ELSEVIER

Journal of Chromatography A, 870 (2000) 45–51

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Full and fractionated experimental designs for robustness testing in the high-performance liquid chromatographic analysis of codeine phosphate, pseudoephedrine hydrochloride and chlorpheniramine maleate in a pharmaceutical preparation

Ruby Ragonese, Mary Mulholland*, John Kalman

Department of Chemistry, Materials and Forensic Science, University of Technology – Sydney, P.O. Box 123, Broadway, Sydney, NSW 2007, Australia

Abstract

This paper describes the testing of a saturated factorial design using a full factorial design. Saturated factorial designs are often used to test the robustness of high-performance liquid chromatography (HPLC) methods, however they are based on several assumptions. A full factorial design relies on fewer assumptions and hence could be used to evaluate the effectiveness of the saturated design. Both designs were used to test a gradient HPLC method for the assay of codeine phosphate, pseudoephedrine hydrochloride and chorpheniramine maleate. Six HPLC conditions, including wavelength, mobile phase pH and ion pairing reagent concentration were tested using the saturated design. Three of these factors were selected for full evaluation using a full factorial design. The results showed that the main effects calculated by each design were comparable. However, the saturated design showed higher standard errors, probably due to the effects of changing several more factors. One interaction effect was indicated as a confounding effect by the saturated design and this was confirmed by the calculation of the same interaction effect using the full design. Overall the method was shown to be robust under the variety of HPLC conditions tested.

Keywords: Experimental design; Robustness testing; Pharmaceutical analysis; Factorial designs; Validation; Codeine; Pseudoephedrine; Chlorpheniramine

1. Introduction

This paper describes an investigation into two statistical methods for the determination of the robustness of an analytical method as part of an overall method validation strategy. Robustness testing is carried out as part of a precision study and the goal is to establish the effect of small changes in the method conditions (such as temperature or instrumental settings) on the qualitative and quantitative abilities of the method [1–3].

The precision is a measure of the random bias of the method. It has contributions from the repeatability of various steps in the analytical method, such as sample preparation and sample injection for high-performance liquid chromatography (HPLC) [4–8], and from reproducibility of the whole analytical method from analyst to analyst, from instrument to instrument and from laboratory to laboratory. Youden et al. set down guidelines for the validation of analytical methods in their book [8,9] and they

*Corresponding author.

0021-9673/00/$ – see front matter
PII: S0021-9673(99)00972-3

specified the testing of ruggedness prior to a reproducibility study for this reason.

For many years the recommended designs for ruggedness testing were saturated fractional factorial designs [8–14]. However, other designs – full, fractional and saturated factorials together with central composite, Box–Behnken and star designs could provide more thorough solutions for some applications. This paper investigates the use of a full factorial design together with a saturated design for a HPLC gradient method.

Plackett–Burman designs for the testing of seven factors are the most commonly used designs for ruggedness testing of HPLC methods [15]. For a ruggedness test it is essential to determine whether a method is rugged to many changes rather than determine the values of each effect. Although these saturated designs assume interaction effects to be negligible and only estimate main effects; they have the feature known as confounding where higher order effects can overwrite the main effects. Thus if a method is not rugged to higher order effects this will be observed in the values of the main effects.

A full factorial design can estimate all higher order interaction effects so this investigation compared the results of this design with the reduced Placket–Burman design. The number of experiments required to perform a full factorial design increases dramatically with the number of factors. For example, a two-level design for seven factors the full design requires 128 experiments, from which 128 statistics can be measured to estimate the effects shown in Table 1. In terms of absolute magnitude the main effects tend to be higher than two-factor interactions which in turn are higher than three-factor interactions and so on. At some point it is true to say that after a certain order interaction effects become negligible and can thus be disregarded in the experimental design. To do this, full factorial designs are fractionated to allow the estimation of only a certain level of interaction effects. This study attempts to evaluate the correctness of these assumptions by simultaneously evaluating factors using saturated and complete designs. Box et al. [15] provide a good introduction to factorial designs; the most thorough ruggedness test would involve the application of a full factorial design that tests all main effects and interaction effects.

Full factorial designs can be fractionated by the exclusion of experiments designed to identify higher order effects and such reduced designs are known as fractional factorial designs. Saturated designs are constructed on the assumption that all interaction effects can be assumed to be insignificant and the number of experiments is now reduced to $k+1$. These designs are particularly useful for an efficient solution to three level designs as they can be reflected. This type of reflected design evaluates the experimental space in only two of the eight potential hypercubes and is only valid with saturated factorial designs as they depend on the assumption that all interaction effects are negligible. With this assumption the results for any of the remaining six hypercubes can be calculated from the results of the two evaluated diagonal hypercubes.

The main limitation around which a ruggedness test is designed is firstly the number of factors and levels that require testing and secondly the number of experiments needed by a particular experimental design. There is clearly a compromise that needs to be made between the thorough study provided by designs such as full factorials and central composites and the efficiency of the reduced designs such as the star and saturated factorial designs. It is unlikely that any ruggedness test could justify the outlay of more than 30 experiments, in fact it is rare that more than 20 experiments would be carried out.

If a method of analysis is fast or can be fully automated and requires the testing of few factors (three or less) then the larger designs can be considered. Good choices are central composite designs, or if a linear factor response is expected a full factorial

Table 1
Number of effects calculated from a full factorial design for seven factors

Average	Main effects	Interaction effects					
		Two-factor	Three-factor	Four-factor	Five-factor	Six-factor	Seven-factor
1	7	21	35	35	21	7	1

design at two levels. These large designs have the advantage of allowing a complete study where all interaction effects are estimated. However, for the majority of applications the large number of experiments required discourages their use. When a large number of factors need to be tested then it is more efficient to select one of the saturated factorials, while bearing in mind the limitations of these designs. This study attempts to show if the assumptions associated with saturated designs can be made for complex HPLC methodology.

2. Experimental

2.1. Chemicals and reagents

All reagents used were of analytical grade; mobile phase solvents were of HPLC grade. The aqueous mobile phase was prepared by weighing 2.0±0.1 g of octanesulphonic acid sodium salt (PIC-B8 ion-pairing reagent) (Biolab Scientific, Clayton, Australia), adding 750 ml of Milli-Q water to dissolve, followed by 0.5 ml of 85% orthophosphoric acid (Univar, APS Ajax Finechem) and making up to 1000 ml with Milli-Q water. The pH of the solution was adjusted to 3.00±0.05 with 2 *M* NaOH (BDH AnalaR-grade, Merck P/l, Australia) and it was then filtered through a 0.45-μm nylon filter (47 mm, Activon). Acetonitrile (Mallinckrodt ChromAR HPLC grade) was used as the organic phase. The samples and reference standard solutions were prepared with a phosphate buffer (27.6 g sodium dihydrogen orthophosphate monohydrate, BDH AnalaR, Merck P/l) in 1000 ml of Milli-Q water, pH adjusted to 2.70±0.05 with 85% orthophosphoric acid. A combined reference standard solution was accurately prepared to contain the following approximate concentrations: 0.095 mg/ml of codeine phosphate, 0.300 mg/ml of pseudoephedrine·HCl and 0.020 mg/ml of chlorpheniramine maleate. The sample solution was prepared by accurately weighing out an amount from a homogeneous sample, equivalent to the above concentrations of standards. Both sample and standard solutions were filtered through 0.45-μm Millex-HV filter discs (Millipore).

2.2. Instrumentation

The method used quantitatively determines the amount of codeine phosphate, pseudoephedrine·HCl and chlorpheniramine maleate in a pharmaceutical tablet. The HPLC system used was the Waters Alliance 2690 Separations Module with a 996 photodiode array detection system, controlled via Millenium32 software. The 10-μl of sample and standard solutions were respectively injected onto a C_8 bonded reversed-phase HPLC column (Waters Symmetry C_8, 5 μm, 150×3.9 mm with guard column). The column temperature was held at 40±2°C. The actives were separated using ion-pair chromatography and gradient elution of the aqueous–acetonitrile mobile phase. The gradient used was linear from 18% to 40% acetonitrile. Each compound was detected at its maximum wavelength, i.e., codeine phosphate at 210 nm, pseudoephedrine·HCl at 205 nm and chlorpheniramine maleate at 223 nm. The optimum flow-rate was 1.0 ml/min.

Table 2 shows the method conditions for the saturated factorial design experiments and Table 3 shows the conditions for the full factorial design. The limits of the factors studied by the designs were selected according to error ranges which would be typically encountered in an analytical laboratory. For example, in both Tables 2 and 3, the phosphate buffer pH optimum value is 2.70. The extreme limits tested were 2.50 and 2.90 respectively, allowing for the pH meter's accuracy limits of ±0.05 (from instrument literature) plus methodology random errors. Ultimately, the limits tested should not be so wide apart so as to purposely cause the ruggedness test to fail, but should represent the type of variability encountered in the analytical laboratory (in this case, a pharmaceutical one).

3. Results and discussion

3.1. Treatment of results

To fully identify quantitative effects, calibration solutions plus sample solutions need to be analysed for each experiment in a ruggedness test. As duplicate determinations are required for the estimation of standard errors a single experiment consisted of the

Table 2
Reflected saturated factorial design for seven factors at three levels

	Wavelength (nm)	Flow-rate (ml/min)	PIC-B8 mobile phase pH	PIC-B8 concentration (g/l)	PO_4 buffer pH	PO_4 buffer concentration (g/l)	Dummy
1	209	1.0	3.00	1.9	2.70	26.6	+
2	209	0.9	3.00	2.0	2.50	27.6	+
3	209	0.9	2.80	2.0	2.70	26.6	−
4	210	0.9	2.80	1.9	2.70	27.6	+
5	209	1.0	2.80	1.9	2.50	27.6	−
6	210	0.9	3.00	1.9	2.50	26.6	−
7	210	1.0	2.80	2.0	2.50	26.6	+
8	210	1.0	3.00	2.0	2.70	27.6	−
9	211	1.0	3.00	2.1	2.70	28.6	+
10	211	1.1	3.00	2.0	2.90	27.6	+
11	211	1.1	3.20	2.0	2.70	28.6	−
12	210	1.1	3.20	2.1	2.70	27.6	+
13	211	1.0	3.20	2.1	2.90	27.6	−
14	210	1.1	3.00	2.1	2.90	28.6	−
15	210	1.0	3.20	2.0	2.90	28.6	+

four chromatographic experiments as shown below. These were then treated as two duplicate series.

The following data were collected for each experiment: (1) the retention times of sample peaks (t). (2) Sample peak area (a) and sample peak height (h). (3) Concentration in the sample (c). This was calculated using both peak areas and peak heights. (4) Mean number of theoretical plates (N), there are several methods to calculate N, the following calculation was employed due to its convenience as it uses values which are previously collected as part of the data handling.

$$N = (2\pi) \cdot \left[\frac{ht}{a} \right]^2 \quad (1)$$

(5) The resolution between each peak and its nearest eluting peak in the sample chromatogram where R_s was calculated as follows:

$$R_s = \frac{1}{2} \cdot \frac{(t_2 - t_1)}{(t_2 + t_1)} \cdot \sqrt{N} \quad (2)$$

where t_1 and t_2 are the retention times for the two peaks and N is the mean plate number.

The main effects were calculated by adding together all the values of a given parameter obtained at one level and subtracting the sum of the values obtained from the other level and divided by the half the number of experiments. For instance, the calculation of the main effect on plate count N for factor A was carried out as follows:

$$(+N_1 + N_2 + N_3 - N_4 + N_5 - N_6 - N_7 - N_8)/4 \quad (4)$$

where N_1 to N_8 are the number of theoretical plates calculated for experiments 1 to 8. The standard errors (SEs) were calculated as follows:

$$\mathrm{SE} = \sqrt{\frac{4\sum (d_1)^2}{2N_{\mathrm{exp}} 2g}} \quad (5)$$

where d_1 is the difference between duplicate experiments and g is the number of degrees of freedom (and is equivalent in this case to the number of experiments).

In order to standardise the units and numerical size of the main effects (MEs) and standard errors for

Table 3
Full factorial design for three factors at two levels

	PIC-B8 mobile phase pH	PIC-B8 concentration (g/l)	PO_4 buffer pH
1	3.00	2.0	2.70
2	3.20	2.0	2.70
3	3.00	2.1	2.70
4	3.20	2.1	2.70
5	3.00	2.0	2.90
6	3.20	2.0	2.90
7	3.00	2.1	2.90
8	3.20	2.1	2.90

each of the measured parameters, all were recalculated as a % of the values obtained at nominal method conditions (x), hence

$$\%\,\mathrm{ME} = \frac{\mathrm{ME}}{x} \cdot 100 \quad (6)$$

and

$$\%\,\mathrm{SE} = \frac{\mathrm{SE}}{x} \cdot 100 \quad (7)$$

3.2. Comparison of results

Table 4 shows the main effects and standard errors obtained for the mobile phase pH, the PIC-B8 reagent concentration in the mobile phase and the phosphate buffer concentration used in the sample extraction. The saturated design also studied the effect of changing flow-rate, wavelength and the pH of sample extraction. The fact that there were more multivariate changes occurring throughout the saturated design resulted in higher standard errors for every HPLC characteristic. The standard errors are typically four-times larger. This is an important result as the variation of multiple conditions is more likely to reflect the day-to-day usage of the HPLC method. A limited full factorial design would hence underestimate the variation associated with the method.

The error associated with the estimation of concentration using peak area was less than 2.2% for all the components. Codeine and pseudoephedrine showed more variation than the chlorpheniramine, this is due to the fact that these compounds are separated by a smaller resolution (3.7).

The factor with the largest influence was the PIC-B8 concentration for the assay of pseudoephedrine; the main effect on concentration was over 3%, shown in Table 4. This result would require further examination to confirm that it is a real effect and not an interaction effect. Table 6 shows the pattern of confounding effects for the Plackett–Burman design. There are three possible interactions: (i) wavelength/flow-rate; (ii) PIC-B8 phase pH/buffer concentration and (iii) buffer pH/dummy

The interaction between PIC-B8 mobile phase pH/PO_4 buffer pH was measured during the full factorial design and found to be 1.146%, as shown in Table 5. This demonstrates that the Plackett–Burman design can identify first order interaction effects and thus are applicable to robustness validation. The purpose of a robustness test t is to establish the effective performance of the HPLC method through a range of changes in the conditions.

Table 4
Summary of main effects and standard errors for a Plackett–Burman and a full factorial design

Factor	R_t		Area		Height		Conc. area		Conc. height		N		R_s	
	Full	Saturated	Full	Saturated	Full	Saturated	Full	Saturated	Full	Saturated	Full	Saturated	Full	Saturated
Codeine phosphate														
PIC-B8 mobile phase pH	1.712	3.236347	−0.652	7.56	−0.594	5.98	−0.243	1.11	−0.594	1.36	2.130	3.67	−10.087	−15.54
PiC-B8 concentration	1.554	−4.30	1.152	−8.67	0.597	−3.98	0.124	−1.86	0.597	−1.34	1.089	−0.07	−7.217	6.10
PO_4 buffer pH	0.011	0.23	1.557	1.87	1.453	−1.06	0.045	−0.84	1.453	−0.70	−1.349	−4.98	7.863	−20.89
% Se	0.109	0.569	0.407	2.371	0.483	2.161	0.402	2.137	0.485	2.056	1.154	2.174	10.972	17.168
Pseudoephedrine														
PIC-B8 mobile phase pH	0.77	0.77	0.28	1.15	−0.26	0.61	0.78	0.84	0.84	−0.38	−0.95	0.08	−11.48	−16.39
PiC-B8 concentration	0.90	0.90	0.93	−4.51	−0.17	0.38	0.07	−3.85	−0.43	−0.56	−1.13	3.26	−8.11	6.99
PO_4 buffer pH	0.73	0.73	0.92	−1.54	0.31	−0.10	−0.10	−2.52	−0.51	0.57	−1.03	−2.29	7.81	−19.86
% Se	0.85	2.75	0.85	2.30	0.91	0.75	0.97	2.21	0.92	0.75	3.08	2.55	11.74	17.14
Chlorpheniramine maleate														
PIC-B8 mobile phase pH	1.76	0.41	14.24	287.51	8.48	287.51	−3.98	3.13	−0.99	2.33	−6.26	−17.68	−0.55	−11.49
PiC-B8 concentration	1.20	−3.49	17.74	255.00	11.38	255.00	−5.15	−1.60	−1.99	−1.37	−7.75	−7.63	−3.25	−5.14
PO_4 buffer pH	−0.01	0.49	3.57	4.84	1.79	4.84	8.80	−1.16	3.78	−0.35	−3.79	−6.78	−4.00	7.20
% Se	0.09	0.36	2.66	7.54	2.36	4.30	1.78	1.71	1.92	1.66	1.24	1.07	−2.32	9.62

Table 5
The interaction effects obtained from the full factorial design

Factor	ME/R_t	ME/area	ME/Ht	ME/conc. (area)	ME/conc. (Ht)	ME/N	ME/R_s
Codeine phosphate							
PIC-B8 mobile phase pH/PIC-B8 concentration	1.077	−0.526	0.098	−0.779	0.081	2.213	1.504
PIC-B8 mobile phase pH/PO_4 buffer pH	−0.187	−0.164	−0.355	0.410	0.223	−2.154	−0.250
PO_4 buffer pH/PIC-B8 concentration	−0.254	−0.421	−0.106	−0.078	−0.660	−0.860	−3.268
PIC-B8 mobile phase pH/PO_4 buffer pH/PIC-B8 concentration	−0.254	−0.008	−0.385	0.192	−0.396	−2.452	8.546
% Se/1	0.109	0.407	0.483	0.402	0.485	1.154	10.972
Pseudoephedrine							
PIC-B8 mobile phase pH/PIC-B8 concentration	1.119	−0.423	−0.296	−0.184	−0.368	1.435	1.430
PIC-B8 mobile phase pH/PO_4 buffer pH	−0.126	−0.973	−0.797	−1.146	−1.111	−1.447	−0.213
PO_4 buffer pH/PIC-B8 concentration	−0.507	−0.268	0.498	−0.313	0.074	−0.381	−3.392
PIC-B8 mobile phase pH/PO_4 buffer pH/PIC-B8 concentration	0.573	0.185	0.517	0.195	0.953	0.634	9.938
% Se/1	0.851	0.851	0.908	0.974	0.923	3.077	11.739
Chlorpheniramine maleate							
PIC-B8 mobile phase pH/PiIC-B8 concentration	1.127	14.394	9.306	−1.136	0.546	−5.959	−3.139
PIC-B8 mobile phase pH/PO_4 buffer pH	−0.284	0.305	0.528	2.249	1.090	−0.956	−0.960
PO_4 buffer pH/PIC-B8 concentration	−0.134	1.963	1.360	9.034	4.580	−1.508	0.273
PIC-B8 mobile phase pH/PO_4 buffer pH/PIC-B8 concentration	−0.122	1.172	1.026	1.039	−0.291	−1.129	−2.519
% Se/1	0.091	2.665	2.360	1.783	1.922	1.241	−2.317

3.3. Summary of results for the reflected Plackett–Burman design

Table 6 presents the confounding pattern for the Plackett–Burman design. Table 7 shows the main effects for the calculation of concentration by peak area for Codeine Phosphate. The standard errors were between 2 and 4%. The only factors that resulted in a statistically significant main effect were the phosphate buffer concentration and the PIC-B8 concentration, indicating that these are important for the assay of codeine phosphate and need to be carefully defined in the analytical method documentation.

The main effects for pseudoephedrine shown in Table 4 are all within acceptable limits. The larger value obtained for PIC-B8 concentration was explained earlier in the text.

Chlorpheniramine maleate proved to be the least robust measurement made by this method. This was expected as the peak is of low concentration and is retained for the longest on the column. Other validation results presented evidence of this. The largest main effect was obtained for changing the concentration of the phosphate buffer used to extract the sample. This value was 8.8% and thus it will be necessary to closely define these conditions in the analytical method documentation. The changing of

Table 6
The confounding pattern for the Plackett–Burman design

Main effects	Confounding effects		
Wavelength	Flow-rate/PIC concentration	PIC phase pH/dummy	Buffer pH/buffer concentration
Flow rate	Wavelength/PIC concentration	PIC phase pH/buffer pH	Buffer concentration/dummy
PIC phase pH	Wavelength/dummy	Flow-rate/buffer pH	PIC phase concentration/buffer concentration
PIC phase concentration	Wavelength/flow-rate	PIC phase pH/buffer concentration	Buffer pH/dummy
Buffer pH	Wavelength/buffer concentration	Flow rate/PIC phase pH	PIC phase concentration/dummy
Buffer concentration	Wavelength/buffer pH	Flow rate/dummy	PIC phase pH/PIC phase concentration
Dummy	Wavelength/PIC Phase pH	Flow rate/buffer concentration	PIC phase concentration/buffer pH

Table 7
Main effects on concentration (area) for codeine phosphate

Factor	Concentration area/1	Concentration area/2
Dummy	−2.03	4.65
Flow-rate	−0.81	−1.36
PIC-B8 mobile phase pH	1.11	1.19
PIC-B8 concentration	−1.86	−4.28
PO_4 buffer pH	−0.84	−3.91
PO_4 buffer concentration	1.49	−2.80
Dummy	0.72	0.72
Standard error	2.14	3.66

the PIC-B8 concentration also had a significant effect (5.15%) therefore, this factor will need close control as well.

4. Conclusions

The results showed that the main effects calculated by each design were comparable. However, the saturated design showed higher standard errors, probably due to the effects of changing several more factors. One interaction effect was indicated as a confounding effect by the saturated design and this was confirmed by the calculation of the same interaction effect using the full design. Overall, it was shown that the assumptions associated with saturated designs could be made for this particular HPLC method.

In general, the method was shown to be robust with a few modifications required to more closely control the sample preparation.

References

[1] M.J. Cardone, J. Assoc. Off. Anal. Chem. 66 (1983) 1257.
[2] E.L. Inman, J.K. Krischman, P.J. Jimenez, G.D. Winkel, M.L. Persinger, B.S. Rutherford, J. Chromatogr. Sci. 25 (1987) 252.
[3] M. Mulholland, J.A. Van Leeuwen, B. Vandegiste, Anal. Chim. Acta 223 (1989) 183.
[4] M. Thompson, Analyst 113 (1988) 1579.
[5] R. McGill, J.W. Turkey, W.A. Larson, Am. Statistician 32 (1978) 12.
[6] W.J. Dixon, Biometrics 9 (1953) 4.
[7] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, in: Chemometrics – A Textbook, Elsevier, 1988, pp. 93–105.
[8] W.J. Youden, E.H. Steiner, in: Statistical Manual of the AOAC, AOAC, Washington, DC, 1975, p. 33.
[9] W.J. Youden, Mater. Res. Stand. November (1961) 862.
[10] M. Mulholland, in: M.M.W.B. Hendriks, J.H. De Boer, A.K. Smilde (Eds.), Robustness of Analytical Chemical Methods and Pharmaceutical Technological Products, Data Handling in Science and Technology, Elsevier, Amsterdam, 1996, pp. 191–232, Ch. 5.
[11] G. Wernimont, ASTM Standardisation News March (1977) 13.
[12] C.D. Hendrix, Chem Technol. March (1979) 167.
[13] B. Fischer, Anal. Proc. 21 (1984) 443.
[14] R.L. Plackett, J.P. Burman, Biometrika 33 (1946) 305.
[15] G.E.P. Box, W.G. Hunter, J.S. Hunter, in: Statistics for Experimenters – An Introduction to Design, Data Analysis and Model Building, Wiley, New York, 1978, pp. 291–453.