Chemometrics and Intelligent Laboratory Systems, 10 (1991) 337–347
Elsevier Science Publishers B.V., Amsterdam

RES, an expert system for the set-up and interpretation of a ruggedness test in HPLC method validation

Part 1: The ruggedness test in HPLC method validation

J.A. van Leeuwen *, L.M.C. Buydens, B.G.M. Vandeginste *
and G. Kateman

Department of Analytical Chemistry, Catholic University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen (The Netherlands)

P.J. Schoenmakers

Philips Research, Prof. Holstlaan, 5656 AA Eindhoven (The Netherlands)

M. Mulholland

Philips Scientific, York Street, CB1 2PX Cambridge (U.K.)

(Received 21 March 1990; accepted 6 October 1990)

Abstract

Van Leeuwen, J.A., Buydens, L.M.C., Vandeginste, B.G.M., Kateman, G., Schoenmakers, P.J. and Mulholland, M. 1991. RES, an expert system for the set-up and interpretation of a ruggedness test in HPLC method validation. Part 1: The ruggedness test in HPLC method validation. *Chemometrics and Intelligent Laboratory Systems*, 10: 337–347.

In method validation, an intralaboratory repeatability study and an interlaboratory reproducibility study can be performed as part of a precision test. In HPLC, an intralaboratory ruggedness test can be performed to detect problems that would otherwise be encountered in a reproducibility study. In a ruggedness test, variations in ambient factors that are expected to occur in practice, are simulated. Several steps determine the success of a ruggedness test. The complexity and lack of standard procedures for some of these steps is the main reason why ruggedness testing is still not widely accepted.

INTRODUCTION

Considering the widespread use of routine high-performance liquid chromatographic (HPLC) analyses in the analytical laboratory, it is very important that HPLC methods are thoroughly validated. Much time and effort are involved in validating an HPLC method. It is therefore important that any possible problems with a method are detected at an early stage during the method's validation. In particular, interlaboratory repro-

* Present address: Unilever Research Laboratory, Vlaardingen, The Netherlands.

0169-7439/91/$03.50

TABLE 1

Tests in a normal method validation procedure

Method validation	
Specificity:	interference
	peak purity
Accuracy:	recovery
	linearity
Sensitivity:	detection limit
Precision:	repeatability
	ruggedness
	reproducibility

ducibility tests or collaborative studies in the final stage of method validation can be very costly and should only be undertaken if there is a reasonable chance that the method will be accepted [1–3]. Moreover, method validation can frustate the work done in method development and the sooner problems with a method are identified, the easier it is to modify the method because the information from the method development process is still available.

In a typical method validation procedure a number of tests must be performed on the method: e.g. tests to quantify the method's accuracy, sensitivity, specificity, precision etc. (see Table 1). The level of testing depends on the later use of the method. If the method is only to be used occasionally, a relatively limited method validation procedure will suffice. If the method is to be submitted to a regulatory body, extensive testing is required.

Should the method be submitted to a regulatory body, thorough testing of the method, in the laboratory where it was developed, is usually followed by an interlaboratory study. To avoid problems in an interlaboratory test, a method can first be tested in the laboratory to an extent that it is expected to pass interlaboratory testing. If problems occur with the method during an interlaboratory test it is often difficult to trace the cause of the problems because relevant factors are not tested in a controlled way and will vary at random. It is therefore advisable to perform an intralaboratory test on factors that will be tested in an interlaboratory reproducibility study. For that purpose the newly developed method can be submitted to a ruggedness test [4,5].

In the present paper the principles of ruggedness testing in HPLC are outlined and the procedures are described that are part of a typical ruggedness test. In part 2 of this paper an expert system is described that guides a user through a ruggedness test in HPLC method validation [6]. The expert system is based on the theory presented in this part. The purpose of the expert system is to demonstrate the use of expert system technology in HPLC method validation [7].

RUGGEDNESS TESTING

In a ruggedness test one tests the effect of small changes in the operating conditions of a method. The changes reflect the possible changes in circumstances when a method is transferred from one laboratory to another. A ruggedness test can be defined as: "an intralaboratory experimental plan, used before undertaking an interlaboratory study, to examine the behaviour of an analytical process when small changes in the environmental and/or operating conditions are made, akin to those likely to arise in different laboratories" [8].

A ruggedness test is advisable for every analytical method that will be submitted to an interlaboratory study. Also, a ruggedness test can be applied to any method which has been optimised, in order to test whether the optimisation process has not led to an unstable method. In HPLC, a rugedness test is particularly useful, because the number of factors that may affect the performance of the method is very large. Such factors can be found in every part of the HPLC method, ranging from the preparation of the sample to the detection. For instance, an HPLC method can be particularly sensitive to changes in the column (e.g. from one batch to another) or it can be sensitive to small changes in the wavelength of the detector. Applications of the ruggedness test to HPLC methods have been described recently [9–11].

In a ruggedness test, a number of essential steps can be identified (Fig. 1). Roughly speaking, a ruggedness test consists of a pre-experimental phase, an experimental phase and a post-experimental phase. In the pre-experimental phase, the

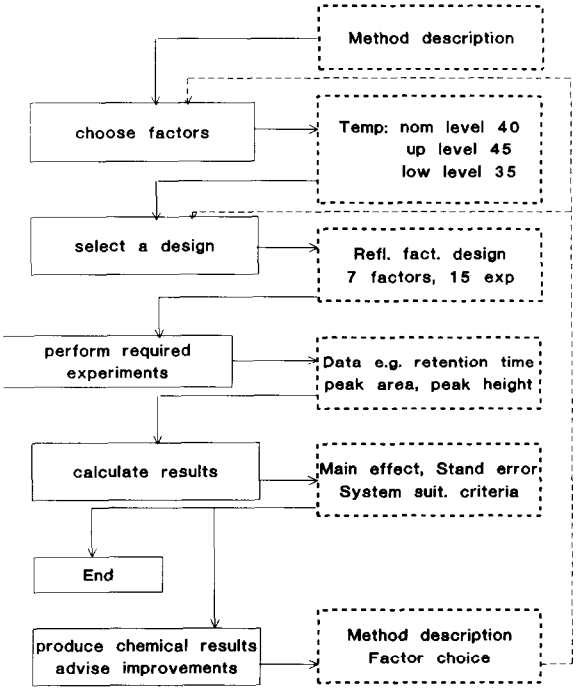

Fig. 1. Steps in ruggedness test set-up and interpretation.

ruggedness test is set up. This phase consists of two important steps: the selection of factors to test, and the selection of an appropriate experimental design. In the experimental phase the experiments are carried out. In the post-experimental phase, the experimental results are interpreted in three steps. The first is to combine information about the factors and the experimental design with the experimental data to evaluate the results statistically. In the next step these results are translated into chemically relevant results. Finally, in some cases, advice can be given on improvements to either the method or the testing procedure.

Each step has its own characteristics and can be seen as a subtask of the ruggedness testing procedure. They differ, for instance, in required knowledge and experience, mode of operation, etc. An outline of each step is given below.

FACTOR CHOICE

Many factors can affect the performance of a method when it is transferred to other laboratories, particularly when the method is complicated. In chromatography, and especially in HPLC the number of factors can be very large (approx. 50, see Table 2). Examples of factors that will undoubtedly vary are temperature, column parameters, the analyst's skill, status of equipment used, etc. Many of these factors will not affect method performance to an extent that the results become unacceptable, because the factor levels vary only moderately in practice. Therefore, in a ruggedness test, only the relevant factors must be tested and testing must take place at relevant levels. For instance, if an operation procedure specifies that a certain step should be carried out at 45 °C, it is useful to test this step at 40 °C and 50 °C but not

TABLE 2

Examples of some possible factors with variation percentages and steps

Sample preparation factors:	
sample weight	1%
shake time	20%
sonicate time	20%
heat temperature	5 °C
wash volume	30%
extraction volume	30%
centrifuge minutes	20%
pore size 1	5 μm
pore size 2	5 μm
extraction 1	
extraction 2	
dilution	
Chromatograph factors:	
pH	1
temperature	5 °C
solvent %	0.1–3%
flow-rate	0.1 ml/min
buffer concentration	1%
additive concentration	0.5%
Column factors:	
manufacturer	
batch	
Detector factors:	
wavelength	5 nm
ri-range	
filter	
time constant	5
Data-handling factors:	
user selected factor	

at 100 °C. Deviations of up to 5 °C may be found in practice but deviations of 50 °C will not, unless great errors are made. Such errors will almost invariably cause obvious malfunctioning of the method.

The level of testing also depends on the purpose of the method. In general, ruggedness testing is performed on methods that will eventually also be used in laboratories other than the one in which they have been developed. Otherwise, the level of testing will be less stringent, usually resulting in the testing of fewer factors.

Choosing the right factors to test is an essential step in setting up a ruggedness test. It largely depends on the expertise and experience of the person selecting the factors whether or not an acceptable set of factors is selected. However, experienced analysts may also overlook important factors. An expert system may rationalise this process, making it consistent and reproducible [12].

SELECTING THE DESIGN

If the factors to be tested have been identified, an experimental design must be chosen on the basis of which the effects of the factors can be tested. The experimental design gives the combinations of factors governing the way in which the experiments are carried out. It is important to keep the number of experiments in a ruggedness test as low as possible. Each HPLC experiment requires a considerable amount of time, especially when conditions are varied after each experiment. In a ruggedness test much time is needed between the experiments for establishing the factors at the specified levels. As a consequence, if a test is extended over a long period of time (e.g. a few weeks), time effects, such as column deterioration, may become important. These factors are then implicitly tested in the test.

Several types of experimental designs can be used in a ruggedness test. However, only a few are applicable. Factorial designs best fit the purpose of ruggedness tests, establishing which of the factors affect method performance and estimating some of the interactions. When using factorial designs a choice must be made between full and (saturated) fractional factorial designs.

TABLE 3

Placket–Burman design to test 7 factors

A Plackett–Burman design for 7 factors at two levels: 0 = factor at nominal level; 1 = factor at extreme level.

Factor	Experiment							
	1	2	3	4	5	6	7	8
1	0	0	0	0	1	1	1	1
2	0	0	1	1	0	0	1	1
3	0	1	0	1	0	1	0	1
4	0	0	1	1	1	1	0	0
5	0	1	0	1	1	0	1	0
6	0	1	1	0	0	1	1	0
7	0	1	1	0	1	0	0	1

The number of factors to test in an HPLC ruggedness test is usually between 3 and 11 [12]. Occasionally, up to 15 factors can be tested [13]. The common type of design for this kind of test is a (saturated) fractional factorial design. With such designs the factors are tested efficiently using a small number of experiments. If, for instance, a so called Plackett–Burmann design is used, seven factors can be tested with only eight experiments [14] (Table 3). These designs assume that all interactions between factors are negligible and provide only the effect of single factors. The extent to which a factor affects the performance of the method is called a main effect. The possibilities for estimating first order interactions in fractional factorial designs are limited because the interac-

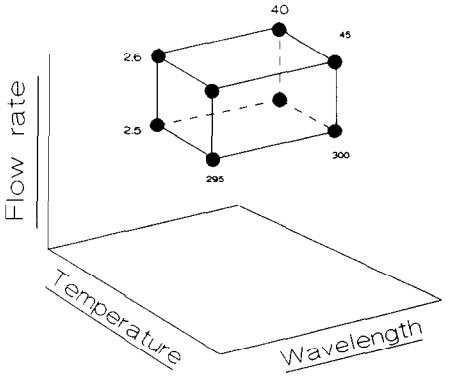

Fig. 2. Factorial design with three factors tested at two levels.

tions are confounded with the main effects. If the number of factors to be tested is low (e.g. 3 or 4), then a full factorial design can be considered (Fig. 2). When using full factorial designs, interaction effects are estimated without increasing the number of experiments to an unacceptable number, provided the number of factors is low.

PERFORMING THE EXPERIMENTS

When the factors have been chosen and a design has been selected, the experimental work can be started. This involves performing the experiments as specified in the design and obtaining the required information from the chromatogram.

In HPLC a number of parameters can be measured (Table 4). The most relevant of these are the concentration calculated from the peak area or from the peak height. These parameters are evaluated to decide on the ruggedness of the method. Other parameters such as the retention time, the resolution or the plate number are also of interest. They indicate whether the method is very sensitive to a certain factor. For instance, a decrease in the plate number indicates that the method deteriorates if it is not performed under the specified conditions. Resolution is a special parameter because, in certain cases, it can cause a method to be non-rugged when the resolution falls below a specified value.

Effects on peak height and peak area should warn the user. If, for instance, a factor affects the peak area or height considerably, the limit of detection may be increased. In general it is important to measure as many parameters as possible. The more information about the method that is collected, the more accurately can possible problems be identified.

TABLE 4

List of possible parameters

1 Concentration calculated with peak area
2 Concentration calculated with peak height
3 Peak area
4 Peak height
5 Retention time
6 Resolution
7 Plate count

CALCULATING THE STATISTICAL RESULTS

After carrying out the experiments, statistical information must be derived from the experimental results. Performing the experiments according to the experimental design leads to a data matrix with four dimensions:

- the number of experiments N_{exp}
- the number of parameters N_{param}
- the number of components in the sample N_{comp}
- the number of duplicates (usually 2)

From this data matrix, the main effect for every component and parameter can be calculated for every factor. Because all measurements are usually performed in duplicate, the standard errors can also be calculated for each component and for each parameter (Table 5).

The calculation of the main effects and the standard errors results in two matrices. The main effect is calculated for each parameter, for each component and for each factor, so that the data matrix of the main effects has three dimensions $[N_{fac}, N_{param}, N_{comp}]$ with N_{fac} being the number of factors tested. The main effects indicate whether a factor affects the response of the method.

The standard errors are calculated for every component, for every parameter and form a two dimensional data matrix $[N_{comp}, N_{param}]$. The standard errors reflect the repeatability of the method at the extreme levels. Normally, a repeatability test under nominal method conditions has already been performed. The repeatability at the extreme levels is investigated because if large errors are found, the main effects become unreliable.

PRODUCING CHEMICALLY RELEVANT RESULTS

The next step is the translation of the statistical results into information on the method. In order to identify significant factors, the main effects and standard errors are compared to critical levels.

TABLE 5

Calculation of main effects and standard errors

Data matrix:

– For every parameter [N_{param}], for every peak [N_{comp}]:

	inj. 1	inj. 2
experiment 1	$x_{1,1}$	$x_{1,2}$
experiment 2	$x_{2,1}$	$x_{2,2}$
experiment 3	$x_{3,1}$	$x_{3,2}$
⋮		
experiment N_{exp}	$x_{N,1}$	$x_{N,2}$

– Experimental design [N_{exp}, N_{fac}]:

	fac *A*	fac *B*	fac ...	fac N_{fac}
experiment 1	–	–	–	–
experiment 2	+	–	–	+
experiment 3	+	+	–	+
⋮				
experiment N_{exp}	–	–	+	+

– Divisor (div): the number of times an effect can be measured in the design

Calculations:

– Main effects for every factor [N_{fac}], for every parameter [N_{param}], for every peak [N_{comp}]:

$$\text{M.E.} = 100 * \frac{\left(-\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \ldots - \bar{x}_{N_{\text{exp}}}\right)}{\bar{x}_1 * \text{div}}$$

$$x_i = \frac{x_{i,\text{inj}1} + x_{i,\text{inj}2}}{2}$$

– Standard errors for every parameter [N_{param}], for every peak [N_{comp}]:

$$\text{S.E.} = 100 * \frac{\sum_{i-1}^{N_{\text{exp}}} \text{Diff}_i^2}{x_1 * N_{\text{dupl}}} \qquad \text{Diff}_i = x_{i,\text{inj}1} - x_{i,\text{inj}2}$$

The critical levels differ depending on the parameter for which the effect or error was measured.

The order in which the errors and effects are interpreted is important. If the method shows a problem with the ruggedness of the concentration, the method fails the ruggedness test and other parameters, such as resolution, become irrelevant.

The standard errors must be checked first (Fig. 3). Standard errors are expressed in percentages of the nominal level. The nominal level is the level at which the method is specified in the operational procedure. In practice it appears that a standard error, found in a ruggedness test in HPLC, of less than 1% is acceptable. If standard errors larger than 1% are found, repeatability is too low. In principle a standard error larger than 1% calls for a diagnosis. However, before diagnosing problems, possible outliers are flagged, having a difference between duplicate measurements larger than 2%. The outlying experiment must be repeated and the standard error is recalculated. If the problem persists, the standard error is listed for diagnosis.

Depending on the size of the standard error,

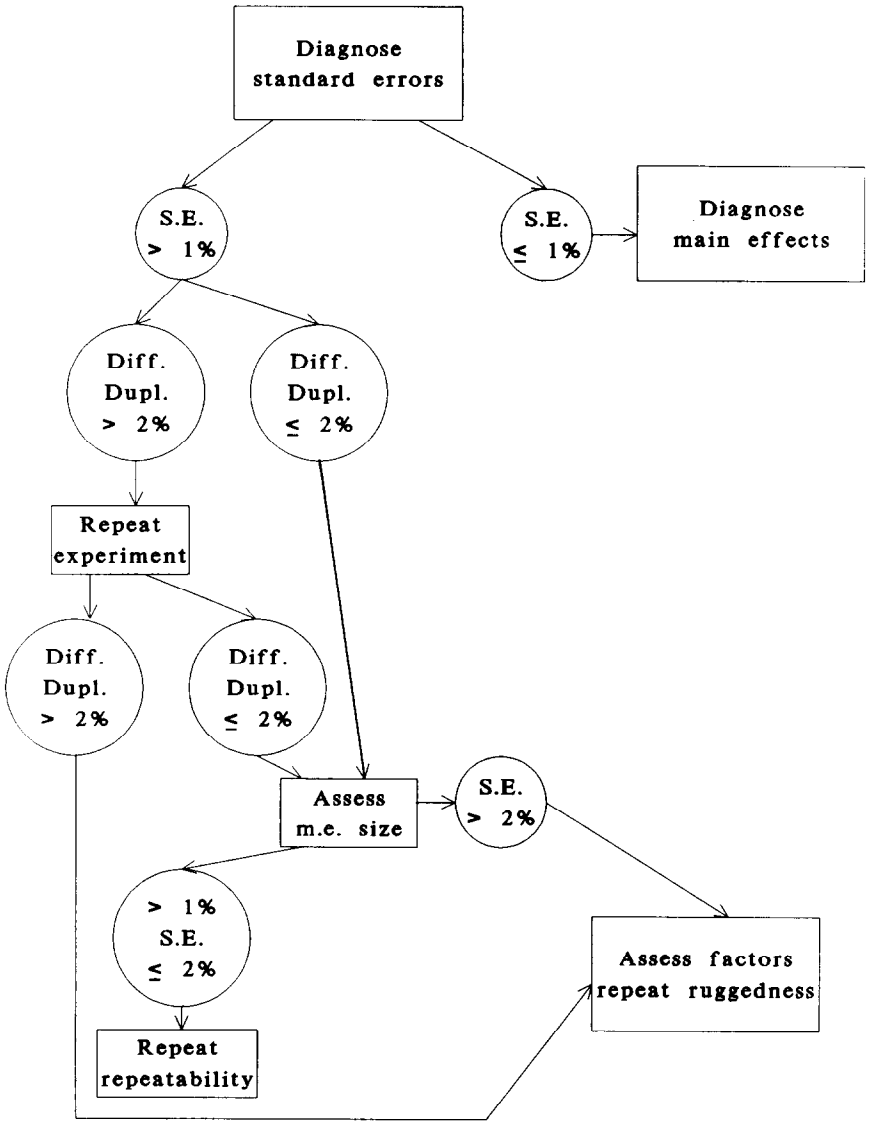

Fig. 3. The diagnosis process for the standard errors.

various actions can be recommended after diagnosis. If the standard error is between 1 and 2%, repeatability is too low. It is then advisable to repeat the repeatability test. If standard errors are larger then 2%, another ruggedness test should follow, the levels of the factors being specified within a narrower range. However, if sample preparation factors are tested, this part of the method must be modified before starting a new ruggedness test. Sample preparation problems are likely to recur if this is omitted. Also, if resolution is a problem, it is suggested that the method should be improved at this point before a new ruggedness test is initiated.

If the standard errors are within the specification, basically the same procedure is applied to the main effects. In statistical terms, a main effect is significant when it is significantly larger than the

standard error for the same component and parameter. This can be tested using an *F*-test. However, comparison of the main effect and the standard error is not enough. If a main effect is smaller than the corresponding standard error, no effect is found for that factor. If the main effect is significantly larger than the standard error a statistically significant main effect exists, but this does not mean that the main effect is relevant from the chromatographic point of view. Therefore, the main effects are also checked against predefined limits. These limits represent acceptable variations in practice. Critical levels for main effects are specific for every parameter (Table 6). For instance, if any main effect found for the concentration is larger than 1%, a problem is revealed that should be subjected to diagnosis. For the other parameters the limits may be less tight. Exceeding these limits is not automatically interpreted as a problem with the method. However, such incidents may be considered in the diagnosis, to see if any warning to the user is appropriate.

All main effects that are subjected to diagnosis can now be related to the factors causing the problems. This list of factors can serve as a basis for the diagnostic process. The factors are identified by the component and the parameter that showed the problem and by the value of the established main effect.

It may be necessary to warn the user that certain factors may affect method performance, without their being critical (Table 7). Such a warning could, for instance, be that changes in temperature may cause a loss of resolution. This is not necessarily critical for the method, but it must be checked regularly.

TABLE 6

List of some critical levels

Parameter	Critical level
Conc. peak area	1%
Conc. peak height	1%
Plate count	50%
Retention time	10%
Peak area	2%
Peak height	2%
Resolution	50% or < 2.5

TABLE 7

List of possible diagnoses

Problem *	Factor	Parameter	Diagnosis
S.E.	all	all	Redevelop sample preparation
M.E.	all	area height	
S.E.	all	all	Improve method (resolution)
M.E.	all	resolution	
S.E.	all	all	Respecify factor levels
M.E.	all	conc. area conc. height	
	drift	height	
	all	resolution	
M.E.	drift	ret. time height	Recalibrate regularly
M.E.	all	p. count	Deterioration in performance
M.E.	all	area height	Reduction of limit of detection
M.E.	all	area height	Peak runs off scale

* M.E. = Main effect; S.E. = standard error.

Main effects on retention time can be due to the effect of drigt in one of the so-called drifting factors, such as temperature or solvent composition. Drift can occur when the laboratory temperature is not controlled or when evaporation of one of the solvent components occurs. The user should be warned, so that the problem can be avoided.

Main effects on the peak area and height can indicate problems in the sample preparation, affecting the recovery. If any factor stemming from sample preparation causes an effect in peak height or area larger than a certain value (default 2%) then the sample preparation is not rugged and must be modified. In addition, main effects on the peak height can indicate if drift factors tend to affect quantification. If this happens, the user should recalibrate regularly. If the effect is too large, the conclusion should be that the method is not rugged.

If no errors or effects are found, system suitability criteria can be calculated. These criteria take the form of upper and lower limits for the resolution, retention time, plate count, etc. A test run of

TABLE 8

System suitability criteria

In a two component system the lowest resolution is found in exp 1. The highest resolution for that component pair is found in exp 5. The system suitability criteria then are, for instance, as given below.

	Experiment 1		Experiment 5	
	comp 1	comp 2	comp 1	comp 2
Rs	2.8	4.5	4.0	6.5
Rt	1.5	3.6	2.5	5.9
N	2000	4000	2500	5600
⋮				

the method, which can be performed at the start of every day, must give results within these limits (Table 8). The system suitability criteria can be found from the extreme values for the parameters obtained during the ruggedness test. First the lowest resolution observed in any experiment is found. This is matched with the maximum resolution for the same component pair. This gives two experiments: one giving the lowest resolution overall and one giving the maximum resolution for the same component pair showing the lowest overall resolution. The results of these two experiments are listed as system suitability criteria.

IMPROVING THE METHOD OR THE TESTING PROCEDURE

Interpretation of the main effects can lead to suggestions for changes in the HPLC method. Because any change in the method will cause much additional experimental work, such as the repetition of previous steps of the method validation process, it is advisable to follow a conservative strategy in implementing such changes. An example of this is the solution of problems with the resolution of the method by changing the flow-rate. In many cases such a problem can also be solved by changing the column dimensions, but this may cause much more dramatic effects than just changing the flow-rate, thus necessitating a great deal of additional testing.

If a method produces significant main effects or standard errors, it may be possible to remedy the problem with a minimum of changes. In particular, in cases where the resolution is the parameter that is affected, a simple adaption of the method may solve the problem.

Resolution is the only parameter in the ruggedness test that not only causes problems if main effects or standard errors are produced: resolution can also cause problems if its value falls below a certain critical value. For instance, a main effect of 5% for resolution will not be listed as a problem (Table 6). However, if the normal resolution is below 3 (e.g. 2.6) and the required resolution for the method is 2.5 it may occur that the resolution becomes $2.6 - 0.05 * 2.6 = 2.47$ which is smaller than the required resolution of 2.5. In such as case a simple solution can be to increase the initial resolution to (for instance) 2.7 by reducing the flow-rate.

A conservative approach to reducing main effects may be to keep factors which failed the ruggedness test under more rigid control. The respecified factors are then tested anew in a second ruggedness test. This can, for instance, happen if the method temperature is fixed at 45 °C and the method is initially tested at 40 °C and 50 °C. If the main effect for one of the concentration parameters is too large upon such changes, the test can be respecified to test temperatures of 43 °C and 47 °C. The ruggedness test is repeated and the method is now more likely to pass the test. The tighter boundary conditions for the temperature should be included in the description of the method, stating that special attention should be paid to controlling the temperature. The tighter control of factors may also be of benefit if there is a standard error larger than 2%. The experiment at which this occurs and the factors that are at the extreme level in this experiment must be identified. The test may be redefined with the factors at a different level. If the method still fails the test after respecification of the levels, then the conclusion should be that it is not rugged.

For some of the factors, respecification of the levels is not appropriate. For instance, if a change of column manufacturer leads to a failure of the test, the method should simply be specified with only one possible manufacturer.

DISCUSSION

Expertise in the field of selecting factors, designs and diagnosing the results is required to perform a satisfactory ruggedness test. Usually, all this expertise is not available in the laboratory, so a ruggedness test is scarcely ever performed. However, ruggedness testing can reduce the chances of failure and thus the cost of reproducibility testing significantly. Previous results of ruggedness tests on HPLC methods indicate that critical factors can indeed be identified [9,10].

Computer programs exist that assist in performing some of the steps described here. For instance, a system exists for selecting the factors and the design [12]. A program also exists for solving resolution problems, although it has been developed for a different purpose [15]. However, an integrated program that covers an entire ruggedness test, from factor choice to improvement of the method does not exist. Because many steps in the ruggedness test are specific for a certain method of analysis, no general approach exists, except for the statistical part [4]. For instance, a ruggedness test in gas chromatography requires other factors to be tested and different conclusions to be drawn than a ruggedness tes in liquid chromatography. Even within liquid chromatography large differences can exist between different types of analysis.

The combination of required expertise and complexity of the test procedures make ruggedness testing an ideal application area for expert systems. Expert systems allow to implement procedures that lack good theoretical models, but on which a considerable amount of experience exists. This is a typical situation that is encountered in ruggedness testing in HPLC and probably also in other areas of HPLC [7].

If expert systems are used in combination with more general algorithmic programs (e.g. statistical packages), difficult problems, such as that of ruggedness testing, can be automated. In Part 2 of this paper a program is described that contains this combination of heuristic and algorithmic processes applied to ruggedness testing in HPLC.

ACKNOWLEDGEMENT

Part of this research is supported by the E.C. under Esprit project 1570 ESCA.

REFERENCES

1 W. Horwitz, Evaluation of analytical methods used for regulation of foods and drugs, *Analytical Chemistry*, 54 (1982) 67A–76A.
2 K.W. Boyer, W. Horwitz and K. Albert, Interlaboratory variability in trace element analysis, *Analytical Chemistry*, 57 (1985) 454–459.
3 Guidelines for collaborative study procedure to validate characteristics of a method of analysis, *Journal of the Association of Official Analytical Chemists*, 72 (1989) 694–704.
4 W.J. Youden and E.H. Steiner, *Statistical manual of the Association of Official Analytical Chemists*, A.O.A.C., Washington, DC, 1975.
5 M.J. Cardone, Detection and determination of error in analytical methodology. Part I The Method Verification Program, *Journal of the Association of Official Analytical Chemists*, 66 (1983) 1257–1282.
6 J.A. van Leeuwen, L.M.C. Buydens, B.G.M. Vandeginste, G. Kateman, P.J. Schoenmakers and M. Mulholland, RES, an expert system for the set-up and interpretation of a ruggedness test in HPLC method validation. Part 2: The ruggedness expert system, *Chemometrics and Intelligent Laboratory Systems*, 11 (1991) in press.
7 D. Goulder, T. Blaffert, A. Blokland, L. Buydens, A. Chhabra, A. Cleland, N. Dunand, H. Hindriks, G. Kateman, J.A. van Leeuwen, D.L. Massart, M. Mulholland, G. Musch, P. Naish, A. Peeters, G. Postma, P.J. Schoenmakers, M. Desmet, B.G.M. Vandeginste and J. Vink, Expert systems for chemical analysis (Esprit project 1570), *Chromatographia*, 26 (1988) 237–243.
8 G.T. Wernimont, *Use of Statistics to Develop and Evaluate Analytical Methods*, A.O.A.C., Arlington, VA, 1985.
9 M. Mulholland and J. Waterhouse, Development and evaluation of an automated procedure for the ruggedness testing of chromatographic conditions in high performance liquid chromatography, *Journal of Chromatography*, 395 (1987) 539–551.
10 M. Mulholland and J. Waterhouse, Investigation of the limitations of saturated fractional factorial experimental designs, with confounding effects for an HPLC ruggedness test, *Chromatographia*, 25 (1988) 769–774.
11 M. Mulholland, Ruggedness testing in analytical chemistry, *Trends in Analytical Chemistry*, 7 (1988) 383–389.
12 J.A. van Leeuwen, B.G.M. Vandeginste, G. Kateman, M. Mulholland and A. Cleland, An expertsystem for the choice

of factors for a ruggedness test in liquid chromatography, *Analytica Chimica Acta*, 228 (1990) 145–153.
13 G. Wernimont, Ruggedness evaluation of test procedures, *ASTM Standardization News*, March (1977) 13–16.
14 P.L. Placket and J.P. Burman, The design of optimum multifactorial experiments, *Biometrika*, 33 (1946) 305–325.
15 P.J. Schoenmakers, N. Dunand, A. Cleland, G. Musch and T. Blaffert, An expert system for the optimization of columns, operating conditions and instrumentation for high pressure liquid chromatography, *Chromatographia*, 26 (1988) 37–44.