

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/112337>

Please be advised that this information was generated on 2021-04-06 and may be subject to change.

Expert Systems for Method Development and Validation in HPLC

Mary Mulholland^{1,*}, N. Walker¹, J. A. van Leeuwen, Luitgard Buydens², F. Maris³, H. Hindriks³, and Peter J. Schoenmakers⁴

¹ Philips Scientific, Cambridge CB1 2PQ, UK

² University of Nijmegen, Nijmegen, The Netherlands

³ Organon International, Oss, The Netherlands

⁴ Philips Research Laboratories, NL-5600 JA Eindhoven, The Netherlands

Abstract. ESCA, Expert Systems Applied to Chemical Analysis, started its research in March 1987, with the aim of building prototype expert systems for HPLC method development. Results of this research have been published as the work has progressed. The project is now completed and this paper summarises some of the overall project conclusions. Seven different expert systems have been built which tackle problems throughout the process of method development, four stand-alone systems and three integrated systems. The object of ESCA was to evaluate the applicability of expert system technology to analytical chemistry and not all the systems were built for commercial uses. Many of the systems tackle problems specific to one or more of the partners and thus may not be useful outside this environment. However, the results of the work are still pertinent to analysts wishing to build their own systems. These results are described, however, the emphasis of the paper is on those systems developed for method validation.

Method validation for HPLC is a complex task which requires many characteristics of the method to be tested, e.g. accuracy, precision, etc. The expert systems built within ESCA concern the validation of precision. Two systems were developed for repeatability testing and ruggedness testing. The method validation process can be divided into several discrete stages, these include: (1) The selection of the method feature to test, for instance which factors can influence the ruggedness of a method. (2) The definition of a test procedure, for instance an efficient statistical design. (3) The execution of experiments and the interpretation of results. (4) A diagnosis of any observed problem. This paper describes these two systems in some detail and summarises some of the results obtained from their evaluation. It concludes that expert systems can be useful in solving analytical problems and the integration of several expert systems can provide extremely powerful tools for the analyst.

* To whom correspondence should be addressed: Department of Chemistry, The University of New South Wales, Sydney, NSW 2033, Australia

Key words: expert systems, method development, method validation, HPLC.

ESCA started its research in March 1987, with the aim of building prototype expert systems for HPLC method development and validation. The starting point of the project involved two separate tasks. Firstly, expert system development software was evaluated to enable the project to select the best available packages for our application [1]. Secondly, the chemical application area had to be defined [2, 3]. Initially, four separate expert system prototypes were built using various expert system development software. The first system tackled the problem of selecting a best starting column and mobile phase for the HPLC of central nervous system drugs [4]. The second system uses expertise to recommend the best criterion for the selectivity optimisation [5]. The third system optimises the physical parameters of an HPLC method, e.g. flow rate and column dimensions [6]. The final system tackled the problem of ruggedness testing HPLC methods [7].

As the project progressed it became apparent that although each of the four systems performed well in its application, analysts require a combination of these tasks. Therefore, the expert systems were integrated in such a way that communication between the various stages of method development was possible.

Three possible integrations were proposed, an integration of the stages of method development, an integration of the ruggedness test with system optimisation and an integration of repeatability with system optimisation.

The purpose of this paper is firstly to summarise the objectives of each integration and to describe the results of testing these systems. Each team is preparing publication of the detailed results from the evaluation of the individual systems. However, some conclusions were common throughout the work. It is the aim of this paper to describe these as they reflect both the successes and limitations of current expert system technology.

Integration of the Method Development Stages

The first step in the definition of an integrated system for HPLC method development was to identify areas of knowledge which were missing from the existing prototype expert systems. The existing prototype for column and mobile phase selection only concerned the specific application of central nervous system (CNS) drugs. In order to widen the scope of this system two other applications were added. Prior to the ESCA project the Free University of Brussels had developed an expert system which could define methods for Label Claim Analysis. This involved developing methods for a very wide range of drug formulations to ascertain the correctly labelled dosage content [8]. This system was implemented in the same expert system shell as the CNS drugs prototype and thus they could be combined in an integrated system. A second system was also added which could refine methods taken from the literature. The addition of these two modules considerably expanded the application scope of the first stage in the method development process.

In order to combine the first guess systems with a selectivity optimisation stage it was necessary to build an extra system which optimised the retention range of the sample components.

ESCA Expert Systems Applied to Chemical Analysis

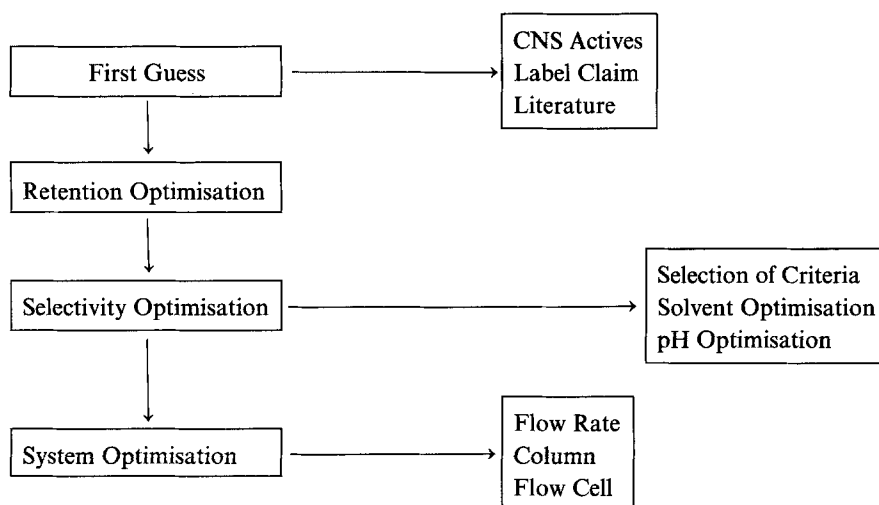


Fig. 1. The integrated system for method development

The selectivity optimisation stage was expanded by including modules for the optimisation of solvent composition and pH. A simplex procedure was used to optimise solvents and the Doehlert design was used to optimise pH.

Finally, the system optimisation prototype could be integrated to define the physical parameters of the method.

These systems were integrated such that communication between the stages was possible via a supervisor system with a common data base [9]. This system is illustrated in Fig. 1.

Integration of the Method Validation Stages

The aim of validating a method is to identify any sources of error in the method and to check whether these errors are within acceptable ranges. However, sometimes a method can show errors which are unacceptable for its application. It is at this stage in the validation process when a link to a re-optimisation module is required. To provide this link the ruggedness test system was integrated to the system optimisation program.

Earlier in the project a small expert system for repeatability testing was built [10]. This system was used as a starting point for a third integrated system which could combine a repeatability test with the system optimisation program.

For both these systems the process of the validation was divided into four stages, which are illustrated in Fig. 2. The first stage required the definition of the characteristic to test. For both repeatability and ruggedness testing this was precision. Next the test was defined as either ruggedness or repeatability and an experimental procedure was recommended. Finally the results are processed in a diagnosis module. This included a pass/fail criterion, a method for interpreting and curing any problems and a means to re-optimize a failed method. The integration of the system

Method Validation

Define Characteristic

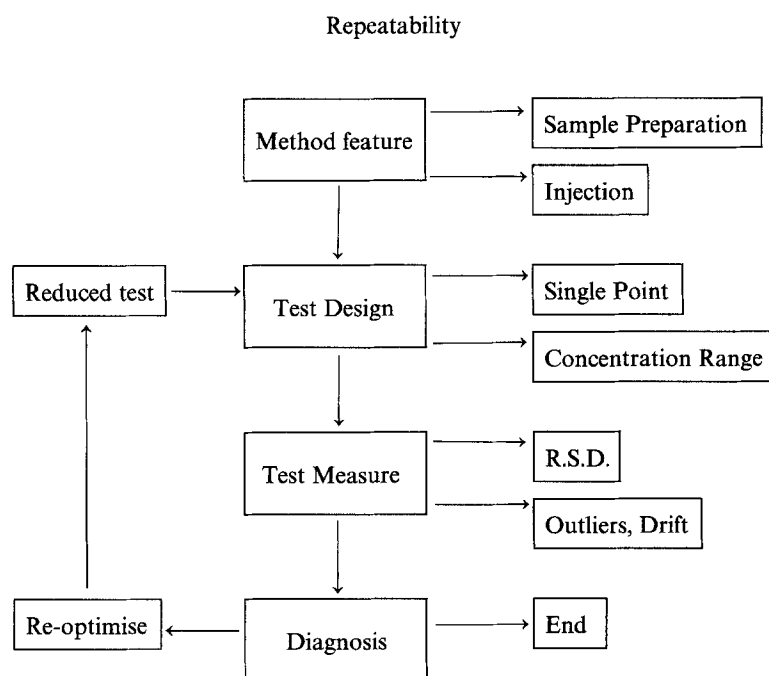
Accuracy
Precision
Sensitivity

Define test

Repeatability
Ruggedness

Experimental Diagnosis

Pass/Fail
Cure Problem
Re-optimize Method

Fig. 2. The four stages of method validation**Fig. 3.** The repeatability text expert system

optimisation program in the diagnosis module allowed the resolution of a critical pair of peaks to be improved.

Fig. 3 illustrates the repeatability test integrated system. The method feature to be tested for repeatability could be the sample preparation or the injection procedure. The experimental design proposed could test the repeatability across a concentration range or at a single concentration point. The measurements made were for the relative standard deviation of concentration, peak heights, retention times and peak areas. These results could then be checked for any outliers or drift. The diagnosis module then could identify any problems and suggest a cure. When the problem was due to inadequate resolution between a pair of peaks, the system

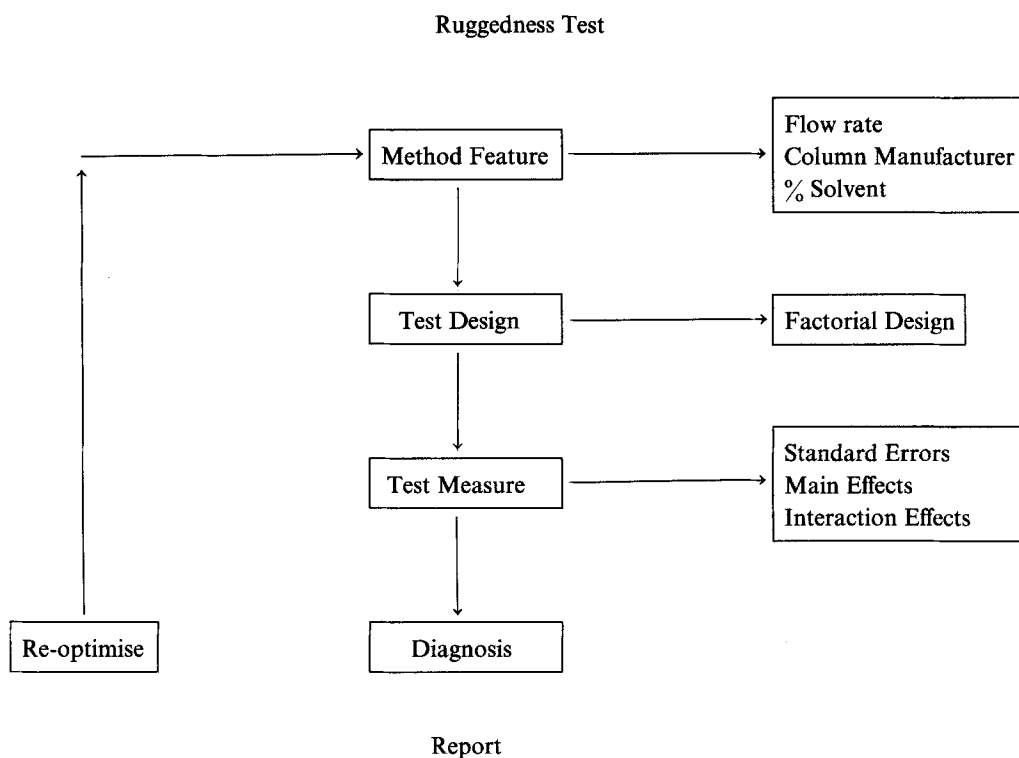


Fig. 4. The ruggedness test expert system

Example Diagnosis	
If column temperature causes a large effect on retention times	Then conclude recommend frequent recalibration
If a factor causes loss of resolution	Then conclude consult system optimisation to increase resolution

Fig. 5. An example diagnosis rule

routed to the re-optimisation module. The new method was then re-tested for repeatability.

The ruggedness test is illustrated in a similar way in Fig. 4. A ruggedness test normally required the testing of between two and ten features of the method. These were chosen by considering their potential influence on the ruggedness of the method. Factorial designs were employed to test the effect of changing these method features, and standard errors, main effects and interaction effects were measured. The diagnosis module then checked these measurements against pass/fail criteria and identified potential problems. When required it routed to the system optimisation module to improve the resolution. An example diagnosis rule is shown in Fig. 5.

Evaluation of the Validation Expert Systems

The expert systems required two stages of testing, validation and evaluation. Validation of the software was performed by the immediate team members and the object

was to test the correct operation of the software and to check that the decisions made by the expert system agree with those of the expert. Evaluation was performed by other team members, who were not directly involved with the development of the software, and by third parties external to ESCA. The objective of the evaluation was to determine how well the expert system performed in practice.

The validation of the software involved checking the system for logical reasoning and testing for bugs. Any inconsistencies or bugs could then be removed at this stage. The expert then selected a number of problems which demonstrated a variety of situations. The expert predicted the answers, and then tested the problem on the software. If the expert and expert system disagreed the reason needed to be identified and a solution found. Knowledge could be incorrect in which case it should be corrected. Knowledge could be missing which then should be added. In this way, the validation process showed whether the systems provided good answers to problems falling within the intended scope of expertise.

In order to validate the ruggedness test expert system, the expert selected ten HPLC methods representing a variety of pharmaceutical applications. The expert selected the factors to test and a statistical design to perform the experiments. These methods were then applied to the expert system and the answers compared. Several modifications and additions were then made to the software until the agreement between consultations was acceptable. The expert and expert system always agreed on the factors to be tested within a difference of two factors, i.e. there were never more than two factors over which the expert and expert system did not agree.

When the software was successfully validated and its performance was considered sufficiently consistent with that of the expert, it could then be given to external evaluators. The evaluation then involved putting problems to the systems in a practical laboratory environment. Two types of evaluators could be distinguished, those who were experts themselves in the application area and those who had HPLC experience but did not have the specific expertise contained in the software. It was important to involve both types of evaluator to assess both the quality and overall usefulness of the advice given by the expert systems.

A list of evaluation criteria was proposed by the knowledge engineers and experts within ESCA. This list is shown in Table 1. The evaluators were then allowed to select relevant criteria from this list to test the software against. The list consisted of three types of criteria, these concerned the user interface, the consistency of advice and the limitations of the system. These criteria allowed the expert evaluators to contribute to identifying the system limits, whereas the non-experts could evaluate the ease of use and consistency of the software.

The time scales of the evaluations varied from a couple of weeks to over a month, so some evaluations were more thorough than others. Due to the fact that the repeatability system was built at such a late stage in the project there was some overlap between the validation and evaluation of this package.

Results of the Evaluation

The Repeatability Expert System

This system was evaluated in two laboratories, at Organon International and Philips Research in Eindhoven.

Table 1. List of evaluation criteria

<i>Man machine interface (user interface)</i>	
	Choice of phrases
	Explanation
	Operation (mouse, keyboard, etc.)
	Ease of use
<i>Consistency testing</i>	
	Accuracy (correct answers, quality of advice)
	Robustness (does the system crash or lock-up)
	Reproducibility (same input, same output)
<i>System limits</i>	
	Conflict (do answers provide conflicting advice)
	Missing knowledge
	Strange or inexplicable advice
	Technical content

The evaluation at Organon was carried out in three separate ways:

1. Method development and validation of new HPLC methods.
2. Repeatability testing of validated methods.
3. Trouble shooting.

For each type of application, different experiments were selected to perform the evaluation and to demonstrate the applicability of this system to the pharmaceutical industry. A full description of this evaluation and its results is currently in preparation for publication. However, a summary of the problems identified and the subsequent action taken is shown in Table 2. This shows the type of problems which were identified, it also demonstrates how the validation of the software had to be combined with the evaluation. Ideally, problems such as incorrect calculations should have been identified and fixed during a separate validation stage.

The second evaluation of the repeatability system applied the software to an HPLC method for the determination of caffeine in coffee. The method was tested at two different injection volumes.

In the process of this evaluation several limitations of the system were highlighted.

Table 2. Summary of some of the problems observed

Problem description	Action
Usage requirement unclear	Text altered
Inaccurate retention times	Calculation corrected
Inaccurate RSD calculation	Calculation corrected
Poor file handling	Load options added
No outlier test	Outlier test added
Drifting baseline not observed	In study

(1) The diagnosis of problems by the system was not completely correct. However, the evaluators understood that this was probably inevitable, because a real expert was also expected to come up with several possible causes for a single observed problem. Hence the criteria used by the evaluators were (a) whether the suggestions were sensible and (b) whether the real cause of the problem was among those suggested. Put in this perspective, the diagnosis module was considered to function correctly.

(2) A design editor which would allow changes to the suggested experiments was requested.

(3) For the system to be used in a practical environment it would need to be linked to a chromatography data station.

(4) The system cannot be used by chromatographers without a certain amount of knowledge. It was not a tool to educate novices.

During the evaluation of this software many of the identified problems were rectified and requested additions were made. This demonstrated the contribution made to the overall value of the software due to the evaluation process.

The Ruggedness Test Expert System

The ruggedness test system was evaluated in several pharmaceutical laboratories. The detailed results will be published in a separate paper. The evaluations were divided into three stages:

1. The input of a variety of methods to the system.
2. A critical evaluation of the factors selected.
3. The collection of data and an evaluation of the diagnosis.

The purpose of applying several different methods to the expert system was to allow the evaluators to find any method features which could not be defined by the system. Also any missing knowledge could be identified. Table 3 summarises some of the results from this stage of testing. The first column lists the method features where knowledge was missing. The second column lists some of the method features which could not be defined in the input module. It was recognised that to some extent this was inevitable as HPLC is very complex and it was difficult to program all the available expertise. However, it was recommended that modules be added to the system at a later date with knowledge on other types of detection. The ruggedness test system did allow the user to use his own knowledge to select factors when necessary. This was invaluable in these cases where knowledge was missing.

Table 3. Results on the input of methods

Missing knowledge	Method features not defined
Fluorescence detection	Dual detection
Ion chromatography	Solid phase extraction
Gradient methods	Both sonicate and shake
Chiral chromatography	Sample solvent composition

The selection of factors module was critically evaluated by comparing the factors selected with the opinions of the evaluator. The following list shows examples of the comments made:

1. The selection of centrifuge time was not necessary.
2. In some instances the levels selected for solvent mix were too extreme.
3. Column batch variation was often preferred to column manufacturer variation.

The selection of factors module was very flexible and any changes in factors or levels required by the user could be made. The analysts used this system in a completely different manner to a conventional software program. They examined the advice and the explanations and accepted or rejected suggestions as they required. This is exactly the same as a typical consultation with an expert and it emphasized the importance of flexibility in expert systems. It was not always possible to overcome conflicts of opinion between evaluators and expert. For instance, some evaluators felt restricted by the allowed statistical designs. However, the expert felt that too much flexibility could lead to errors.

The diagnosis module of the system was also found to have missing knowledge and some differences in opinion were noted. The following list summarises some comments:

1. The main effects were not tested on their statistical significance.
2. The pass/fail levels were sometimes too strict.
3. The handling of internal standard methods was inadequate.
4. The system suitability criteria did not include noise.

To some extent the flexibility of the system could overcome these problems. However, it was clear that additional knowledge is required, particularly for handling internal standards and noise measurements.

Conclusions

The Repeatability Test Expert System

It was clear that there was insufficient time to evaluate the complete system and especially the link from the diagnosis module to the system optimisation module. This part of the system would only be consulted if the diagnosed problem was due to inadequate resolution. This did not occur for any of the test cases. For this reason, it was difficult to conclude on the usefulness of the integrated expert system. Although the evaluators did not use the re-optimisation module, they agreed that it could prove invaluable for the relatively few times it would be required.

Initially, the software had several problems with flexibility and missing knowledge, but these were solved in later updates of the software. This process of evaluation, recommendations and improvements was invaluable to the quality of the final software.

The user interface of this package was considered excellent, with good edit and explain facilities. The quality of advice was good and it could deal with relatively difficult problems. The evaluators concluded that the system was of use to analysts with little software knowledge. However, some experience of HPLC was required and training in the use of the package would be necessary.

The Ruggedness Test Expert System

As for the repeatability system the integrated module for re-optimisation was not fully evaluated. None of the test cases required an improvement in resolution.

The evaluation of this expert system clearly illustrated the difference between an expert system and a conventional software package. Analysts were happy to accept advice or modify it when required. This is exactly how a typical consultation with an expert would be. Generally the expert system was found to be useful, particularly for analysts who had not been involved in the method development process.

There were areas of missing knowledge which should eventually be added to enhance its usefulness.

General Conclusions

As the ESCA research progressed the team built a total of seven different expert systems and it was concluded that these systems provided solutions to many problems in HPLC method development. They were successfully used at each stage of the method development process and an integration of these stages provided useful communication links.

When the evaluation of these packages began many problems were predicted. This was due to the unique nature of expert systems. They do not give right or wrong answers, but good or bad advice. The quality of this advice would be very difficult to measure. Users are expected to interact with expert systems in a completely different way to other software packages. They need to question the system and confirm or modify advice. In some cases they were required to supplement the system with their own knowledge. In fact, we found that analysts welcomed this kind of interaction and were perfectly capable of critically evaluating the advice given. The overall conclusion was that expert systems do not replace analysts as indeed experts cannot directly replace analysts. However, they provide easy access to expertise which can help them work more efficiently.

References

- [1] J. A. van Leeuwen, B. G. M. Vandeginste, G. Postma, G. Kateman, *Chemom. Intel. Lab. Syst.* **1989**, 6, 239.
- [2] D. Goulder, T. Blaffert, A. Blokland, L. Buydens, A. Chabra, A. Cleland, N. Dunand, H. Hindriks, G. Kateman, J. A. van Leeuwen, D. Massart, M. Mulholland, G. Musche, P. Naish, A. Peeters, G. Postma, P. Schoenmakers, M. de Smet, B. Vandeginste, J. Vink, *Chromatographia* **1988**, 26, 37.
- [3] P. J. Schoenmakers, M. Mulholland, *Chromatographia* **1988**, 25, 737.
- [4] H. Hindriks, F. Maris, J. Vink, A. Peeters, M. de Smet, D. L. Massart, L. Buydens, *J. Chromatogr.* **1989**, 485, 255.
- [5] A. Peeters, L. Buydens, D. L. Massart, P. J. Schoenmakers, *Chromatographia* **1988**, 26, 101.
- [6] P. J. Schoenmakers, N. Dunand, A. Cleland, G. Musche, T. Blaffert, *Chromatographia* **1988**, 26, 37.
- [7] M. Mulholland, N. Dunand, A. Cleland, J. A. van Leeuwen, B. G. M. Vandeginste, *J. Chromatogr.* **1989**, 485, 283.
- [8] M. de Smet, A. Peeters, L. Buydens, D. L. Massart, *J. Chromatogr.* **1989**, 485, 283.

- [9] P. Conti, H. Piryns, N. Van den Driessche, M. de Smet, T. Hamoir, F. Maris, H. Hindriks, P. Schoenmakers, D. L. Massart, *Chromatographia* (accepted).
- [10] M. Mulholland, J. A. van Leeuwen, B. G. M. Vandeginste, *Anal. Chim. Acta* **1989**, 223, 183.

Received August 31, 1990. Revision February 12, 1991.