036	−0.310

time, suggesting that 0.5 is slightly too slow to detect the peak without distortion. However, this did not affect the quantitative results.

Having examined the results obtained for the ruggedness test, the analyst can now use these in several ways. If the results did not prove ruggedness over the ranges tested the method could be redefined, using the results obtained. Table V shows the optimum conditions selected for aspirin and salicylic acid and a compromise for both compounds. Optimum levels are chosen for (a) increased sensitivity, (b) increased N

TABLE V

RESPECIFICATION OF THE CHROMATOGRAPHIC METHOD

Method parameter	*Original method value*	*Optimum value for aspirin*	*Optimum value for salicylic acid*	*Optimum compromise value*	*Method ruggedness range of values*
Acetonitrile (%)	25	25	27	25	23–27
Acid type	Orthophosphoric acid	Orthophosphoric acid	Perchloric acid	Perchloric acid	Any of the three acid types
Flow-rate (ml/min)	1.5	1.4	1.4	1.4	1.4–1.6
Temperature (°C)	40	45	45	45	35–45
Wavelength (nm)	295	300	295	290	290–300
Response time (s)	0.5	0.12	0.12	0.12	0.12–2

and optimum speed of analysis with given resolution. The redefined method now requires full validation, including a ruggedness test. This is a very simplistic way of optimising the method, which may not be adequate for methods that show very poor ruggedness. Table V also shows the specification range of values which can be defined if the ruggedness test has shown tolerable effects.

If the results are examined for fluctuations in qualitative parameters that do not affect the quantitative results, it is possible to select a range of values for the system suitability parameters. For instance, Table VI shows the complete set of results generated for effects on N measured for aspirin (N_{exp} = 8). N varies from 2450 to 2860 without affecting the quantitative results. The full set of suitability parameters can be derived in this way, and these are shown in Table VII. It is important to note

TABLE VI

STATISTICS REPORT FOR THE RUGGEDNESS OF N CALCULATED FOR ASPIRIN

Component: Aspirin. Parameter: number of plates.

Exp.	*Replicate 1*	*Replicate 2*	*Mean*	*Diff. 2*
1	2594.498	2651.988	2623.239	3304.201
2	2591.226	2563.362	2577.294	776.444
3	2465.545	2470.684	2468.114	26.406
4	2494.086	2413.076	2453.581	6562.543
5	2524.794	2586.986	2555.890	3867.801
6	2487.411	2385.297	2436.354	10427.222
7	2611.641	2616.789	2614.215	26.501
8	2871.246	2857.535	2864.390	187.976
	% Std error = 0.692	R.S.D. = 5.362		
8	2871.246	2857.535	2864.390	187.976
9	2839.625	2778.809	2809.217	3698.665
10	2583.133	2646.321	2614.727	3992.722
11	2644.862	2611.584	2628.223	1107.414
12	2646.538	2657.371	2651.954	117.359
13	2786.368	2872.239	2829.303	7373.887
14	2629.070	2614.051	2621.560	225.550
15	2786.726	2839.296	2813.011	2763.638
	% Std error = 0.609	R.S.D. = 3.980		

% Main effects

Main effect (Acetonitrile) max value =	0.628
Main effect (Acetonitrile) min value =	0.303
Main effect (Acid type) max value =	3.152
Main effect (Acid type) min value =	3.489
Main effect (Flow-rate) max value =	1.787
Main effect (Flow-rate) min value =	−0.055
Main effect (Temperature) max value =	1.985
Main effect (Temperature) min value =	0.036
Main effect (Wavelength) max value =	0.984
Main effect (Wavelength) min value =	0.328
Main effect (Resp. time) max value =	1.349
Main effect (Resp. time) min value =	0.386

TABLE VII

SUITABILITY PARAMETERS DERIVED FOR THE ANALYSIS OF ASPIRIN AND SALICYLIC ACID

	Aspirin	*Salicylic acid*
Retention time (s)	200–320	240–320
Number of theoretical plates	2450–2860	2900–3700
Resolution	2.8–6.0	
Peak area response (units/mg)	1.88–26.86	287.386
Peak height response (units/mg)	0.18–1.98	14.6–37.3

that all results must be within these ranges before a given system could be considered suitable for this application.

The effects calculated for the quantitative results depended on calibration solutions injected under identical conditions, hence they did not account for drifts in chromatographic conditions. It was therefore important to observe effects of variables which could shift peak areas and heights (*e.g.*, the temperature could drift, if the method operates at ambient temperature and the solvent composition of premixed solvents can change with evaporation). The analyst can then predict the frequency of calibration injections and the maximum number of injections needed in a set of sample concentration calculations. For instance, this ruggedness test revealed that a 1% reduction in the acetonitrile composition caused an average decrease in the peak height of aspirin of 4.165%. The method specification should therefore caution the user to minimize solvent evaporation. It is also recommended that no more than ten injections should be included in the calculation of sample concentrations by peak height (*e.g.* calibration × 2, sample × 6, calibration × 2), due to this effect.

CONCLUSIONS

The proposed method for ruggedness testing successfully revealed the effects caused by changing chromatographic conditions for the HPLC analysis of aspirin and salicylic acid. The tested method was shown to be rugged with respect to the chromatographic variables evaluated. The value of employing a ruggedness test as part of a method validation programme was clearly demonstrated by the increased understanding of the method which was achieved. The automation of the test makes it simple to use and suitable for routine use.

ACKNOWLEDGEMENT

The authors wish to thank D. R. Stout, Philips Analytical, for writing the software necessary to automate the proposed ruggedness test.

REFERENCES

1 G. E. P. Box, W. G. Hunter and J. Stuart Hunter, *Statistics for Experimenters, An Introduction to Design, Data Analysis and Model Building*, Wiley, New York, Chichester, Brisbane, Toronto, 1978, Part III, Ch. 9–13, pp. 291–453.
2 R. L. Plackett and J. P. Burman, *Biometrika*, 33 (1946) 305.
3 M. Otto and W. Wegscheider, *J. Chromatogr.*, 258 (1983) 11.
4 W. Lindberg, E. Johansson and K. Johansson, *J. Chromatogr.*, 211 (1981) 201.
5 J. S. Kiel, S. L. Morgan and R. K. Abramson, *J. Chromatogr.*, 320 (1985) 313.
6 J. Hudec, M. Polievka and J. Balak, *Petrochemia*, 22(1) (1982) 12.
7 B. S. Aswathan and V. J. Victor, *Indian J Technol.*, 12 (1974) 367.
8 R. C. Sarkar and M. Das, *J. Indian Chem. Soc.*, 57 (1980) 787.
9 A. Caputi and P. A. Stafford, *J. Assoc. Off. Anal. Chem.*, 60(5) (1977) 1044.
10 V. Y. Taguchi, M. L. Cotton, C. H. Yates and J. F. Millar, *J. Pharm. Sci.*, 70(1) (1981) 64.
11 P. V. Mroso, A. Li Wan Po and W. J. Irwin, *J. Pharm. Sci.*, 71(10) (1982) 1096.
12 L. Y. Lo and A. Bye, *J. Chromatogr.*, 181 (1980) 473.
13 K. J. Williams, A. Li Wan Po and W. J. Irwin, *J. Chromatogr.*, 194 (1980) 217.
14 J. Fogel, P. Epstein and P. Chen, *J. Chromatogr.*, 317 (1984) 507.