Norms and guidelines for
the practical estimation of uncertainty
in analytical measurements

Programme d'appui scientifique à la normalisation et aux réglementations techniques

Rapport final

N° NM/12/23
NM/03/24
NM/01/25

Politique scientifique fédérale

Jacques CROMMEN, Philippe HUBERT, Patrice CHIAP & Roland Djang'Eing'A MARINI

D. Luc MASSART, Johanna SMEYERS-VERBECKE, Yvan VANDER HEYDEN & Edelgard HUND

Jos HOOGMARTENS, Edwin ROETS, Anne VAN SCHEPDAEL, E. ADAMS & Pieter DEHOUCK

DEPARTMENT OF ANALYTICAL PHARMACEUTICAL CHEMISTRY, INSTITUTE OF PHARMACY
C.H.U., Bât. B36, Avenue de l'Hôpital 1
B-4000 LIEGE 1

DEPARTMENT OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
VRIJE UNIVERSITEIT BRUSSELS
Laarbeeklaan 103
B-1090 BRUSSELS

LABORATORIUM VOOR FARMACEUTISCHE CHEMIE EN ANALYSE VAN GENEESMIDDELEN
Van Evenstraat 4
B-3000 LEUVEN

Jaques DE BEER (WETENSCHAPPELIJK INSTITUUT VOOR VOLKSGEZONDHEID – LOUIS PASTEUR)

Attilio CECCATO (GALEPHAR)

Bruno BOULANGER (LILLY DEVELOPMENT CENTRE, STATISTICAL & MATHEMATICAL SCIENCES)

Liviu MASALAR (SERVICE SYSTÈMES DE PRODUCTION MÉCANIQUE ET MÉTROLOGIE, UNIVERSITE DE LIEGE)

Ilias JIMIDAR (JANSSEN RESEARCH FOUNDATION, METHOD DEVELOPMENT - GLOBAL ANALYTICAL DEVELOPMENT)

Greta VANMARCKE (SOLVAY R & T, DCRT-AN)

Carine VANOETEREN (BAYER ANTWERPEN NV)

E. J. VAN DE VAART (SCIENTIFIC INSTITUTE OF THE DUTCH PHARMACIST)

TABLE OF CONTENTS

RESUME 5
A. CONTEXTE 5
B. OBJECTIFS 5
C. CONCLUSION 5
D. APPORT DU PROJET DANS UN CONTEXTE D'APPUI AUX PROCESSUS DE NORMALISATION ET DE REGLEMENTATIONS TECHNIQUES 7
E. MOTS-CLES 7
SAMENVATTING 8
F. ONDERZOEKSKADER 8
G. DOELSTELLINGEN 8
H. CONCLUSIES 9
I. BIJDRAGE VAN HET PROJECT MET BETREKKING TOT DE ONDERSTEUNING VAN HET STANDAARDISATIEPROCES EN TECHNISCHE REGELGEVINGEN 10
J. SLEUTELWOORDEN 10
SUMMARY 11
K. CONTEXT 11
L. OBJECTIVES 11
M. CONCLUSIONS 11
N. CONTRIBUTION OF THE PROJECT IN A CONTEXT OF SUPPORT TO THE PROCESSES OF STANDARDISATION AND TECHNICAL REGULATIONS 12
O. KEYWORDS 13
1 INTRODUCTION 14
2 METHODOLOGY 15
3 RESULTS 19
3.1 EVALUATION OF THE EXISTING SITUATION 19
3.1.1 The error-budget approach 19
3.1.2 Uncertainty evaluation based on precision assessment 21
3.1.2.1 Introduction and definitions 21
3.1.2.2 Intra-laboratory precision assessment 22
3.1.2.3 Inter-laboratory precision assessment 25
3.1.2.4 Application of precision data in uncertainty evaluation 25
3.1.3 Operational definition of uncertainty 26
3.1.4 Comparison of different methods to evaluate the uncertainty of an HPLC assay 27
3.2 DETERMINATION OF UNCERTAINTY IN ANALYTICAL MEASUREMENTS FROM INTERLABORATORY RESULTS 29
3.2.1 Introduction 29
3.2.2 Set-up of the interlaboratory studies 30
3.2.3 Interlaboratory studies 31
3.2.3.1 Interlaboratory on the analysis of a pen V sample 31
3.2.3.2 Interlaboratory on a LC method for the analysis of erythromycin 37
3.2.3.3 Interlaboratory study related to timolol maleate by LC 41
3.2.3.4 Interlaboratory on a CE method for the analysis of metacyclin 62
3.2.3.5 Interlaboratory on a LC method for the analysis of ketoprofen 66
3.2.3.6 Interlaboratory on a CE method for the analysis of amoxicillin 71
4 DIFFUSION AND VALORISATION 71
4.1 GUIDELINE 71

4.2 DIFFUSION THROUGH INTERNATIONAL CONGRESSES AND PUBLICATIONS 72
4.2.1 International congresses 72
4.2.2 Publications in international scientific journals 73
5 PERSPECTIVES AND OUTCOMES 74
6 REFERENCES 75

RESUME

A. CONTEXTE

Il est essentiel de disposer de méthodes analytiques harmonisées et fiables lors de l'analyse de routine. Le but de celles-ci est d'obtenir des informations de façon à prendre des décisions importantes en ce qui concerne par exemple la conformité aux réglementations gouvernementales. Il est généralement accepté parmi les analystes qu'une interprétation correcte des résultats d'une mesure demande une estimation fiable de leur incertitude. L'incertitude est un paramètre de performance pour les résultats de mesures qui est fondamental si ces résultats doivent pouvoir être comparés et si les utilisateurs veulent pouvoir s'y fier.

B. OBJECTIFS

Le but du projet a été de mettre au point des normes et recommandations pour l'estimation de l'incertitude de mesures analytiques obtenues à partir des méthodes appliquées dans l'industrie chimique, pharmaceutique et apparentée. Il a été principalement question de méthodes de chromatographie liquide ainsi que celles d'électrophorèse capillaire destinées à l'analyse des matières premières médicamenteuses, pour la bioanalyse de médicaments dans le cadre d'études pharmacocinétiques et au test de pureté énantiomérique des médicaments chiraux.
En vue d'atteindre les objectifs du projet, les démarches suivantes ont été appliquées :

- Examen de la situation actuelle et principalement les deux approches différentes employées pour déterminer l'incertitude d'une mesure à savoir l'approche ISO communément appelée « bottom-up » et celle du Comité des Méthodes Analytiques communément appelée « top-down ».
- Etude de certains domaines d'application (études de cas) et de certains problèmes techniques, à savoir l'échantillonnage, le contrôle de qualité, la limite de détection et de quantification dans le but de définir des lignes directrices générales.
- Développement des recommandations générales incluant une approche par arbre décisionnel et proposition de procédures standardisées finalisées, aisées à mettre en œuvre et d'un coût abordable.
- Incitation à la reconnaissance du bien-fondé des recommandations proposées.

C. CONCLUSION

Dans la première partie de ce projet, la collecte des matériels et des situations existantes ainsi que l'étude des thèmes particuliers ont été évaluées. L'approche « bottom-up » (aussi appelée « erro-budget ») pour l'évaluation de l'incertitude a été examinée. Elle est effectuée en quatre étapes à savoir: spécification du mesurande, identification des sources d'incertitudes, quantification des composantes de l'incertitude pour chaque source potentielle d'incertitude identifiée à la deuxième étape et calcul de

l'incertitude combinée. L'évaluation de l'incertitude basée sur l'estimation de la précision a été aussi examinée considérant les études intra-laboratoire aussi bien qu'inter-laboratoires. La seconde partie a consisté en la détermination pratique de l'incertitude d'un résultat analytique individuel obtenu à partir des études inter-laboratoires. Dans cet ordre, de nouvelles méthodes de chromatographie liquide (CL) et d'électrophorèse capillaire (EC) ont été développées, validées et appliquées pour l'analyse des substances médicamenteuses. Trois études inter-laboratoires ont été accomplies. Deux sont complètement terminées et la dernière est en cours. Les méthodes CL et EC développées ont été inclues dans les différents protocoles élaborés pour ces études.

La première étude inter-laboratoires traitant de l'analyse des échantillons de la phénoxyméthylpénicilline (Pen V) consiste en la semi-micro détermination de l'eau par la technique de Karl-Fischer et un test de chromatographie liquide pour la détermination de la 4-hydroxy phénoxyméthylpénicilline ainsi que les autres substances apparentées. Les deux méthodes ont été basées sur la monographie de la Pharmacopée Européenne. La première étude inter-laboratoires consistait aussi à la titration potentiométrique acide-base pour déterminer le contenu de la Pen V. L'étude a montré comment différents estimés de l'incertitude de mesures analytiques peuvent être déterminés et comment les résultats d'une étude inter-laboratoires peuvent être utilisés pour estimer l'incertitude des résultats futures d'un laboratoire analysant des échantillons similaires.

La seconde étude a été réalisée dans le but de valider une nouvelle méthode CL pour l'analyse de l'érythromycine. La reproductibilité de cette méthode a été examinée lors de l'étude inter-laboratoires. Tous les laboratoires ont obtenu des résultats adéquats au niveau de la sélectivité permettant ainsi de déterminer le contenu de l'érythromycine A et de toutes les autres substances apparentées identifiées. L'analyse de la variance (ANOVA), réalisée sur ces résultats, a démontré une bonne reproductibilité de la méthode. De ce fait, elle convient pour remplacer la méthode officielle de la Pharmacopée Européenne. Les résultats de validation peuvent être aussi utilisés par n'importe quel laboratoire qui voudrait effectuer une estimation de l'incertitude pour cette méthode.

Une autre étude inter-laboratoires relative au maléate de timolol a été proposée. Elle comprend la détermination du R-timolol et des autres substances apparentées par chromatographie liquide en phase normale. Une méthode CL pour la détermination du R-timolol, décrite dans la 4ème édition de la Pharmacopée Européenne, a été adaptée, validée et testée au niveau de la robustesse. Les résultats du test de la robustesse ont été utilisés pour l'estimation de l'incertitude.

Quant aux méthodes EC, deux études inter-laboratoires ont été proposées. La première consiste en l'analyse de la métacycline et la seconde au test de pureté énantiomérique du kétoprofène. Pour la première étude, des problèmes sérieux quant au transfert de la méthode en utilisant différents équipements ont été observés. Ces résultats n'ont pas permis une analyse statistique. D'autre part, une autre étude inter-laboratoires basée sur la méthode EC développée pour le kétoprofène a été proposée. Considérant les résultats obtenus avec la première étude, le test de robustesse s'est avéré nécessaire d'être réalisé pour la méthode EC du kétoprofène après avoir validé cette méthode.

D. APPORT DU PROJET DANS UN CONTEXTE D'APPUI AUX PROCESSUS DE NORMALISATION ET DE REGLEMENTATIONS TECHNIQUES

A la fin du projet, les résultats attendus ont été le développement d'une stratégie générale incluant une approche par arbre décisionnel et l'établissement des procédures standardisées. Ces procédures doivent être aisées à mettre en oeuvre et d'un coût abordable. Ainsi, un guide de recommandations a été préparé. Le réseau du projet tentera d'en obtenir l'acceptation dans le but de valoriser les résultats attendus.

Le réseau tentera aussi d'obtenir l'acceptation des recommandations au sein du groupe ISO et leur application dans la commission de la Pharmacopée Européenne.

Le réseau tentera d'initier une nouvelle commission de la "Société Française des Sciences Techniques" (SFSTP) sur l'"Harmonisation des procédures analytiques quantitatives".

Finalement, la dernière activité concernera la définition et l'évaluation d'un guide de recommandations minimales pour la détermination de l'incertitude dans un laboratoire analytique et son évaluation dans les rapports analytiques.

Les collaborations internationales aussi bien avec les industries pharmaceutiques et chimiques que les Universités ont été faites.

E. MOTS-CLES

- Estimation de l'incertitude
- Méthodes de chromatographie liquide
- Méthodes d'électrophorèse capillaire
- Etude inter-laboratoires
- Validation
- Robustesse
- Recommandations.

SAMENVATTING

F. ONDERZOEKSKADER

Voor routineanalysen is het essentieel te beschikken over betrouwbare analytische methoden. Het doel van vele analytische methoden is het verwerven van informatie zodat belangrijke beslissingen genomen kunnen worden, bijvoorbeeld inzake de naleving van overheidsvereisten. Het is algemeen aanvaard onder analisten dat een correcte interpretatie van experimentele resultaten een betrouwbare schatting van hun onzekerheid vereist. Onzekerheid is een primaire karakteristiek voor de kwaliteit van experimentele data die fundamenteel is om te verzekeren dat de resultaten vergeleken kunnen worden en om eindgebruikers toe te laten om op deze data te vertrouwen.

G. DOELSTELLINGEN

Het doel van dit project was het ontwikkelen van normen en richtlijnen voor de praktische schatting van de onzekerheid van analytische metingen verkregen via methoden die gebruikt worden in de chemische, farmaceutische en verwante industrieën. Het betreft voornamelijk vloeistofchromatografische (LC) en capillaire electroforetische (CE) methoden voor de analyse van farmaceutische bulkproducten, voor de bioanalyse van geneesmiddelen in het kader van farmacokinetische studies en voor zuiverheidtesten m.b.t. enantiomeren voor chirale geneesmiddelen.
Teneinde de doelstellingen van dit project te verwezenlijken, gebruikte het netwerk de volgende strategie:

- Eerst en vooral werd de bestaande situatie bestudeerd en in het bijzonder de twee benaderingen die gebruikt worden om de onzekerheid op een meting te bepalen, namelijk de ISO benadering die algemeen gekend is als “bottom-up” en de Analytical Method Committee benadering, ook “top-down approach” genoemd.
- Ten tweede werden enkele toepassingsgebieden onderzocht tezamen met een paar technische problemen zoals staalname, kwaliteitscontrole, detectie- en kwantificatielimieten, teneinde algemene richtlijnen te definiëren.
- Ten derde werden algemene richtlijnen, met inbegrip van een “menu-driven” benadering, en afgewerkte standaardprocedures ontwikkeld. Deze procedures zouden praktisch en rendabel moeten zijn.
- Tenslotte zal het netwerk proberen om de aanvaarding van de voorgestelde documenten te verkrijgen.

H. CONCLUSIES

In dit project werd in het eerste gedeelte het materiaal en de bestaande situaties verzameld, alsook werden de specifieke thema's geëvalueerd. De bottom-up benadering (ook error-budget benadering genoemd) voor het evalueren van onzekerheid werd onderzocht en bestond uit vier stappen: de specificatie van de te meten component, de identificatie van bronnen van onzekerheid, de kwantificatie van de onzekerheidscomponenten voor elk van de in stap 2 geïdentificeerde mogelijke bronnen van onzekerheid en de berekening van de gecombineerde onzekerheid. De onzekerheidsevaluatie gebaseerd op de schatting van de precisie werd tevens bestudeerd via zowel intra- als interlaboratoriumstudies. Het tweede deel betrof de praktische schatting van onzekerheid van een individueel analytisch resultaat verkregen door een interlaboratoriumstudie. In deze context werden enkele nieuwe vloeistofchromatografische (LC) en capillaire electroforetische (CE) methoden voor de analyse van geneesmiddelen ontwikkeld, gevalideerd en toegepast. Drie interlaboratoriumstudies werden uitgevoerd. Twee van de studies werden met succes beëindigd, terwijl de andere bijna afgelopen is. De ontwikkelde LC en CE methoden werden opgenomen in de verschillende protocolen die opgesteld werden voor deze studies.

De eerste interlaboratoriumstudie, omtrent de analyse van een fenoxymethylpenicilline (Pen V) staal bestond uit een Karl-Fischer semi-micro bepaling van water en een vloeistofchromatografische test voor de bepaling van 4-hydroxyphenoxymethylpenicilline en andere gerelateerde onzuiverheden. De twee methoden zijn gebaseerd op de monografie van de Europese Farmacopee. De eerste studie bestond tevens uit een potentiometrische zuur-base titratie voor de gehaltebepaling van Pen V. De studie toonde aan hoe verschillende schattingen van de onzekerheid van analytische metingen bepaald kunnen worden en hoe de resultaten van een interlaboratoriumstudie gebruikt kunnen worden voor de schatting van de onzekerheid op toekomstige resultaten bepaald door een individueel laboratorium voor analoge stalen.

De tweede interlaboratoriumstudie werd uitgevoerd om een nieuwe LC methode te valideren voor de analyse van erythromycine. De reproduceerbaarheid van deze methode werd bestudeerd in een interlaboratoriumstudie. Alle labo's verkregen een adequate selectiviteit om de gehaltebepaling van erythromycin A en alle geïdentificeerde gerelateerde substanties toe te laten. De variantie-analyse (ANOVA), uitgevoerd op deze resultaten, toonde de goede reproduceerbaarheid van de methode aan. De methode is geschikt om de bestaande officiële methode van de Europese Farmacopee te vervangen. De resultaten van de validatie kunnen tevens gebruikt worden door elk labo dat een onzekerheidsbepaling wil doen voor deze methode.

Een bijkomende interlaboratoriumstudie betreffende de chirale scheiding van timololmaleaat werd voorgesteld. Ze bestond uit de bepaling van R-timolol en andere gerelateerde substanties door middel van vloeistofchromatografie in normaalfase modus. Een LC methode voor de bepaling van R-timolol,

beschreven in de 4de editie van de Europese Farmacopee werd aangepast, gevalideerd en getest op robuustheid. De resultaten van de robuustheidtest werden gebruikt voor de onzekerheidsbepaling.
Wat de CE methoden betreft, werden twee interlaboratoriumstudies voorgesteld. De eerste betreft de analyse van metacycline en de tweede de enantiomeren-zuiverheidstest van ketoprofen. Voor de eerste CE studie werden ernstige methode-transferproblemen vastgesteld bij gebruik van verschillende apparatuur. De resultaten lieten geen statistische analyse toe. Anderzijds werd een tweede interlaboratoriumstudie gebaseerd op CE methodeontwikkeling voor ketoprofen voorgesteld. Rekening houdend met de resultaten verkregen in de eerste CE studie, werd het nodig geacht om een robuustheidtest uit te voeren voor de CE methode op ketoprofen na de validatie van de methode.

I. BIJDRAGE VAN HET PROJECT MET BETREKKING TOT DE ONDERSTEUNING VAN HET STANDAARDISATIEPROCES EN TECHNISCHE REGELGEVINGEN

De verwachte resultaten bij het einde van het project waren de ontwikkeling van een algemene strategie, inclusief een “menu-gedreven” benadering, en de definitie van standaardprocedures, die praktisch en rendabel zouden moeten zijn. Daarom werd een voorstel van richtlijn gemaakt. Het netwerk van het project zal proberen aanvaarding te verkrijgen om de resultaten te valoriseren.

Het netwerk zal tevens proberen om de aanvaarding van de klad richtlijnen in de ISO groep te verkrijgen alsook hun toepassing in de Europese Farmacopeecommissie.

Het netwerk zal proberen een nieuwe commissie samen te stellen van de “Société Française des Sciences Techniques” (SFSTP) omtrent “Harmonization of quantitative analytical procedures”.

Tenslotte bestond de laatste activiteit uit de definitie en bepaling van een set van minimale richtlijnen voor de bepaling van onzekerheid in een analytisch laboratorium en hun rapportering in analytische rapporten.

Er werden internationale samenwerkingen gerealiseerd met de farmaceutische en chemische industrie alsook met universitaire laboratoria.

J. SLEUTELWOORDEN

- Onzekerheidsschatting
- Vloeistofchromatografische methoden
- Capillaire electroforetische methoden
- Interlaboratoriumstudies
- Validatie
- Robuustheid
- Richtlijnen

SUMMARY

K. CONTEXT

In routine analysis, it is essential to dispose of reliable analytical methods. The goal of many analytical methods is to gain information so that important decisions can be taken, for instance, about compliance with governmental regulations. It is generally accepted among analysts that a correct interpretation of measurement results requires a reliable estimation of their uncertainty. Uncertainty is a basic performance characteristic of measurement results that is fundamental in order to assure comparability among results and allow end users to rely on them.

L. OBJECTIVES

The aim of the project has been to develop norms and guidelines for the practical estimation of uncertainty in analytical measurements obtained from methods applied in the chemical, pharmaceutical and related industries. It mainly concerns liquid chromatographic (LC) and capillary electrophoretic (CE) methods for the analysis of bulk pharmaceutical compounds, for the bioanalysis of drugs in the frame of pharmacokinetic studies and for the enantiomeric purity testing of chiral drugs.

In order to achieve the objectives of the project, the network has used the following strategy :

- First, the existing situation and in particular the two approaches used to determine the uncertainty on a measurement has been examined, namely the ISO approach commonly known as “bottom-up” and the Analytical Method Committee approach commonly known as “top-down”.
- Second, some fields of application have been investigated and a few technical problems such as sampling, quality control, limits of detection and quantisation, have been studied in order to define general guidelines.
- Third, the general guidelines including a « menu-driven » approach and finalised standard procedures are being developed. These procedures should be practical and cost-efficient.
- Finally, the network will try to obtain acceptance of the proposed guidelines.

M. CONCLUSIONS

In this project, the collection of the material and the existing situations as well as the study of particular themes was evaluated in the first part. The bottom-up approach (also called error-budget approach) for uncertainty evaluation was examined and consisted in four steps : specification of the measurand, identification of uncertainty sources, quantification of the uncertainty components for each potential source of uncertainty identified in step 2 and calculation of the combined uncertainty. The uncertainty evaluation based on precision assessment was also examined considering intralaboratory as well as interlaboratory studies. The second part concerned the practical estimation of uncertainty of an

individual analytical result obtained from an interlaboratory study. For this purpose, some novel liquid chromatographic (LC) and capillary electrophoretic (CE) methods for the analysis of drugs were developed, validated and applied. Three interlaboratory studies were performed. Two were successfully completed while another one is nearly finished. The developed LC and CE methods were included in the different protocols elaborated for these studies.

The first interlaboratory study dealing with the analysis of a phenoxymethylpenicillin (Pen V) sample consisted of a Karl-Fischer semi-micro determination of water and a liquid chromatography test for the determination of 4-hydroxyphenoxymethylpenicillin and other related impurities. The two methods were based on the European Pharmacopoeia monograph. The first interlaboratory study consisted also of a potentiometric acid-base titration to assay the content of Pen V. The study showed how different uncertainty estimates of analytical measurements can be determined and how the results of an interlaboratory study can be used to estimate the uncertainty on future results by a single lab analysing similar samples.

The second interlaboratory study was performed in order to validate a new LC method for the analysis of erythromycin. The reproducibility of this method was examined in an interlaboratory study. All labs obtained adequate selectivity allowing the content determination of erythromycin A and all identified related substances. The analysis of variance (ANOVA), performed on these results, demonstrated the good reproducibility of the method. The method is suitable to replace the existing official method of the European Pharmacopoeia. The results of the validation can also be used by any lab that would like to make an uncertainty statement for this method.

Another interlaboratory study related to timolol maleate was proposed. It involved the determination of R-timolol and other related substances by liquid chromatography in the normal phase mode. A LC method for the determination of R-timolol, described in the 4th edition of the European Pharmacopoeia, was adapted, validated and tested for the robustness. The results of the robustness testing were used to assess uncertainty.

As for CE methods, two interlaboratory studies were proposed. The first concerns the analysis of metacycline and the second the enantiomeric purity testing of ketoprofen. For the first CE study, serious problems of method transfer using different equipments were observed. These results did not allow a statistical analysis. On the other hand, another interlaboratory study based on the CE method developed for ketoprofen was proposed. Considering the results obtained for the first CE study, it was found necessary to perform robustness testing for the CE method on ketoprofen after having validated this method.

N. CONTRIBUTION OF THE PROJECT IN A CONTEXT OF SUPPORT TO THE PROCESSES OF STANDARDISATION AND TECHNICAL REGULATIONS

At the end of this project, the expected results were to develop a general strategy including a “menu-driven” approach and to establish standard procedures. These procedures should be practical

and cost-efficient. Therefore, a set of draft guidelines have been prepared. The network of the project will try to obtain their acceptance in order to valorise the expected results.

The network will try also to obtain the acceptance of the draft guidelines in the ISO group and its application in the European Pharmacopoeia commission.

The network will try to initiate a new commission of the “Société Française des Sciences Techniques” (SFSTP) on “Harmonization of quantitative analytical procedures”.

Finally, the last activity concerned the definition and assessment of a set of minimal guidelines for the determination of uncertainty in an analytical laboratory and its assessment in analytical reports.

International collaborations with pharmaceutical and chemical industries as well as University laboratories were made.

O. KEYWORDS

- Uncertainty assessment
- Liquid chromatography methods
- Capillary electrophoretic methods
- Interlaboratory studies
- Validation
- Robustness
- Guidelines

1 INTRODUCTION

The aim of this project was to develop norms and guidelines for the practical estimation of uncertainty of analytical measurements obtained from methods applied in the chemicals, pharmaceuticals, agro-food and related industries. In routine analysis it is essential to have harmonised, reliable analytical methods. In this context, validation of methods is an important regulatory problem for several industries and it is necessary to demonstrate the validity of methods that are applied to assess the conformance of the products to specifications. The goal of many analytical methods is to gain information so that important decisions can be taken, for instance, about compliance with governmental regulations. However, at the same time there is a need for acceptable norms to define the uncertainty in an analytical measurement. It is also important that the analysis results are reliable to verify if products are within the required specifications, which is a topic that concerns everybody who is involved in a consumer-manufacturer relationship. This is a problem both for the industry and for the national and international authorities. Reliability of analytical results requires on the one hand that methods are validated, while on the other hand, the traceability and uncertainty are two basic performance characteristics of measurement results that are fundamental if comparability among results is to be assured and end users are to be able to rely on them. The importance of traceability is recognised world-wide; the concept and the need for its practical implementation are included in quality assurance standards such as the European Norms (EN) 45001 or the ISO Guide 25. However, although the concept of uncertainty is well established, its practical estimation in chemical analysis presents several problems. Different international committees, such as the ISO/CASCO working group 10, are at present discussing the extent and clarification of the determination of uncertainty in testing laboratories so that a revised definition can be included in the new ISO/DIS 17025 standard, the successor to ISO guide 25. Two approaches for calculating uncertainty in quantitative analysis are defined : the ISO approach also called "*bottom-up*" and the one presented by the Analytical Methods Committee commonly known as "*top-down*" [2]. The ISO approach was originally proposed for quantifying uncertainty in physical measurements. It is based on identifying, quantifying and combining all sources of uncertainty on the measurement. Although the ISO approach improves knowledge of measurement procedure, the difficulty of applying it in chemical measurements hampers its widespread use.

Thus, this project will be mainly focused on the development of norms and guidelines for the estimation of measurement uncertainty in pharmaceutical and biomedical laboratories. It will concern capillary electrophoretic (CE) and liquid chromatographic (LC) methods to analyse bulk pharmaceutical compounds, for the bioanalysis of drugs for pharmacokinetic purposes, for the chiral analysis of drugs.

However, according to the new ISO/IEC F DIS 17025 document [43], testing laboratories such as those involved in the present project or those of the industrial Belgian Science Policy project partners should have to apply procedures for estimating uncertainty of measurement. In addition, in certain cases

the nature of the test method may preclude rigorous, metrologically and statistically valid, calculation of uncertainty of measurement. In these cases the laboratory should at least attempt to identify all the components of uncertainty and make a reasonable estimation and ensure the way of reporting the result does not give a wrong impression of the uncertainty. Reasonable estimation should be based on knowledge of the performance of the method and on the measurement scope and should make use of, for example, previous experience and validation. In this context, the main objective of this project is to determine how to make a statement about the uncertainty of an individual chemical analysis result. The term "chemical" should be interpreted in a large context, which means that also results from the pharmaceutical and agro/nutrition industry will be considered. The measurements could also have a physical background (e.g. measuring a colour) instead of a chemical one.

It is clear that standardisation in this matter is necessary, not only on the national but even more on the international level and therefore our aim is to define standards to be used throughout Europe or even world-wide. In other words there is a need to have agreements (i.e. norms and guidelines) about the way in which practical estimation of uncertainty should be established.

Seen from an organisational point of view the aim of this project is to bring together the knowledge of Belgian researchers working in the field of method validation and related subjects in order to allow the Belgian researchers and industries to play a prominent role in the development of norms and guidelines.

2 METHODOLOGY

In order to achieve the objectives of the project, the network intends to use the following strategy:

- First, the existing situation and in particular the two different approaches used to determine the uncertainty on a measurement will be examined, namely the ISO approach (bottom-up) and the Analytical Method Committee approach (top-down) [2]. However, these guidelines are difficult to apply in practice, which prevents their widespread use. Therefore the existing proposal will be studied more in detail.
- Second, some fields of application will be investigated and a few technical problems such as sampling, matrix effects, quality control, limits of detection and quantisation will be studied in order to define general guidelines, since different amounts of method validation data can be obtained from which conclusions about the uncertainty should be drawn, before going to generalisation.
- Third, the general guidelines including a "menu-driven" approach (to reduce the skill level required in the implementation of the ISO principles) and finalised standard procedures will be developed. These procedures should be practical and cost-efficient which is not always the case for the existing ones. For this reason, the network will try to define a set of minimal

guidelines. These general guidelines will also have to be supplemented by more specific advice for the different specific fields. In the additional guidelines the principles outlined in the general guidelines will be followed to ensure harmonisation between the different fields.

- Finally, the network will try to obtain acceptance of the proposed guidelines.

The correct interpretation of a measurement result requires knowledge about the uncertainty of the result. Although important, the determination of uncertainty is not evident. Several approaches for estimating uncertainty in analytical measurements are proposed. Two of them are described above (bottom-up and top-down approaches) and the approach where uncertainty is measured by using information from the validation process. In the ISO approach all sources of uncertainty are estimated individually. The ISO approach was originally proposed for quantifying uncertainty in physical measurements and was subsequently adapted by EURACHEM for chemical measurements. However the practical application of the ISO approach in chemical measurements is cumbersome and this prevents its widespread use. The "top-down" approach uses interlaboratory studies to measure uncertainty. In this approach the lab is seen from a higher level and thus both systematic and random errors within one laboratory become random errors when they are considered from this higher level. Uncertainty can also be derived from validation data. During the validation of an analytical method, the trueness and precision are tested. The results from this validation process might be used to make an uncertainty statement. However the relationship between validation and uncertainty statement is not clear for many analysts. There is a need to clarify this relationship and to show how the existing data can be applied in practice.

Depending on the situation under which the analyst is validating, different operational definitions of uncertainty have been proposed. Within-laboratory uncertainty, as proposed by Hund et al. [34], only considers the intermediate precision and contains the repeatability and the between-run effect. If an analysis is performed by different laboratories, a between-laboratory effect will exist. The combination of the between-laboratory effect and the repeatability then determines the reproducibility uncertainty on an analytical measurement. When, moreover, the method bias is taken into account as source of uncertainty, bias-included uncertainty is obtained.

It is generally accepted among analysts that a correct interpretation of measurement results requires a reliable estimation of their uncertainty [2, 16]. Influenced by the trend towards more international standardisation, the Comité International des Poids et Mesures (CIPM) in 1977 initiated a broader discussion on the assessment and expression of uncertainty in measurement. After an inquiry among national metrology laboratories, the Bureau International des Poids et Mesures (BIPM) gathered a group of experts to establish a "uniform and generally acceptable procedure for the specification of uncertainty" [27]. Although the International Union of Pure and Applied Chemistry (IUPAC) and the International Federation of Clinical Chemistry (IFCC) participate in the ISO Technical Advisory Group on Metrology, they were not involved in the working group that finally developed the guidance document. In fact, the working group was only composed of members of BIPM, IEC (International Electrotechnical

Commission), ISO and OIML (International Organisation for Legal Metrology). Thus, the group that established the “Guide to the expression of uncertainty in measurement” (GUM) in 1993 only involved experts from the metrological field, but not from chemistry [27]. The GUM proposes an error-propagation or error-budget approach to estimate the uncertainty related to a measurement result. The analytical process is split into its components. One of the least complex examples for a chemical measurement process is the standardisation of a solution against a titrimetric standard, e.g. NaOH against a potassium hydrogen phthalate (KHP) solution. The uncertainty in the concentration of the NaOH standard is split into the uncertainties related to the mass of the titrimetric standard (solid), to the purity of the titrimetric standard, to the molar mass of KHP and to the titration volume of the NaOH [17]. The uncertainties related to these components are separately quantified in the form of a standard deviation and then combined in an error-budget. However, as these uncertainty components are influenced by various parameters, they can be split further. For more complex situations, so-called cause and effect diagrams facilitate the split-up [17, 39]. Due to their shape, they are also referred to as fishbone diagrams.

It is generally accepted that physical measurements and chemical measurements have entirely different error patterns that behave different on replication [32]. According to ISO [45], a random error is a component of the error, which in the course of a number of analyses for the same sample varies in an unpredictable way. The systematic error is the component of error, which in a comparable situation either remains constant or varies in a predictable way [45]. While systematic errors predominate in physical measurements, analytical chemical measurements are more affected by random errors [32]. The error related to a particular analytical measurement x can be decomposed as follows:

+ method bias
+ laboratory bias
+ run effect
x = true value + measurement error

Dependent on the rung chosen as the viewpoint, errors are considered as systematic error or random. For instance, for a particular laboratory, the laboratory bias is a systematic error, but as a member of a set of biases from various laboratories, it is a random error.

Some chemists adopted the error-budget approach [15, 17], but only few go so far as to propose it as a replacement for method validation [50]. In fact, the error-budget approach is not generally accepted as the most suitable way to evaluate the uncertainty related to a measurement result in analytical chemistry [32, 49]. As an alternative method to measure the uncertainty related to a chemical measurement result, the Analytical Methods Committee of the Royal Society of Chemistry (AMC) proposed the top-down approach, which is based on precision data assessed in an inter-laboratory study [2]. Further approaches based on validation data have been proposed as well. The Nordic committee for food analysis for instance proposed to base the uncertainty evaluation on precision data acquired within

a single laboratory [56]. Both approaches relate the uncertainty estimation with the precision evaluation. A further element of method validation that gained importance in uncertainty estimation is trueness evaluation, mostly performed with recovery experiments [5, 29, 54]. A workshop organised by FAO (Food and Agriculture Organisation), IAEA (International Atomic Energy Agency), AOAC int. (Association of Official Analytical Chemists) and IUPAC also stressed the advantages related to a reconsideration of validation data in uncertainty evaluation. The report on this workshop does not only refer to inter-laboratory studies, but also mentions robustness tests as valuable sources of information about the uncertainty of measurement results [61]. A more detailed proposal for the consideration of robustness data in uncertainty evaluation is provided by Barwick and Ellison [5].

A broad discussion about the two different main approaches took place over the last five to seven years. Some of the arguments will also be addressed in the following. Besides the already mentioned difference in measurement structure of physical and chemical measurements, the availability of certified reference materials (CRMs) is also very different for both types of measurement [11]. While for about 90% of physical measurement procedures, CRMs are available, only about 8-10% of chemical measurement procedures can be performed with CRMs [66]. A further point of discussion is the requirement of traceability to the SI (système international) by ISO/IEC 17025 [46]. Traceability is formally defined as "*the property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties*" [45]. Traceability plays an important role for the comparability of measurement results and the uncertainties related to them.

Attempts for reconciliation between the different uncertainty approaches can be observed during the last years [30, 71]. A cornerstone is for instance the International standard ISO/IEC 17025 "General requirements for the competence of testing and calibration laboratories" [46]. In the section about uncertainty of measurement results, it states that "*reasonable estimation shall be based on knowledge of the performance of the method and on the measurement scope and shall make use of, for example, previous experience and validation data*". Consequently, it refers both to ISO 5725, the international standard about the accuracy (trueness and precision) of measurement methods and results [42], and to the GUM [27].

Unfortunately, the discussions on uncertainty evaluation are mainly theoretical. Literature provides only few case studies, which allow a comparison of the different approaches [3, 60, 69]. For techniques prevailing in pharmaceutical analysis, especially for chromatography, the GUM approach is only scarcely applied [31]. A main reason for this is certainly the complex measurement structure of chromatographic techniques. While inter-laboratory studies are frequently reported for chromatographic methods, e.g. [67, 72, 75], among others due to the requirement of reproducibility assessment for official methods [41], the error-budget approach was only recently introduced to chromatographic methods [31]. An important difficulty is obviously the quantification of the uncertainty components [4]. This might also

be the reason why the error budget in [31] does for instance not consider the uncertainty related to signal processing.

Uncertainty can be expressed in two forms. The standard deviation obtained in the uncertainty assessment can be considered, which is referred to as standard uncertainty. In the following, unless otherwise specified, this standard uncertainty is considered. The expanded uncertainty U(x) defines an interval around the result of a measurement x±U(x) with U(x)=ku(x). The factor k is called the coverage factor. Usually, k=2, so that the expanded uncertainty is roughly equivalent to half the length of a 95 % confidence interval.

3 RESULTS

The results are regrouped in two main parts namely the study of the existing situation as well as the study of the particular themes and the determination of uncertainty in analytical measurements from interlaboratory study results. In this purpose, some liquid chromatographic (LC) and capillary electrophoretic (CE) methods were developed and validated.
The following interlaboratory studies were performed :

3.1 EVALUATION OF THE EXISTING SITUATION

3.1.1 The error-budget approach

The “guide to the expression of uncertainty in measurement” (GUM) was elaborated by a group of experts from different fields of metrology [27]. It suggests an error-budget approach in order to estimate the uncertainty related to a measurement result. The guideline is intended for all areas of measurement. The US National Institute of Standards and Technology (NIST) has published a comparable document [63]. Introduction and interpretation of the error-budget approach for analytical chemistry was performed by EURACHEM [16, 58]. The first edition of the EURACHEM guideline [58] closely followed the GUM [27] and rose a broad discussion on uncertainty evaluation in analytical chemistry. As a result of this discussion, the second edition of the EURACHEM guideline [17] was considerably adapted. It tried to make a compromise between the requirements of the GUM and the needs of analytical chemists. This search for a compromise will mainly be considered in section 4. Here, the discussion focuses on the classical error-propagation approach, as proposed in [27, 58].

The generalised procedure for the evaluation of uncertainty in analytical chemistry based on the error-budget approach consists of 4 steps [5, 17]:

Specification of the measurand:

The specification of the measurand encloses both a clear and unambiguous statement of the measurement subject and a quantitative expression of the relation between the measurand and the parameters (other measurands as well as constants and quantities, which cannot directly be measured) on which it depends.

Identification of uncertainty sources:

The identification of the uncertainty sources usually starts with the basic expression used to calculate the measurand from intermediate values. All these intermediate values are potential sources of uncertainty, but a comprehensive identification should also account for other parameters that only indirectly influence the calculation of the measurand, such as the time or temperature used for a certain step of the analytical process. A list of typical sources of uncertainty is given in [17, 58]. BCR, the Community Bureau of Reference of the European Union, stresses that one should always be aware of accounting for the whole analytical process, which includes sampling, storage, preparation, separation but also uncertainties caused by rounding during the calculation of the measurement result [11].

Quantification of the uncertainty components:

For each potential source of uncertainty identified in step 2, the magnitude of the uncertainty has to be measured or estimated. According to the GUM [27], the contributions of the different uncertainty components should be separately evaluated, using either repeated measurements (type A evaluation) or by other methods (type B evaluation). The GUM explicitly does not prefer one of the types of evaluation over the other [17]. If the uncertainty is derived in a type A evaluation, the experimental standard error on the mean is used as the estimator of the uncertainty of this component. For a type B evaluation, the GUM suggests to consider previous measurement data, experience with the behaviour of relevant materials and instruments, manufacturer's specifications, data provided in calibration, other certificates or uncertainties assigned to reference data taken from handbooks.

Calculation of the combined uncertainty:

The uncertainty contributions assessed in step 3 all have to be expressed in the form of a standard deviation. An error propagation approach is then applied to combine the different components to the combined standard uncertainty, u.

EURACHEM stresses that not all of the uncertainty components give a significant contribution to the combined uncertainty [17]. As a consequence, the guideline suggests that "*unless there is a large number of insignificant uncertainty sources*", those components that are "*less than one third of the largest need not be evaluated in detail*" [17]. However, the estimation whether an uncertainty contribution will exceed this one-third limit also requires an evaluation. How a less detailed and therefore less expensive evaluation shall be performed is not outlined in [17]. An approach by Caruso for a distinction between relevant and irrelevant contributions is based on a three-step process of brainstorming of experts with a fishbone diagram, factor weighing and Pareto analysis [12]. However, it is obviously too complex to be accepted in practice. Consequently, a complete assessment of all uncertainty

components is almost unavoidable except for those contributions, which are known to be negligible from previous experience.

3.1.2 Uncertainty evaluation based on precision assessment

3.1.2.1 Introduction and definitions

The precision expresses the closeness of agreement between independent test results obtained under stipulated conditions [45]. It is a measure for the variability within a series of measurement results. It is defined in relation to the central value of the distribution - usually a normal distribution – as the extent of dispersion of a series of results. Precision depends only on the distribution of random errors and does not relate to the true value or the specified value. It is usually expressed in terms of imprecision, either the standard deviation s or the variance s^2 of the test results [43]. Notice that for qualitative analyses, which lead to a “yes/no” decision for a given threshold concentration, the precision cannot be expressed as a standard deviation or variance. For this situation, IUPAC suggests to express the precision as the ratio between the number of false positive (negative) and the number of known negative (positive) results [61]. Consequently, the following is only valid for quantitative analyses.

Depending on the experimental conditions, the precision is usually evaluated as repeatability, intermediate precision or reproducibility standard deviation, respectively. The most extremes are the repeatability and reproducibility conditions. The repeatability and reproducibility conditions are defined as follows [43]:

Repeatability Conditions: Conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.

Reproducibility Conditions: Conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.

In addition to these most extreme precision conditions, ISO [44] defines intermediate precision conditions. In contrast to the reproducibility, intermediate precision can be evaluated within a single laboratory. Different factors can be modified between the measurements such as:

- the measurements can be performed at different times (on different days)
- the measurements can be performed by different operators
- the measurements can be performed using different equipment.

Of course, one can either modify only one of the factors or simultaneously modify several of them.

The repeatability, intermediate precision and reproducibility standard deviation are abbreviated as s_r, $s_{I(X)}$ and s_R, respectively. The symbol X in the intermediate precision refers to the factor(s), which are modified. For instance $s_{I(T)}$ expresses the time-different intermediate precision standard deviation, $s_{I(TOE)}$ is the (time+operator+equipment)-different intermediate precision standard deviation.

3.1.2.2 Intra-laboratory precision assessment

a. Repeatability assessment

The repeatability standard deviation can be estimated from independent measurements performed under repeatability conditions:

$$s_r = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \qquad \text{(Eq. 1)}$$

with x_i an individual measurement result and $\bar{x}$ the average result of n determinations. Independence of the measurement means that each test result is obtained in such a way that it is not influenced by any previous result on the same test object.

b. Assessment of intermediate precision

While the separate evaluation of the repeatability requires no special design for the experimental set-up, intermediate precision and reproducibility estimates are best evaluated in a balanced experimental design. The main advantage of such a balanced experimental set-up lies in the fact that it simultaneously allows obtaining a repeatability estimate. Usually, nested designs are applied; fully-nested designs are more common than staggered-nested designs [43]. In principle, intermediate precision can be estimated from intra- as well as from inter-laboratory studies. If an inter-laboratory approach is used, it has to be verified that the variances in the different laboratories are similar, so that they can be pooled [44]. As this requirement is often not fulfilled, and the intra-laboratory approach requires less experimental expense, the latter approach is usually preferred. The formulae given in the following apply to the intra-laboratory situation.

To estimate the **time-different intermediate precision**, a two-factor fully-nested design is applied: one operator performs n experiments on each of p days. An analysis of variance (ANOVA) yields the following mean squares:

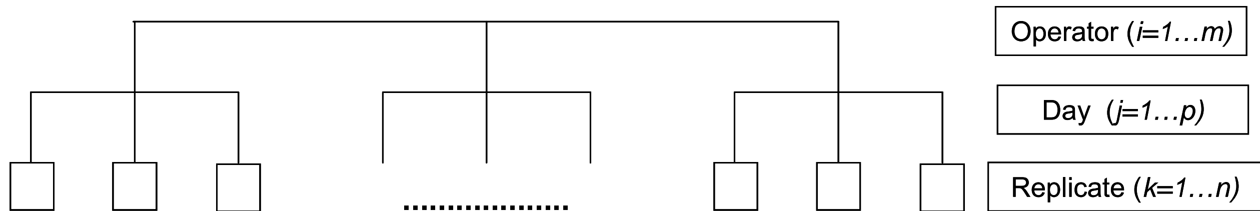
Source of variation	Mean square	Estimate of
Day	$MS_D = \frac{n\sum_{i=1}^{p}\left(\bar{x}_i - \bar{\bar{x}}\right)^2}{p-1}$	$\sigma_r^2 + n\sigma_D^2$
Residual	$MS_E = \frac{\sum_{i=1}^{p}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2}{p(n-1)}$	σ_r^2

with $\bar{\bar{x}} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{n} x_{ij}}{pn}$ and $\bar{x}_i = \frac{\sum_{j=1}^{n} x_{ij}}{n}$ the grand mean and the day means, respectively.

The mean squares given in the ANOVA table allow the estimation of different precision terms:

- the repeatability variance $s_r^2 = MS_E$ (Eq. 2)
- the between-day variance component $s_D^2 = \frac{MS_D - MS_E}{n}$ (Eq. 3), which is set to 0 if it is negative
- the time-different intermediate precision variance $s_{I(T)}^2 = s_r^2 + s_D^2$ (Eq. 4)

The estimation of the **(time+operator)-different intermediate precision** is based on a three-factor fully-nested design. The experimental set-up can be illustrated as follows:



With *m* operators each performing *n* replicates on each of *p* days, the following mean squares can be estimated:

Source of variation	Mean square	Estimate of
Operator	$MS_O = \frac{np\sum_{i=1}^{m}\left(\bar{x}_i - \bar{\bar{x}}\right)^2}{(m-1)}$	$\sigma_r^2 + n\sigma_D^2 + np\sigma_O^2$
Day within operator	$MS_D = \frac{n\sum_{i=1}^{m}\sum_{j=1}^{p}\left(\bar{x}_{ij} - \bar{x}_i\right)^2}{m(p-1)}$	$\sigma_r^2 + n\sigma_D^2$
Replicates within days (residual)	$MS_E = \frac{\sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\left(x_{ijk} - \bar{x}_{ij}\right)^2}{mp(n-1)}$	σ_r^2

with $\bar{\bar{x}} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n} x_{ijk}}{mpn}$, $\bar{x}_i = \frac{\sum_{j=1}^{p}\sum_{k=1}^{n} x_{ijk}}{pn}$ and $\bar{x}_{ij} = \frac{\sum_{k=1}^{n} x_{ijk}}{n}$ the grand mean, the mean value obtained by operator i and the day mean of operator i on day j, respectively.

The following precision estimates are obtained:

- the repeatability variance $s_r^2 = MS_E$ (Eq. 5)
- the between-day variance component $s_D^2 = \frac{MS_D - MS_E}{n}$ (Eq. 6), which is set to 0 if it is negative
- the time-different intermediate precision variance $s_{I(T)}^2 = s_D^2 + s_r^2$ (Eq. 7)
- the between-operator variance component $s_O^2 = \frac{MS_O - MS_E}{np}$ (Eq. 8), which is set to 0 if it is negative
- the (operator+time)-different intermediate precision variance $s_{I(OT)}^2 = s_O^2 + s_D^2 + s_r^2$ (Eq. 8)

In the same way, further factors (such as the factor instrument) can be added to the model in order to obtain other intermediate precision estimates.

Often, only a small number of e.g. operators will be available, leading to a poor estimate for s_O^2. In such a situation, it can be advantageous to evaluate the combined (time+operator) variance component instead of the individual components. Examples for this way of precision evaluation can be found in [51].

3.1.2.3 Inter-laboratory precision assessment

The reproducibility standard deviation, s_R, is assessed in an inter-laboratory precision evaluation study. These studies, often also referred to as collaborative trials, are considered in detail in the following review [34].

Notice that the calculations for the ANOVA to derive the reproducibility standard deviation are comparable to the calculations described above to derive an intermediate precision standard deviation with one factor modified. Again, a balanced design also allows estimating the repeatability standard deviation s_r.

3.1.2.4 Application of precision data in uncertainty evaluation

The use of precision data in uncertainty evaluation has a long tradition in analytical chemistry. Different approaches to derive the uncertainty of a measurement result from precision data have been proposed as alternatives to the GUM [2, 56].

a) The top-down approach: consideration of reproducibility

In the course of the uncertainty discussion following the presentation of the GUM [27], AMC - as already mentioned - proposed the top-down approach as an alternative uncertainty evaluation method specially tailored for the measurement structure of analytical chemical measurements [2]. It is based on reproducibility data and fulfils the requirement to include systematic errors, because systematic errors within a laboratory become random errors if a population of several laboratories is considered. The reproducibility standard deviation s_R derived in an inter-laboratory study includes all uncertainty contributions that randomly occur in a population of laboratories in the determination of a certain analyte. Consequently, AMC proposed it as an estimate of the uncertainty of the measurement result. However, as was discussed in the preceding section, the true value of the analyte concentration is usually unknown, so that an assigned value is mostly used as the best estimate of the concentration. This assigned value is also related to an uncertainty, u_a, which AMC also considers in the derivation of the uncertainty of the measurement result x: $u_x = \sqrt{s_R^2 + u_a^2}$ (Eq. 9)

b) Uncertainty evaluation based on intermediate precision

Parallel to the AMC in the United Kingdom, other national or multinational committees for standardisation developed alternative approaches to the GUM. The Nordic Committee for Food Analysis

(NMKL) also agreed that the error-budget approach better suited to physical than chemical measurements. It designed a procedure that derives the uncertainty of a measurement result from intermediate precision data [56].

NMKL is fully aware that intermediate precision data cannot account for systematic errors. For the assessment of systematic errors, it proposes comparison with a certified reference material, with an established reference method or the use of recovery experiments. The measurement result is then corrected for this systematic error and the uncertainty is estimated from what NMKL refers to as “internal reproducibility”, in ISO terms the operator+time-different intermediate precision standard deviation: $u(x)=s_{I(OT)}$ [56]. Notice that this approach does not include any uncertainty in the evaluation of the systematic error.

NMKL proposes a further approach for uncertainty estimation for situations, in which it is not possible to estimate an intermediate precision, e.g. unstable analytes. It is based on repeatability data measured in duplicate determinations of different materials. However, NMKL immediately warns that this approach will result in an underestimation of the measurement uncertainty [56].

3.1.3 Operational definition of uncertainty

Sections 2 and 3 provide an overview on the main approaches for uncertainty estimation. The term “uncertainty” is applied independent of the way the uncertainty is assessed. For instance, in the error-budget approach, no distinction is made whether covariances are taken into account or not and whether the uncertainty contribution, e.g. due to sampling, is taken into account. Section 3 showed that precision can be evaluated either in an intra- or in an inter-laboratory approach. The Nordic committee for food analysis (NMKL) prefers intermediate precision rather than repeatability estimates [56], whereas the Analytical Methods Committee of the Royal Society of Chemistry (AMC) bases its uncertainty evaluation on reproducibility estimates [2]. These examples clearly show that the term uncertainty does not have the same meaning in all situations. The reproducibility variance must be expected to be larger than the repeatability variance, because the former is the sum of the latter and the between-laboratory variance component. Accordingly, the uncertainties derived with the different approaches clearly refer to different estimates, but this difference does not become clear from the general term “uncertainty”.

ISO defines different terms for the precision estimation according to the circumstances under which the precision is evaluated. No such terms are defined for uncertainty assessment, but a distinction between the different estimates seems recommendable as well. Valcarcel and Rios [66] propose to distinguish between “specific uncertainty”, which only considers the precision evaluation, and “generic uncertainty”, which includes the trueness evaluation as well. However, this “specific uncertainty” provides no information about the type of precision evaluation considered, neither does “generic

uncertainty" express how the trueness is assessed and to which reference the result is traceable. In order to clarify the term uncertainty, some more detailed operational definitions of uncertainty are proposed in the following article [37].

3.1.4 Comparison of different methods to evaluate the uncertainty of an HPLC assay

The preceding sections demonstrate that in the past, uncertainty evaluation was mainly discussed on a theoretical base. For the different approaches, only few case studies are available. While the top-down approach [2] is set-up in a similar way for all analytical methods, the error-budget approach [26] splits the analytical process considered into its uncertainty sources and is therefore method-dependent. Recently, a case study for the top-down approach has been published [21]. It considers a HPLC analysis with refractometric detection for simple sugars [21]. Some case studies can be found for the error-budget approach. However, most of them focus on analytical methods, which have a very clear measurement structure, so that the identification and usually also the quantification of the uncertainty contributions is straightforward [3, 60, 69]. It is therefore not surprising that the standardisation and titration examples given by EURACHEM [18] already mentioned in chapters 7 and 8 are very popular examples for demonstrating uncertainty estimation, as follows for instance from [62]. For chromatographic methods, which play an essential role in many analytical areas and which are known to have a complex measurement structure, only few case studies have been published yet. There are two case studies, which apply the error-budget approach in HPLC [31] or GC/MS [59], respectively. A further case study for HPLC reconsiders validation data in a modified error-budget approach [52]. However, the number of uncertainty sources considered in these case studies is rather restricted. While the above-mentioned case study for the top-down approach [21] can serve as a base for all applications, it is clear that more case studies are required in order to demonstrate to analysts how uncertainty estimation can be performed for a particular analytical technique with the error-budget approach.

Only few comparisons have been performed yet between the uncertainty approaches. A comparison study was performed on initiation of the UK Ministry of Agriculture, Fisheries and Food and focused on some selected methods for food analysis. It revealed that the error-budget and the top-down approach yield comparable uncertainty estimates [74]. A further comparison has been performed for coal analysis [69]. The uncertainty estimates obtained for HPLC with the error-budget approach [51] were compared with the results of an inter-laboratory study; the results were indeed comparable. It is desirable that comparisons be also performed in other measurement areas and with different analytical methods.

Within the project a case study has been performed, which applies both the error-budget [26] and the top-down approach [2] to a chromatographic method. It reconsiders the HPLC assay of the

European Pharmacopoeia for tylosin for veterinary use [19] already considered in the derivation of SST limits from a robustness test in chapter 5. Besides the robustness data, data from an inter-laboratory study [67] were available for this method, too. The reconsideration of these data was expected to reduce the experimental expense required for uncertainty estimation. Since robustness data were available, an uncertainty approach presented by Barwick and Ellison [6] was also considered. It combines precision, trueness and robustness data. Moreover, it is also tried to estimate the uncertainty only from robustness data, which at least in the pharmaceutical field should be available for all validated methods [41]. This alternative approach is based on the idea that robustness tests can be considered as intra-laboratory simulations of inter-laboratory studies if the modifications introduced in the robustness test correspond to the variations to be expected in an inter-laboratory transfer of the method.

In terms of the operational definitions of uncertainty proposed in the preceding chapter, the error-budget approach should lead to a within-laboratory uncertainty, while the top-down approach results in a reproducibility uncertainty. In the approach based on robustness data only, one only uses uncertainty estimates assessed within a single lab. However, as the robustness test is considered a simulation of an inter-laboratory study, the corresponding uncertainty estimate should rather be considered a reproducibility uncertainty than a within-laboratory uncertainty. The uncertainty estimate obtained from the approach by Barwick and Ellison can be considered a bias-included uncertainty since it comprises a bias (trueness) assessment.

The comparison is summarized in the following article [38].

The case study did not reveal relevant differences between the different uncertainty approaches. However, it contradicts the misgiving by the GUM [26] that other approaches than the error-budget approach tend to an underestimation of the uncertainty because it is supposed [26] that they should overlook some of the uncertainty sources. In contrast, these other approaches usually yield slightly larger uncertainty estimates than the error-budget approach.

The case study allows some more conclusions for the error-budget approach. If it is not known whether the interaction between different uncertainty sources plays an important role, the experimental set-up should account for the estimation of their covariances. Both the fishbone diagram and a reconsideration of interaction effects from robustness tests can help in the decision whether covariances have to be taken into account. The problem observed with the covariances in the example discussed confirms that with the error-budget approach, the results depend to a large degree on the estimations of the analyst [16], as he decides for instance whether covariances can be considered negligible or not. The example considered also showed that for some uncertainty contributions, e.g. the absorption, it can be advantageous not to completely split the uncertainty into its basic sources but to assess a common uncertainty estimate.

As the case study did not reveal important differences in the uncertainty estimates obtained with the different approaches, practical and economic aspects predominate the decision about the uncertainty approach. If only robustness data are available, estimation based on the approach considering only

robustness data will be preferred. The approach by Barwick and Ellison [6] in addition also requires precision and trueness data, which should however be available if the method has been fully validated. In the same way as the error-budget approach [26], these two approaches provide information about the contribution of the different uncertainty sources. The top-down approach [2] might be preferred if the precision of the method has already been estimated in an inter-laboratory study and if no further information about the uncertainty contribution of the different sources is required. The error-budget approach [26] might be preferred if only few data are available from the method validation and if the measurement structure does not comprise too many uncertainty sources. It depends on the individual situation whether the data required for the error-budget approach are all available from manufacturers' specification or whether some additional measurements are required.

3.2 DETERMINATION OF UNCERTAINTY IN ANALYTICAL MEASUREMENTS FROM INTERLABORATORY RESULTS

3.2.1 Introduction

The correct interpretation of a measurement result requires knowledge about the uncertainty of the result. Although important, the determination of uncertainty is not evident. Several approaches for estimating uncertainty in analytical measurements are proposed as already mentioned in the introduction : the bottom-up approach and the top-down approach where uncertainty is measured by using information from the validation process. In the ISO approach all sources of uncertainty are estimated individually. The ISO approach was originally proposed for quantifying uncertainty in physical measurements and was subsequently adapted by EURACHEM for chemical measurements. However the practical application of the ISO approach in chemical measurements is cumbersome and this prevents its widespread use. The "top-down" approach uses interlaboratory studies to measure uncertainty. In this approach the lab is seen from a higher level and thus both systematic and random errors within one laboratory become random errors when they are considered from this higher level. Uncertainty can also be derived from validation data. During the validation of an analytical method, the trueness and precision are tested. The results from this validation process might be used to make an uncertainty statement. However the relationship between validation and uncertainty statement is not clear for many analysts. There is a need to clarify this relationship and to show how the existing data can be applied in practice.

Depending on the situation under which the analyst is validating, different operational definitions of uncertainty have been proposed. Within-laboratory uncertainty, as proposed by Hund et al. [34], only considers the intermediate precision and contains the repeatability and the between-run effect. If an

analysis is performed by different laboratories, a between-laboratory effect will exist. The combination of the between-laboratory effect and the repeatability then determines the reproducibility uncertainty on an analytical measurement. When, moreover, the method bias is taken into account as source of uncertainty, bias-included uncertainty is obtained.

In a practical view point, this project aims to evaluate reproducibility uncertainty from interlaboratory study results. Six interlaboratory studies were planned :

1) An interlaboratory study on the analysis of a phenoxymethylpencillin (pen V) sample consisting of a Karl-Fischer water determination, of a liquid chromatography (LC) test to determine 4-hydroxyphenoxymethylpenicillin and other impurities, both methods based on the European Pharmacopoeia [20] monograph and of a potentiometric acid-base titration to assay the content of pen V;
2) An interlaboratory study in order to validate a new LC method for the analysis of erythromycin;
3) An interlaboratory study in order to determine R-timolol maleate and other related substances in S-timolol maleate samples using a developed chiral LC method in normal phase mode adapted from European Pharmacopoeia [20] monograph;
4) An interlaboratory study on a method by capillary electrophoresis (CE) to analyse metacycline;
5) A next interlaboratory study on a CE method for the enantiomeric determination of S-ketoprofen and R-ketoprofen is being prepared;
6) Another interlaboratory study on a CE method for the analysis of amoxicillin is also being prepared.

3.2.2 Set-up of the interlaboratory studies

In order to allow a statistical evaluation of the results of the interlaboratory study, a specific set-up is needed. Figure 1 shows a typical set-up used in the studies.

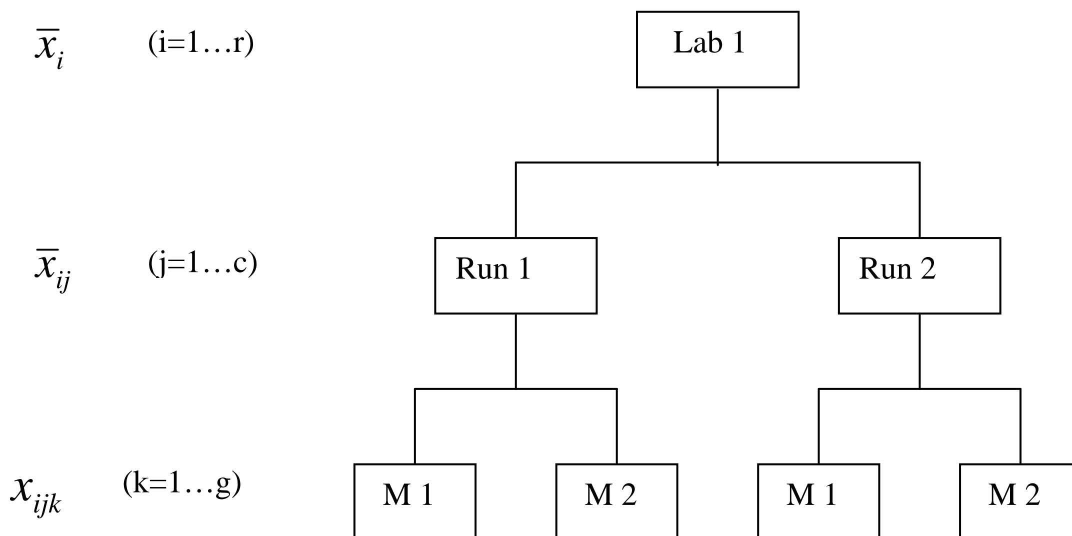


Figure 1: Typical set-up of an interlaboratory study.

Each lab (i=1…r) participating in the study, performed two independent runs and in each run two replicate measurements were carried out. This allows the estimation of the variance between measurements (s^2_m), the variance between runs (s^2_{run}) and the variance between laboratories (s^2_L).

If runs are performed under repeatability conditions, which means in a short period of time (= on the same day), by the same analyst and on the same equipment, the repeatability variance can be calculated as: $s^2_r = s^2_m + s^2_{run}$. (Eq. 10).

If two runs are carried out on different days, $s^2_{run} = s^2_{day}$, and the intermediate precision is estimated and more specifically the time-different intermediate precision is considered. It is calculated as: $s^2_{I(T)} = s^2_m + s^2_{day}$ (Eq. 11) When the between-laboratory variance is taken into account, an estimate of the reproducibility variance is obtained:

$s^2_R = s^2_r + s^2_L$ or $s^2_R = s^2_{I(T)} + s^2_L$ (Eq. 12)

The standard uncertainty of an individual measurement done by a laboratory becomes:

$$u_x = s_R = \sqrt{s^2_m + s^2_{run} + s^2_L} \qquad \text{(Eq. 13)}$$

If a lab performs one run of *g* measurements:

$$u_{\bar{x}} = \sqrt{s^2_m/g + s^2_{run} + s^2_L} \qquad \text{(Eq. 14)}$$

If a lab performs c runs of g measurements:

$$u_{\bar{x}} = \sqrt{s^2_m/(c \times g) + s^2_{run}/c + s^2_L} \qquad \text{(Eq. 15)}$$

The expanded uncertainty for an approximate level of confidence of 95%, U_x, is calculated as: $U_x = ku_x = 2u_x$. A result x can than be given as $x \pm U_x$.

3.2.3 Interlaboratory studies

3.2.3.1 Interlaboratory on the analysis of a pen V sample

3.2.3.1.1 Introduction

Nine labs participated in this interlaboratory study and followed an identical protocol. First a training round was held as prescribed in ISO 5725-2 [47], in order to familiarise the labs with the methods and to evaluate the protocol. In the final study, each lab performed two independent runs of experiments, what means that for the second run all solutions were newly prepared. In each run the same solution was measured twice (Figure 1, r = 8, c = 2, g = 2).

The two runs of the acid-base titration and of the Karl-Fischer water determination were performed on the same day under repeatability conditions. For the LC, it was not possible to perform two runs on the same day and they were analysed on different days, but anyway in the shortest time possible.

Additionally for the acid-base titration a separate experimental set-up was used to evaluate the within-laboratory uncertainty. Three labs (labs 1, 6 and 8) performed two measurements during three days (Figure 2).

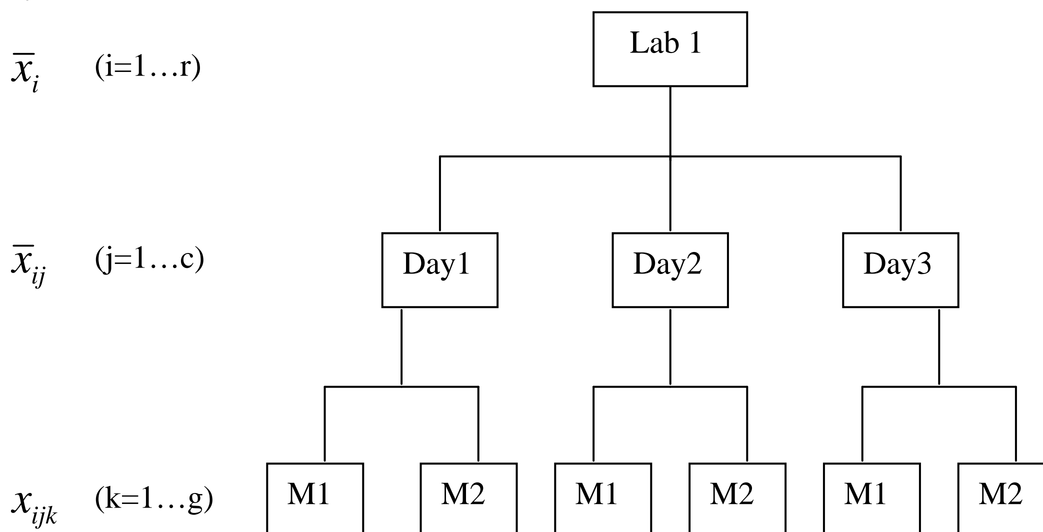


Figure 2: Set-up of the study for the determination of the intermediate precision, r = 3, c = 3 and g = 2.

3.2.3.1.2 Acid-base titration for the determination of pen V

I. Repeatability and reproducibility uncertainty

Figure 3 shows the results for the acid-base titration from 8 different labs (r = 8). The results from the study of the within-laboratory uncertainty, which are also shown (lab 1', 6' and 8'), are discussed later. The statistical analysis was performed according to ISO 5725-2 [47] and started with an outlier evaluation. Both Cochran's test (within-laboratory variability) and Grubbs' test (between-lab variability) were carried out. After the outlier evaluation, the analysis of variance allowed to calculate s^2_m = 0.0887, s^2_{run} = 0.140, and s^2_L = 0.526.

Results of the variance analysis are summarized in Table 1. The repeatability variance is then estimated to be s^2_r = 0.0877 + 0.140 = 0.229 and the reproducibility variance is s^2_R = 0.229 + 0.526 = 0.755. The ratio between the reproducibility and the repeatability variance is about three, what can be considered as normal. The variance between measurements, s^2_m, is lower than the variance between runs, s^2_{run} which also can be seen in Figure 3. For most labs the variability between runs is larger than within runs. This is logical and can be explained by the fact that for every run the standardisation of the newly prepared 0.1 M NaOH solution had to be carried out, introducing an additional variability.

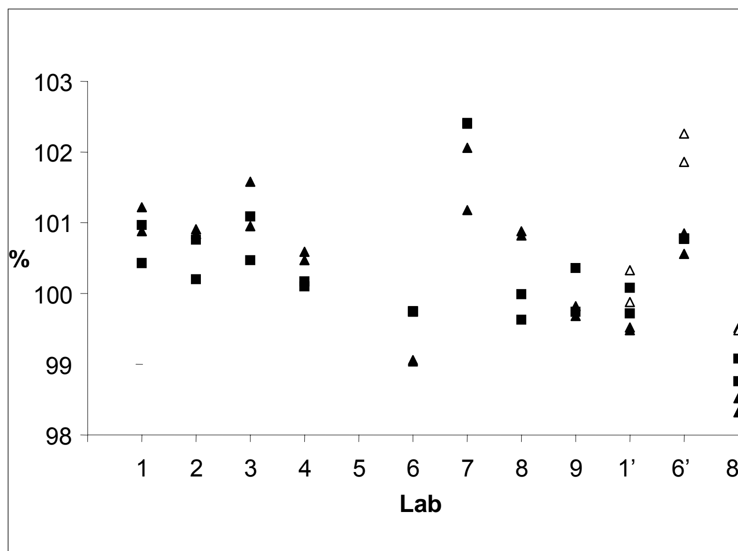


Figure 3 : Acid-base titration: Results of the interlaboratory study (labs 1-9) combined with results of the intermediate precision study (labs 1’, 6’, 8’). Lab 5 did not have results for this

A laboratory that applies the acid-base titration method for the determination of penV can use the information obtained from this interlaboratory study to make an uncertainty statement. The lab then has to show that it is sufficiently proficient. This implies that the in-house repeatability of the lab is similar to the repeatability in the interlaboratory exercise. This can be evaluated by comparison of the repeatability variances by means of an F-test. If repeatabilities are similar, the reproducibility standard deviation from the collaborative study can be used in the uncertainty statement. The standard uncertainty of a single result *x* obtained in the laboratory, taking into account the variance between laboratories, becomes:

$$u_x = s_R = 0.869$$

The expanded uncertainty, using a coverage factor k = 2, is then:

$$U_x = 2u_x = 1.738\ .$$

The result x is then reported as $x \pm 1.7$. If the laboratory reports the mean of one run of two measurements, the standard uncertainty is calculated as:

$$u_{\bar{x}} = \sqrt{s_m^2/2 + s_{run}^2 + s_L^2} = 0.834$$

The standard uncertainty for a laboratory mean obtained from e.g. 2 runs of 3 measurements is:

$$u_{\bar{x}} = \sqrt{s_m^2/(2\times 3) + s_{run}^2/2 + s_L^2} = 0.782$$

II. Within-laboratory uncertainty

From this study the variance between measurements, s^2_m, and the variance between runs, s^2_{run} (= variance between days, s^2_{day}), were calculated to be 0.040 and 0.302, respectively. It follows that the time-different intermediate precision is estimated to be $s^2_{I(T)} = s^2_m + s^2_{day} = 0.342$. Notice that in this expression s^2_{day} is not a pure variance component due to time since besides the effect of changes of time, s^2_{day} also includes effects due to the fact that each day the solutions were newly prepared, the latter being part of the repeatability variance. An estimate of the pure day component can therefore be obtained from $s^2_{I(T)} – s^2_r$ with s^2_r as observed in the previous study. Since $s^2_{I(T)} = 0.342$ and $s^2_r = 0.229$ it can be concluded that the contribution of the pure day-to-day variance for the titration ($s^2_{day\text{-}to\text{-}day} = 0.113$) is similar to the between-run variance.

3.2.3.1.3 Karl-Fischer water determination

Six labs performed the Karl-Fischer water determination. No outliers were found with Cochran's test nor with the Grubbs tests. The results of the analysis of variance are shown in Table 1.

Method	s^2_m	s^2_{run}	s^2_L	s^2_r	s^2_R	$s^2_{I(T)}$
Acid-base titration (100.6 %)	0.0887	0.140	0.526	0.229	0.755	0.342
Karl-Fischer (0.31 %)	$1.48\ 10^{-3}$	0.00[a]	$13.9\ 10^{-3}$	$1.48\ 10^{-3}$	$15.4\ 10^{-3}$	---[b]
LC: 4 –hydroxy (2.61 %)	$1.19\ 10^{-3}$	$3.20\ 10^{-3}$	$77.1\ 10^{-3}$	---[b]	$81.5\ 10^{-3}$	$4.39\ 10^{-3}$
LC: other impurities						
Peak 2 (0.14 %)	$2.40\ 10^{-3}$	0.00[a]	$15.2\ 10^{-3}$	---[b]	$17.6\ 10^{-3}$	$2.40\ 10^{-3}$
Peak 3 (0.28 %)	$4.39\ 10^{-4}$	$1.03\ 10^{-3}$	$8.60\ 10^{-3}$	---[b]	$10.1\ 10^{-3}$	$1.47\ 10^{-3}$
Peak 4 (0.051 %)	$9.22\ 10^{-5}$	0.00[a]	$28.4\ 10^{-5}$	---[b]	$37.6\ 10^{-5}$	$9.22\ 10^{-5}$
Peak 5 (0.22 %)	$1.06\ 10^{-4}$	$1.14\ 10^{-4}$	$14.9\ 10^{-4}$	---[b]	$17.1\ 10^{-4}$	$2.20\ 10^{-4}$
Peak 6 (0.28 %)	$1.55\ 10^{-4}$	$6.28\ 10^{-4}$	$8.98\ 10^{-4}$	---[b]	$16.8\ 10^{-4}$	$7.83\ 10^{-4}$

Table 1: Variance components obtained from the analysis of variance for the different methods. Values in parentheses show the general mean $\overline{x}$. [a] Calculated value < 0, which is set equal to zero, [b]set-up did not allow to calculate this value

The most remarkable result is that s^2_{run} is estimated to be zero, i.e. the standardisation of the iodosulphurous reagent did not introduce additional variability. This is different from the acid-base titration, where an influence of the determination of the NaOH concentration was noticed. A possible explanation is the small amount of water present in the sample. The results found fluctuate between 0.18 % and 0.56 %. Due to this small amount of water the variability of the replicated water determination of the sample will be higher than the variability introduced by an additional standardisation. The standard reproducibility uncertainty of an individual result *x*, obtained in a single laboratory, is calculated as $u_x = s_R = 0.124$.

3.2.3.1.4 Liquid Chromatography

Nine labs performed the liquid chromatography. The liquid chromatography consisted of two parts: the determination of the 4-hydroxypen V and of the other impurities. The impurities 4-hydroxypen V (peak 1 in Figure 4), penicilloic acids (two diastereoisomers, peaks 2 and 3), benzylpenicillin (peak 4) and penilloic acids (two diastereoisomers, peaks 5 and 6) elute in the isocratic part of the chromatogram and most of the unidentified impurities after the main peak, during the gradient elution (Figure 4).

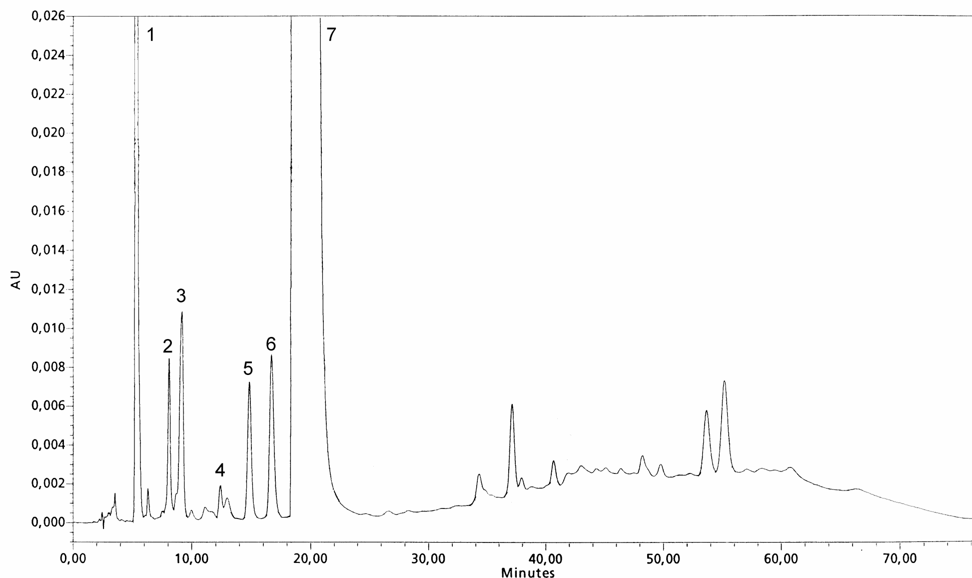


Figure 4 : Typical chromatogram of the pen V test sample (4.0 mg/ml) on a Symmetry C18 column. 1 = 4-hydroxypen V, 2,3 = penicilloic acids (2 diastereoisomers), 4 = benzylpenicillin, 5,6 = penilloic acids (2 diastereoisomers), 7 = pen V.

The analysis of variance was carried out on all known impurities. The results are shown in Table 1. For the determination of the 4-hydroxypen V the reproducibility variance, s^2_R, is 18 times larger than the intermediate precision, $s^2_{I(T)}$. The mean lab results for the 4-hydroxypen V content fluctuate from 2.22 % to 3.16 %. The larger variability on these results can be explained. To improve the sensitivity of the LC

method, the wavelength of detection was changed from 254 nm to 225 nm after the training round (see experimental). For this study however, the 4-hydroxypen V was measured against a dilution of pen V in a reference solution. When a spectrum of both compounds, dissolved in the mobile phase, was taken, it was clear that they were not exactly the same. Between 220 nm and 230 nm, the curve of pen V decreases sharply while for 4-hydroxypen V a small shoulder can be observed. These results explain the between-lab variability: all labs set their detector at a wavelength of 225 nm, but not all detectors will measure exactly at this wavelength. If a detector measures at a wavelength different from 225 nm, a different ratio A (4-hydroxypen V) / A (pen V) will be obtained, resulting in variability on the 4-hydroxypen V content found. The intermediate precision however will be good, as a lab does not change the wavelength of its detector during the experiments. This study shows the importance of controlling the equipment but also of the use of the reference substance corresponding to the impurity, which has to be quantified. The above also demonstrates that it is not always without risk to change the conditions (here the wavelength) of a prescribed and validated method. In our situation it causes an unexpectedly high difference between the reproducibility and intermediate precision. However, these results do not prevent from demonstrating the approach to derive uncertainty estimates. The ratios reproducibility variance, s^2_R, to intermediate precision, $s^2_{I(T)}$, for the other known impurities (Table 1) fluctuate between 2 and 8.

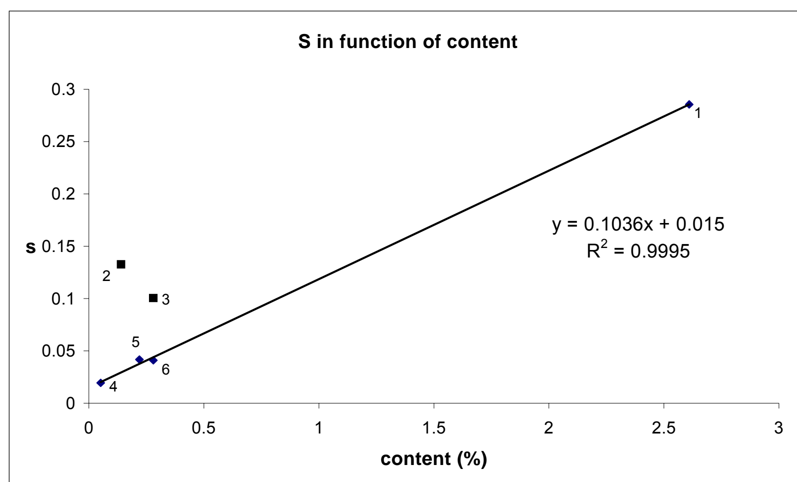


Figure 5: Reproducibility standard deviations as a function of content. ◆s_R on content results for 4-hydroxypen V, benzylpenicillin and penilloic acids, ■ s_R on content results for the penicilloic acids

In Figure 5 the reproducibility standard deviations are plotted as a function of the content found for the different substances. The *s* or the u_x of substances 1, 4, 5 and 6 seem to fall on a straight line. The uncertainty of the penicilloic acids seems to be considerably larger. The latter can be explained by the fact that these substances probably were formed in solution from pen V during the execution of the study. Even though the protocol mentioned a maximum storage time of the test solution, different amounts of substances 2 and 3 may be present in the solution to be analysed, increasing their variability and uncertainty. If *s* or u_x of substances 1, 4, 5 and 6 indeed fall on a straight line, it would allow (i) to make a final prediction of u_x based on the straight line equation, and (ii) to make predictions about the

uncertainty of individual measurements of substances 1, 4, 5 and 6 when their concentration is situated in the interval [0.05 %; 2.6 %].

3.2.3.1.5 Conclusion

This study shows how different uncertainty estimates of analytical measurements can be determined and how the results of an interlaboratory study can be used to estimate the uncertainty on future results by a single lab analysing similar samples.

3.2.3.2 Interlaboratory on a LC method for the analysis of erythromycin

3.2.3.2.1 Introduction

In the second interlaboratory study a new LC method for the analysis of erythromycin was validated and it was shown how the validation data can be used in future by a lab that wants to make an uncertainty estimation concerning the same analysis.

Erythromycin is a mixture of macrolide antibiotics produced by Saccharopolyspora erythreas during fermentation. The main component of erythromycin is erythromycin A (EA). The following related substances are formed during the fermentation process: erythromycin B (EB), erythromycin C (EC), erythromycin D (ED), erythromycin E (EE), erythromycin F (EF) and N-demethylerythromycin A (NdMeEA). In mild acidic conditions EA degrades to erythromycin A enol ether (EAEN) and anhydroerythromycin A (AEA). Pseudoerythromycin A enol ether (PsEAEN) and pseudoerythromycin A hemiketal (PsEAHK) are formed at (slightly) alkaline pH. Erythromycin A N-oxide (EANO) and erythronolide B may also be found in bulk substance and commercial products.

Two new liquid chromatography methods for the analysis of erythromycin have been developed recently. The first method was developed in our laboratory and uses an Xterra RP18 column [13, 70]. Acetonitrile-0.2 M K_2HPO_4 pH 7.0 –water (35:5:60) is used as a mobile phase. The second method was developed by Abbott and uses an Astec C18 polymeric column. Both showed clear improvements compared to the method currently prescribed by European Pharmacopoeia and United States Pharmacopoeia. After comparison of the methods, the Astec method was chosen for this interlaboratory study because of the good column stability. The mobile phase consists of acetonitrile-0.2 M K_2HPO_4 pH 9.0 –water (40:6:54) at a flow rate of 1.0 ml/min. The column is kept at a temperature of 50 °C and detection is done at 215nm. The injection volume is 100 µl.

3.2.3.2.2 Set-up of the study

Eight laboratories participated in this interlaboratory study and each lab analysed four samples. Each sample was analysed twice under repeatability conditions (Figure 6).

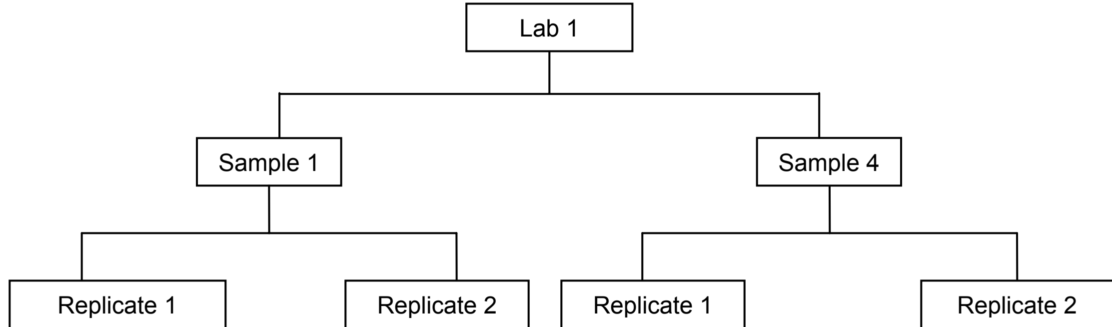


Figure 6: Set-up of the interlaboratory study on a method for erythromycin

3.2.3.2.3 Results

Before a variance analysis was carried out, the outliers were detected. Outliers were removed by Cochran's test and Grubbs' tests. Also Mandel's k (within-laboratory variability) and Mandel's h (between-lab variability) were used as a graphical consistency technique. They were calculated as described in ISO 5725-2 [47]. The figures derived with these values give an overview of the results and possible problems. Results for one sample are shown in Figure 7.
After removing the outlying values, a variance analysis was carried out. Repeatability, between-laboratory and reproducibility variances were estimated. Repeatability variance is calculated as:

$$s_{rj}^2 = \frac{1}{2p}\sum_{i=1}^{p}(y_{ij1} - y_{ij2})^2 \qquad \text{(Eq. 16)}$$

with *p* = total number of labs, *i* = 1 to p, y_{ij1} the first and y_{ij2} the second replicate in lab *i* for substance *j*.
This equation can only be used in the particular case where $n_{ij}=n=2$, which was the case in this study. The equation to calculate the between-laboratory variance is:

$$s_{Lj}^2 = \frac{1}{p-1}\sum_{i=1}^{p}(\bar{y}_{ij} - \bar{\bar{y}}_j)^2 - \frac{s_{rj}^2}{2} \qquad \text{(Eq. 17)}$$

with *p* = total number of labs, *i* = 1 to p, *j* = substance, $\bar{y}_{ij}$ the lab mean for substance j and $\bar{\bar{y}}_j$ the general mean of all labs for substance *j*. Also this equation can only be used if $n_{ij}=n=2$.
Reproducibility variance is calculated as the sum of repeatability variance and between-laboratory variance: $s^2_{Rj} = s^2_{rj} + s^2_{Lj}$ (Eq. 18)

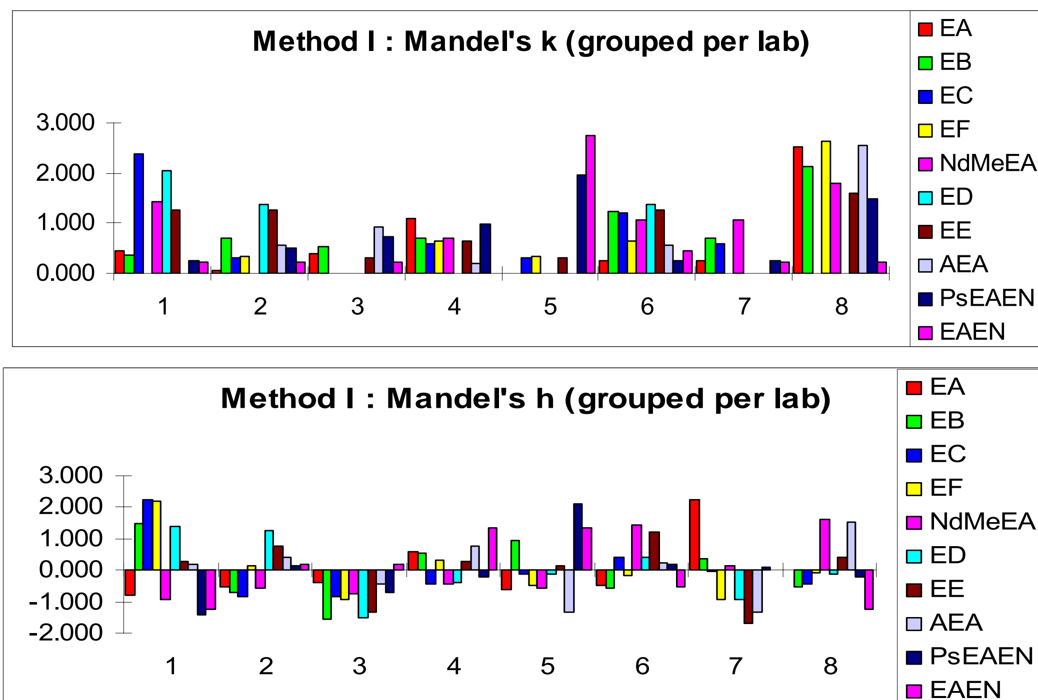


Figure 7: Mandel's *k* statistic for a sample, used to examine within-laboratory consistency and Mandel's *h* statistic for a sample, used to examine between-laboratory consistency, grouped per laboratory.

A typical chromatogram of one of the samples is shown in Figure 8. No selectivity problems were mentioned nor observed by any laboratory. The contents of EA and all identified substances were calculated

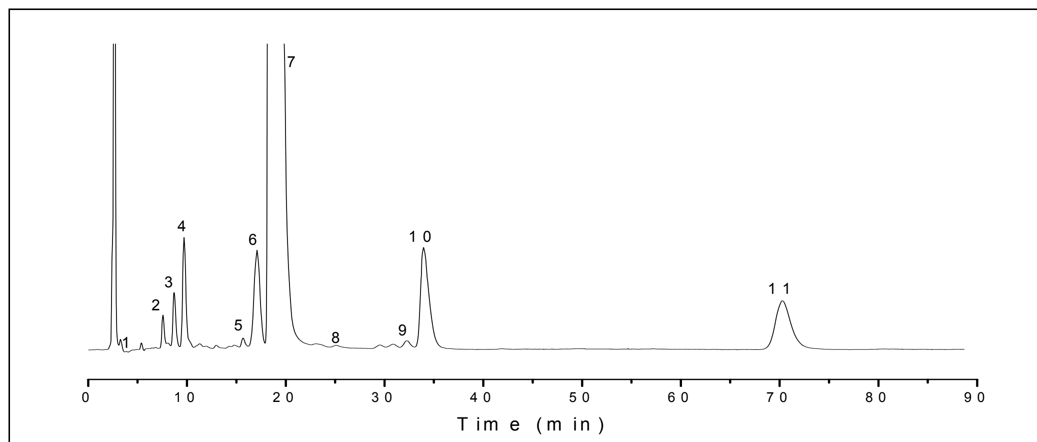


Figure 8: Typical chromatogram of erythromycin sample solution. 1 = EANO (no peak in this sample), 2 = EF, 3 = NdMeEA, 4 = EC, 5 = ED, 6 = EE, 7 = EA, 8 = AEA, 9 = PsEAEN, 10 = EB, 11 = EAEN

The four samples in this study showed similar contents for erythromycin and related substances and therefore these reproducibility variances were pooled. The pooled reproducibility variances for EA and all known impurities are shown in Table 2. They give an idea about the reproducibility of the method.

The ratio reproducibility variance to repeatability variance was also calculated. This ratio depended on the concentration of the substance in the sample. For the main peak (EA) the ratio is found to be 3.5, which indicates the good reproducibility of the determination of EA compared to the repeatability achieved with this method. For the impurities it can be noticed that the ratios are larger, as their concentrations in the sample are smaller. The relatively small ratios for PsEAEN and EAEN, although present in small amount and eluted late in the chromatogram, can be explained by their higher response factors, which lead to higher peak areas compared to other impurities, resulting in a smaller variability on their content determination.

	Substance									
Pooled Variances	EA	EB	EC	EF	NdMeEA	ED	EE	AEA	PsEAEN	EAEN
Repeatability	3.11E-01	1.31E-03	5.69E-04	1.46E-04	2.54E-04	2.20E-04	1.17E-03	4.85E-04	1.18E-05	1.44E-04
Between-lab	7.68E-01	1.76E-02	1.70E-02	5.63E-03	4.27E-03	1.25E-03	1.61E-02	5.46E-03	1.20E-04	1.40E-03
Reproducibility	1.08E+00	1.89E-02	1.75E-02	5.78E-03	4.53E-03	1.47E-03	1.72E-02	5.94E-03	1.32E-04	1.54E-03
Mean content (%)	85.24	3.72	1.16	0.39	0.68	0.15	2.46	0.09	0.04	0.28
Ratio	3.5	14.4	30.8	39.6	17.8	6.7	14.7	12.3	11.1	10.7

Table 2: Pooled variances from the four samples. Ratio = repeatability to reproducibility variance ratio

A laboratory that applies this LC method to analyse erythromycin can use the information obtained from this interlaboratory study to make an uncertainty statement on the results. Therefore, the lab first has to prove that it is sufficiently proficient, i.e. the in-house repeatability has to be similar to the repeatability obtained in this interlaboratory study. This can be evaluated by comparison of the repeatability variances by means of an F-test. If the repeatability is similar, the reproducibility standard deviation from the collaborative study can be used in the uncertainty statement. The standard uncertainty of a single result x from a single lab for the content determination of EA becomes $u_x = 1.039$. The expanded uncertainty, using a coverage factor k=2, is $U_x = 2.078$. The result x of this lab can than be written as $x \pm 2.1$.

If, in an individual laboratory, the content of EA is measured three times under repeatability conditions, the standard uncertainty on the mean result becomes:

$$u_{\bar{x}} = \sqrt{s_{r(EA)}^2 / 3 + s_{L(EA)}^2} = 0.934$$

The expanded uncertainty is $U_{\bar{x}} = 1.868$ and the mean result of the lab can be written as $\bar{x} \pm 1.9$. It is observed that analysing the sample three times instead of once does not improve the uncertainty considerably. In fact this is logical as the between-laboratory uncertainty contributes most to the total uncertainty. Nevertheless, the analysis should be done several times in order to prove that the in-house repeatability is similar to the repeatability found in this interlaboratory study.

If a company has two sites, both analysing the sample three times under repeatability conditions, the standard uncertainty on the content determination of EA becomes:

$$u_{\bar{x}} = \sqrt{s_{r(EA)}^2 / 6 + s_{L(EA)}^2 / 2} = 0.660$$

The expanded uncertainty will now be $U_{\bar{x}} = 1.320$ and the company can report the mean result with the corresponding uncertainty interval as $\bar{x} \pm 1.3$. In this case, an improvement of the uncertainty on the result is seen, because the between-lab uncertainty also decreased.

3.2.3.2.4 Conclusion

The reproducibility of a new LC method for the analysis of erythromycin was examined in an interlaboratory study. All labs achieved a good selectivity allowing the content determination of EA and all identified related substances. The variance analysis, carried out on these results, showed a good reproducibility of the method. The method is suitable to replace the existing official method. The results of the validation can also be used by any lab that wants to make an uncertainty statement for this method.

3.2.3.3 Interlaboratory study related to timolol maleate by LC

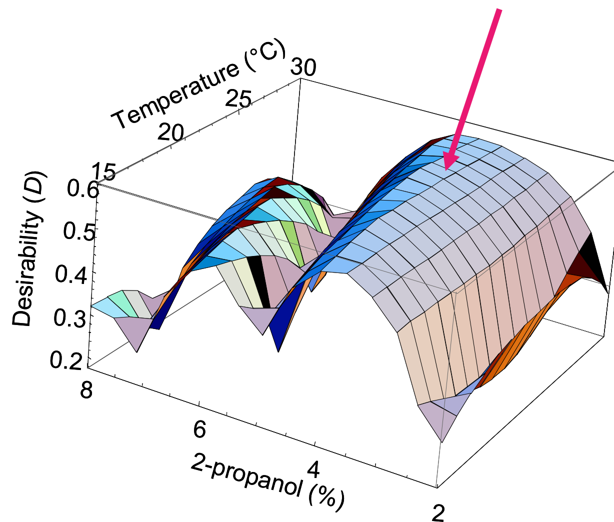
An interlaboratory study related to timolol maleate was proposed. It included the determination of R-timolol and other related substances by liquid chromatography in normal phase mode. The liquid chromatographic (LC) method is described in the 4th edition of the European Pharmacopoeia, Addendum 2004 for the purity enantiomeric testing of S-timolol maleate [20]. The aim of the development was firstly to adapt the LC method for a simultaneous separation of R-timolol and other related substances and secondly to validate the method for the determination of R-timolol and other related substances. S-timolol maleate is a β-adrenergic blocker used in the treatment of hypertension, arrhythmia, angina pectoris, for the prevention of myocardial infarctions and for the topical treatment of increasing intraocular pressure [14].

I. Optimisation of the LC method

During the optimisation of the LC method which was performed by applying a multivariate approach, three factors were selected and tested at three levels on the basis of preliminary studies namely the proportion of 2-propanol (X_1) and diethylamine (X_2) in the mobile phase and the column temperature (X_3). The linear and quadratic effects of these factors as well as the interactions effects between them were studied. The responses selected were the different retention times at upslope half-height, apex and downslope half-height of the peaks corresponding to R-timolol, isotimolol, S-timolol, dimorpholinothiadiazole and dimer maleate. These responses allowed to calculate the following defined responses of interest : the minimum resolution (Rs_{min}) between all peaks, the resolution values between

the peaks pairs R-timolol / S-timolol, S-timolol / dimorpholinothiadiazole and the retention time at apex for the last peak. These responses were estimated from the model and were selected in order to predict the LC conditions which meet the specifications for the enantiomeric separation as described in the monograph of timolol maleate (minimum value of 4.0 for the resolution between R-timolol / S-timolol) and also lead to an adequate quantisation of the late eluting compound in S-timolol maleate samples (minimum value of 4.0 for the resolution between R-timolol / dimorpholinothiadiazole) within an acceptable analysis time. The optimum values of 20 - 25 °C for the column temperature, 3.0 % and 0.1% of proportion of 2-propanol and diethylamine, respectively, in the mobile phase were deduced from the combination of the different desirability used for the selected responses (Figure 9) [53].

Figure 9: Global desirability



Under the optimal LC conditions, R-timolol and other related substances were sufficiently separated to allow their quantisation (Figure 10) [53].

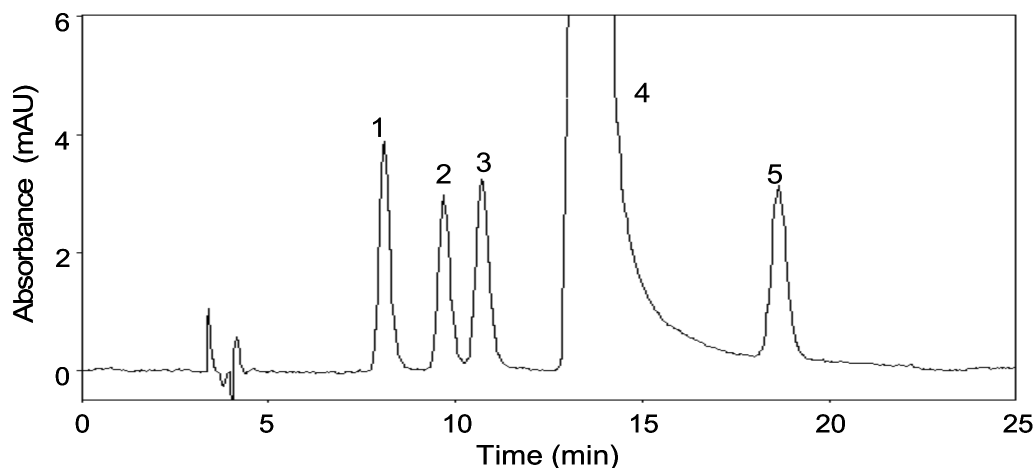


Figure 10 : Chromatogram of a spiked solution in the optimal conditions

Four steps are described in the Eurachem/ CITAC guide in order to assess the uncertainty : 1.- the specification of the measurand, 2.- the identification of the sources of uncertainty, 3.- simplifying by regrouping the sources covered by the known data, then quantifying the grouped components, the residual components to be converted in standard deviations, 4.- calculations of combined standard uncertainty and finally the expanded uncertainty. The optimisation step allowed the identification of the sources of uncertainty which were the proportion of 2-propanol in the mobile phase and the column temperature.

II. Validation of the LC method

The objective of the validation is to demonstrate the method selectivity, to determine the precision (repeatability and intermediate precision) and accuracy at each concentration level of the validation standards, to estimate the limit of quantisation and to verify the linearity of the relationship between the concentrations and the responses [40, 41]. The developed LC method was validated and the following criteria were tested : the method selectivity, which was assess by comparing the chromatogram of a blank solution with that of a 1% S-timolol maleate solution (Figure 11).

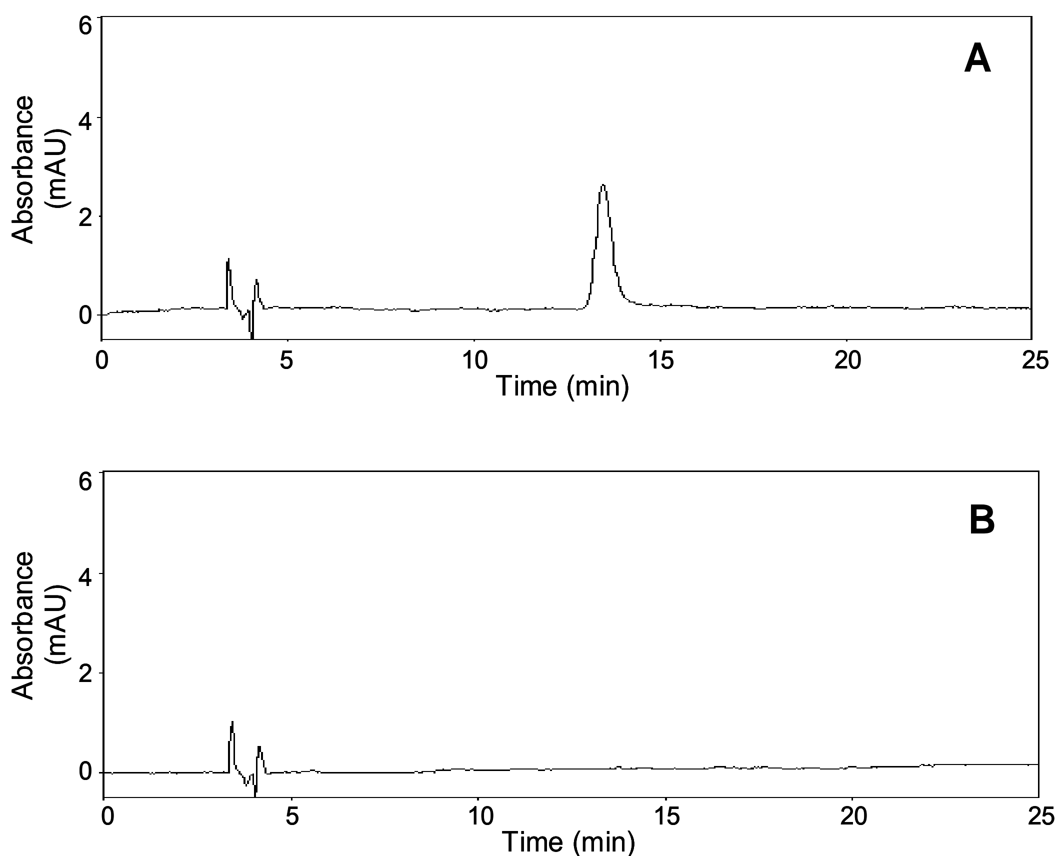


Figure 11 : Chromatogram of a solution containing 1.0 % of S-timolol maleate (A) and organic solvent (B)

As can be seen in the blank chromatogram, no interference was observed which demonstrates the selectivity of the LC method.

The response function as the relationship between the analytical response (peak area) and the analyte concentration was evaluated as well as the method precision. The later was evaluated at two levels: the repeatability and the inter day precision, the trueness allowed to estimate the systematic error or bias and the accuracy considering the total error of a measure (bias and variance). The limit of quantisation (LOQ) was also determined. Instead of defining the acceptance criteria based only on the observed bias and the variances, a novel approach considering the accuracy profile based on the 90% confidence interval of the total error (bias + standard deviation) (ß-expectancy at 90%) was applied [33]. This approach reflect more precisely the performances to be achieved in routine therefore minimise the risk of rejection.

According to a commission of « Analytical validation of quantisation methods» of the “Société Française des Sciences Techniques et Pharmaceutiques”, six steps are defined in order to establish the accuracy profile :

1. Fitting of a regression model from the calibration samples.
2. Back-calculation of the concentrations of validation samples according to the selected model.
3. Determination of the mean bias at each concentration level.
4. Calculation of two-sided 90 % confidence limits of the mean bias at each concentration level considering the standard deviation for intermediate precision.
5. Selection of the acceptance limits taking into account the intended use of the method.
6. Plotting of the accuracy profile, representing as a function of the concentration, the mean bias, the confidence intervals as well as the acceptance limits.

Therefore, the accuracy profile is a decision tool to accept or reject a method according to its attended use, to select a suitable regression model, to determine the quantisation limit and to select an interval of concentration for the assay.

The concentration range from 0.1 to 1.6 % of S-timolol maleate which corresponds to 1.5 $\mu g.mL^{-1}$ to 24.0 $\mu g.mL^{-1}$ was selected for the validation process. This range covers the maximum contents of both R-timolol and other related substances which are respectively 1.0% and 0.4% [20].

Samples were prepared by dilution from independent stock solutions. Tables 3 shows the different concentration levels of solutions prepared.

By using this novel approach, the linear regression model (Figure 12) as well as the interval of assay from the limit of quantisation (0.2%) to the superior concentration (1.6%) were selected.

Concentration level	Calibration samples	Validation samples
1.	0.1%	0.1%
2.	0.2%	0.2%
3.	0.4%	-
4.	0.8%	0.8%
5.	1.6%	1.6%
n = 3 independent replicates per level, for 3 days		
Total	15 samples / day	12 samples / day

Table 3 : Sample preparation for validation process

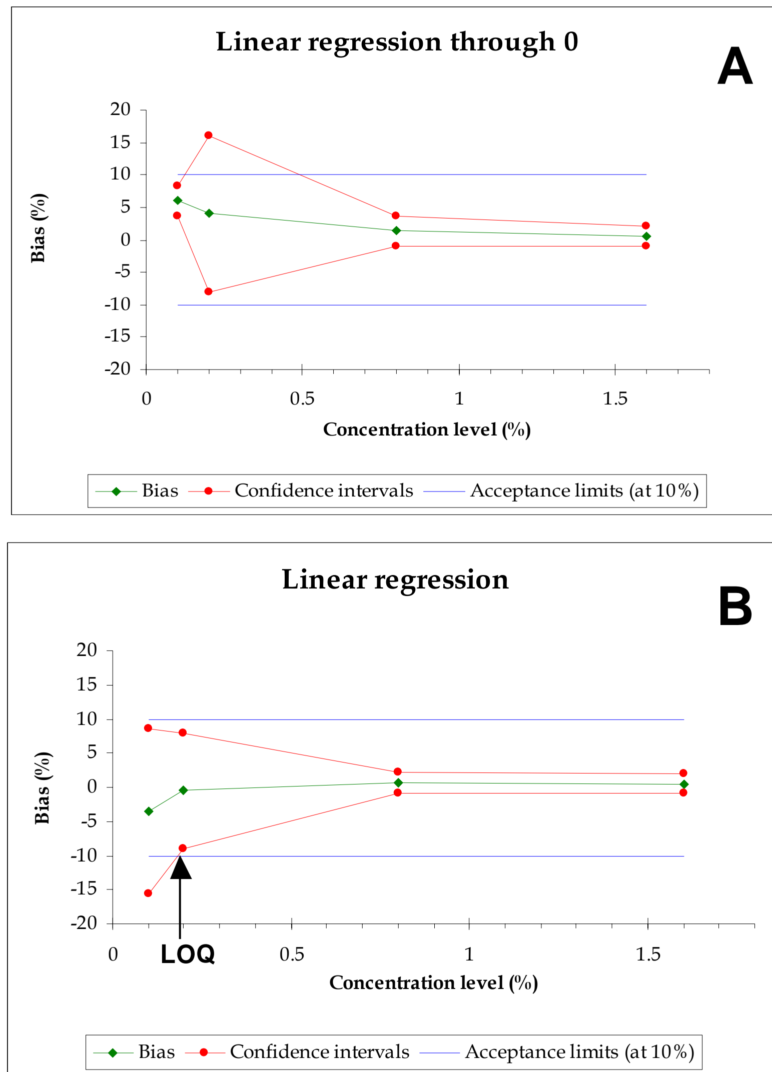


Figure 12 : Figures of the different accuracy profiles (**A** and **B**) for the selection of the suitable regression model, the determination of the LOQ and the assessment of method accuracy.

Thus, the precision, trueness and accuracy were studied at this interval. As can be seen in the Tables 4 and 5, excellent results were obtained with respect to precision, trueness and accuracy. This approach guarantees that 95% of the results obtained in routine analysis or interlaboratory study will be in the

acceptance limits which were fixed at 10% for the determination of the impurities. The estimates obtained from precision, trueness and accuracy are in great importance in the statistical treatment of interlaboratory study results.

	Regression coefficients		Coefficient of determination (r^2)
	Slopes	*y*-intercepts	
Series 1	75120	418.5	0.9999
Series 2	75473	821.6	0.9997
Series 2	74641	1442	0.9998

Table 4 : Linear regression in the concentration range from 0.2 to 1.6 %

Concentration level (%)	Trueness	Precision	
	Bias (%)	Repeatability (%)	Intermediate precision (%)
0.2 %	-0.5	0.4	2.6
0.8 %	0.6	0.8	0.8
1.6 %	0.5	0.7	0.7

Table 5 : Trueness and precision of the method

The quantisation limit of 0.2 % found relatively high was reduced to half by doubling the concentration of the test solution in order to rich a value of 0.1% as mentioned by the International Conference on Harmonisation [40]. The selectivity of the method was also demonstrated by comparing a chromatogram of the dissolution medium with a chromatogram of a 0.2 % of S-timolol maleate solution. Indeed, no interference was also observed in the retention times of the peaks corresponding to S-timolol, R-timolol and the other related impurities.

	Samples									
	8059	8060	11350	11351	11483	11484	11486	A5798	A5799	A5900
R-timolol (C.V. %)	< 0.20	< 0.20	< 0.20	0.20 (2.1)	0.25 (0.6)	0.30 (3.4)	0.20 (4.7)	< 0.20	0.25 (3.7)	0.50 (1.2)
Isotimolol	-	-	-	< 0.20	-	-	-	-	< 0.20	-
DMTDZ	-	-	-	-	-	-	-	-	-	-
Dimer maleate	-	-	-	-	-	-	-	-	< 0.20	-

Table 6 : Content in % of R-timolol maleate and other related substances in (S)-timolol maleate samples (n = 3 replicates; concentrations = 1.5 $mg.ml^{-1}$) with their coefficient of variation (%)

In order to evaluate the applicability of the developed LC method, ten S-timolol maleate samples from different sources were analysed. Results shown in Table 6 indicates that the maximum content of R-timolol was found to be 0.5% and none of the other compound were found to be present in concentrations higher than the LOQ that is 0.2%.

III. Robustness testing of the LC method related to timolol maleate and assessment of uncertainty from robustness testing results

At the last stage of a method development, it is important to demonstrate the reliability of the method reliability when deliberate variations of different parameters are introduced. The robustness is also important in order to identify the factors that could have an influence on the results and thus to anticipate the problems that may occur during the applicability of the analytical method such as interlaboratory studies, routine analysis… Another importance of the robustness is to study the relationship between the different factors and responses selected. If the measures are sensitive to the variations of analytical conditions, it will be convenient to maintain the conditions constant or to introduce a warning in the method description. [35, 36, 40, 48, 68].

The objective of this part was (*i*) to examine the potential sources of variability, the sensitivity of the method to operational factors, to determine the limits of control on the factors and evaluate the effect of the operating factors on the qualitative and quantitative factors. Another objective was (*ii)* to assess the uncertainty from the results obtained in the robustness study.

The operating conditions are described as following : a chiral stationary phase consisted of cellulose tris-3, 5 dimethylphenyl carbamate coated on silica (5 µm; dp) and packed in an analytical column of 250 mm x 4.6 mm, i.d. (ex. Chiralcel OD-H) was used. The mobile phase consisted to a mixture of n-hexane, 2-propanol and diethylamine (DEA) in the proportion of 965 / 35 / 1 (v / v / v). The flow rate was 1.0 $mL.min^{-1}$. The detection wavelength was 297 nm and the injection volume was 10 µL. Samples were dissolved and diluted in 2–propanol containing 1.0 % of DEA.

1) Selection of the factors, the responses and the experimental design

Among the 7 factors were selected, two were qualitative and 5 quantitative. Since the robustness study is simulated to an interlaboratory test, the factors to be selected had to reflect what could happen during this exercise, i.e. different LC equipments, columns of different ages, different volumes of solvents in the mobile phase. Moreover, the wavelength of a detector even well set could give different values. Other factors were also selected.

Factors	Units	Limits	Level (-1)	Level (+1)	Nominal
Type of LC equipment			Shimadzu	Agilent	Shimadzu
Age of the column			Old	New	New
Percentage of 2-propanol in the mobile phase	%	± 0.5	3.0	4.0	3.5
Percentage of diethylamine in the mobile phase	%	± 0.05	0.05	0.15	0.10
Flow rate of the mobile phase	ml/min.	± 0.1	0.9	1.1	1.0
Detection wavelength	nm	± 5	292	302	297
Column temperature	°C	-3 ; +7	20	30	23

Table 7 : Experimental domain

As can described in the Table 7, two qualitative factors or non-procedure related factors [36] comprised the LC instrument manufacturer and the age of the Chiralcel OD-H column. Two LC systems from two manufacturers, Shimadzu and Agilent, were used. The (-1) and (+1) levels were attributed to Shimadzu and Agilent, respectively, in an arbitrary way. These coded levels were also attributed to a new (+1) and an old (-1) column. A new column was defined as a column that has received less than 500 injections and an old one was defined as a column that has received more than 500 injections. Five quantitative factors or procedure-related factors [36] comprised the proportions of 2-propanol and of diethylamine (DEA) in the mobile phase, the flow rate of the mobile phase, the detection wavelength and the column temperature. Their limits and levels are indicated in Table 7.

Some qualitative and quantitative responses were envisaged. The qualitative responses concerned the chromatographic performance parameters, namely the values of the resolution between R-timolol and isotimolol peaks (Rs2-3, critical pair) and between R-timolol and S-timolol peaks (Rs2-4, enantiomeric pair) as well as the retention factor (k'), the tailing factor (T_f) and the theoretical plate number or column efficiency (N) calculated for the peak corresponding to R-timolol. The quantitative responses concerned the content of R-timolol in two samples (A1 and A2) of S-timolol maleate to be examined.

A two-level Plackett-Burman design [57 ,68] was elaborated for the seven factors by means of the JMP software version 3.2 for Windows. Eight experimental conditions were generated as indicated in Table 8. Figure 13 shows the set-up as it was followed for the execution of the analysis.

	Factors						
Exp.	Type of LC equipment	Age of the column	Proportion of 2-propanol in the mobile phase	Proportion of diethylamine in the mobile phase	Flow rate of the mobile phase	Detection wavelength	Column temperature
N° 1	+1	+1	+1	-1	+1	-1	-1
N° 2	-1	+1	+1	+1	-1	+1	-1
N° 3	-1	-1	+1	+1	+1	-1	+1
N° 4	+1	-1	-1	+1	+1	+1	-1
N° 5	-1	+1	-1	-1	+1	+1	+1
N° 6	+1	-1	+1	-1	-1	+1	+1
N° 7	+1	+1	-1	+1	-1	-1	+1
N° 8	-1	-1	-1	-1	-1	-1	-1

Table 8 : Experimental design

As can be seen in the figure 13, the solutions for SST-1 and SST-2 (reference solutions) were firstly run followed by those of the samples A1 and A2 (test solutions). Three replicates per sample were analysed. Before the execution of each experiment, the LC system was equilibrated with the corresponding LC mobile phase for minimum three hours. Due to the duration of the analytical procedure, two days were needed to realise one series. The overall analysis was repeated three times constituting three series.

The minimum number of repetition per replicate, per experiment and per series was set to 3 in order to identify the different sources of variability.

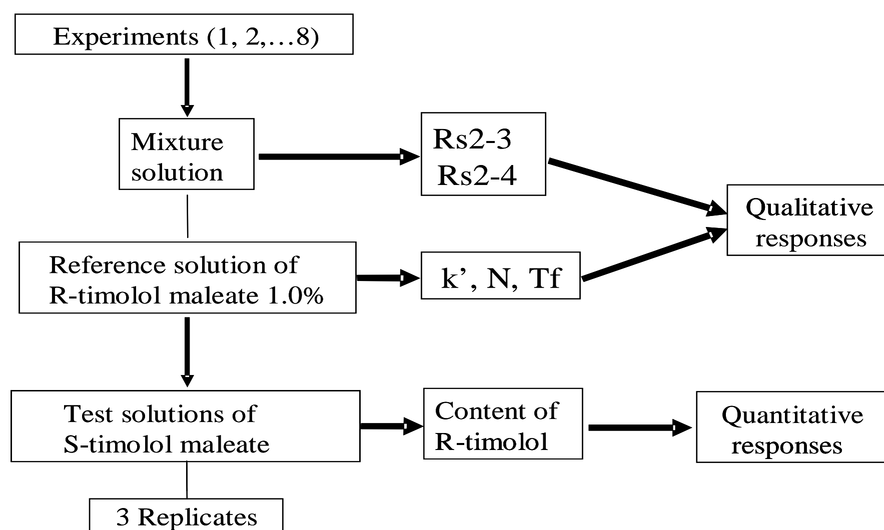


Figure 13 : Schematic design for the execution of the robustness study

2) Results of experiments for qualitative responses

As can be seen in Table 9, the mean values for the different chromatographic performance parameters obtained with the eight experiments were fluctuating between 0.98 and 1.65, 0.81 and 1.39, 1600 and 2400 for k', T_f, and N, respectively.

Experiments	Chromatographic performance parameters				
	k'	**T_f**	**N**	**Rs2-3 ± s.d.**	**Rs2-4 ± s.d.**
N° 1	1.65	1.25	2393	1.28 ± 0.03	5.82 ± 0,30
N° 2	1.11	1.17	1938	0.83 ± 0.11	4.31 ± 0,20
N° 3	1.11	1.32	2283	0.89 ± 0.10	3.98 ± 0,23
N° 4	1.50	1.15	1771	0.44 ± 0.12	3.92 ± 0,16
N° 5	1.46	1.39	1667	0.87 ± 0.07	4.94 ± 0,26
N° 6	1.20	0.81	1786	1.25 ± 0.02	4.93 ± 0,22
N° 7	0.98	1.32	1801	0.59 ± 0.02	3.45 ± 0,19
N° 8	1.57	1.17	1603	0.78 ± 0.04	4.46 ± 0,11
Genera l mean	1.32	1.20	1905	0.87 ± 0,29	4.48 ± 0,74

Table 9 : Chromatographic performance responses : capacity factor (k'), tailing factor (Tf) and plate number (N) of R-timolol, resolutions (Rs) of the critical and enantiomeric pairs (mean values) with their respective standard deviations (s.d.)

The resolution values (Rs2-3) of the critical pair resolution were varied between 0.4 and 1.3 while for the enantiomeric pair, only one resolution value (Rs2-4) was not above 4.0, which is the minimum value for this resolution as specified in the European Pharmacopoeia monograph of S-timolol maleate concerning the enantiomeric purity testing [20].

3) Results of experiments for quantitative responses

The two samples A1 and A2 were assayed and the content of R-timolol was determined by comparing the normalised area of R-timolol peak obtained by running the two S-timolol test solutions against the normalised area of R-timolol peak in the reference solution. As can be seen in Figure 14, the individual values of R-timolol content are fluctuating between 0.45 % and 0.65 % and between 0.25 % and 0.33 % in the samples A1 and A2, respectively.

These values were below 1.0 %, which is the maximum content of this chiral impurity as specified in the European Pharmacopoeia monograph of S-timolol maleate concerning the enantiomeric purity testing

[20]. These contents were around those observed under nominal conditions (0.50 % and 0.27 % for samples A1 and A2, respectively).

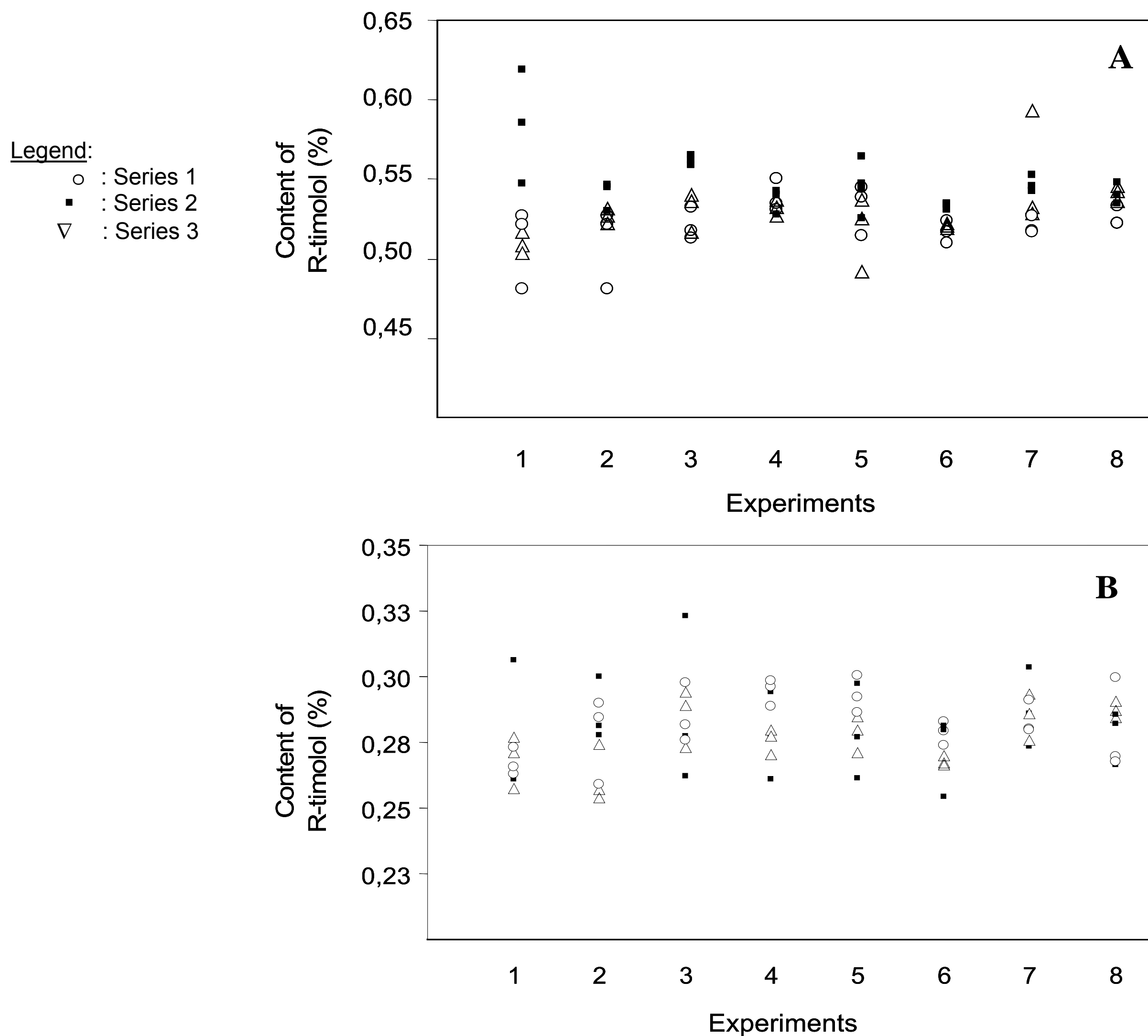


Figure 14 : Content (%) of R-timolol maleate quantified in A1 (**A**) and A2 (**B**) samples of S-timolol maleate

4) Statistical evaluation of the results

This step concerned on the one hand, the analysis of the effects of different factors on the selected responses and on the other hand, the study of the relationship between these responses and the factors in terms of prediction profile.

As can be seen in Table 10, the two qualitative factors have significant effects on the different qualitative responses at 5 % level (p-value < 0.05), except the type of LC equipment for which no significant effect was observed on Rs2-4.

Factors	Chromatographic performance parameters					Quantitative responses	
	k'	Tf	N	Rs 2-3	Rs 2-4	Sample A1	Sample A2
Type of LC equipment	**0.0204**	**< 0.0001**	**< 0.0001**	**0.0130**	0.1208	0.2761	0.2988
Age of the column	**< 0.0001**	**0.0015**	**0.0002**	**< 0.0001**	**< 0.0001**	0.5848	0.2262
Proportion of 2-propanol	**< 0.0001**	**0.0138**	0.0577	0.1317	**< 0.0001**	0.5848	0.8620
Proportion of diethylamine	**0.0249**	0.1380	**< 0.0001**	0.6287	0.4909	0.6619	0.4875
Flow rate	0.4804	0.0696	0.2295	0.7877	0.4353	0.5848	0.3860
Detection wavelength	0.2521	**0.0006**	0.1725	0.1446	0.2547	0.8268	0.2988
Column temperature	0.0021	**0.0154**	0.0925	**< 0.0001**	**< 0.0001**	0.1921	**0.0069**

Table 11 : Statistical significance (p-value) of the coefficients of the factors for each response. The significant values at 5% level are in bold.

All the quantitative factors have some significant effects on the qualitative responses at 5% level (p-value < 0.05) except the flow rate for which no effect was noticed and the wavelength for which it was observed a significant effect only on T_f. Indeed, 2-propanol has a significant effect on Rs2-4 as previously observed [53], on k' and on T_f. A significant effect of DEA on k' and on N was also observed. Moreover, the column temperature has a significant effect on T_f, on Rs2-3 and on Rs2-4.

As can be noticed in Figure 14 showing the influence of the different operating factors on the chromatographic performance criteria, when the type of LC equipment was changed (Agilent instead of Shimadzu), a decrease of T_f, N and Rs2-3 value was observed while the retention factor of R-timolol peak was increased. On the other hand, the Rs2-4 value was not significantly affected (see also Table 10).

Thus, the transfer of the developed method to any other LC equipment should not be a problem to reach the specific value of 4.0 for Rs2-4. When a new column was used, it was observed an increase of Rs2-3 and Rs2-4 values as well as of k' and N. However, a decrease of T_f was noticed. Consequently, the use of a new column is recommended for the improvement of these chromatographic performance criteria.

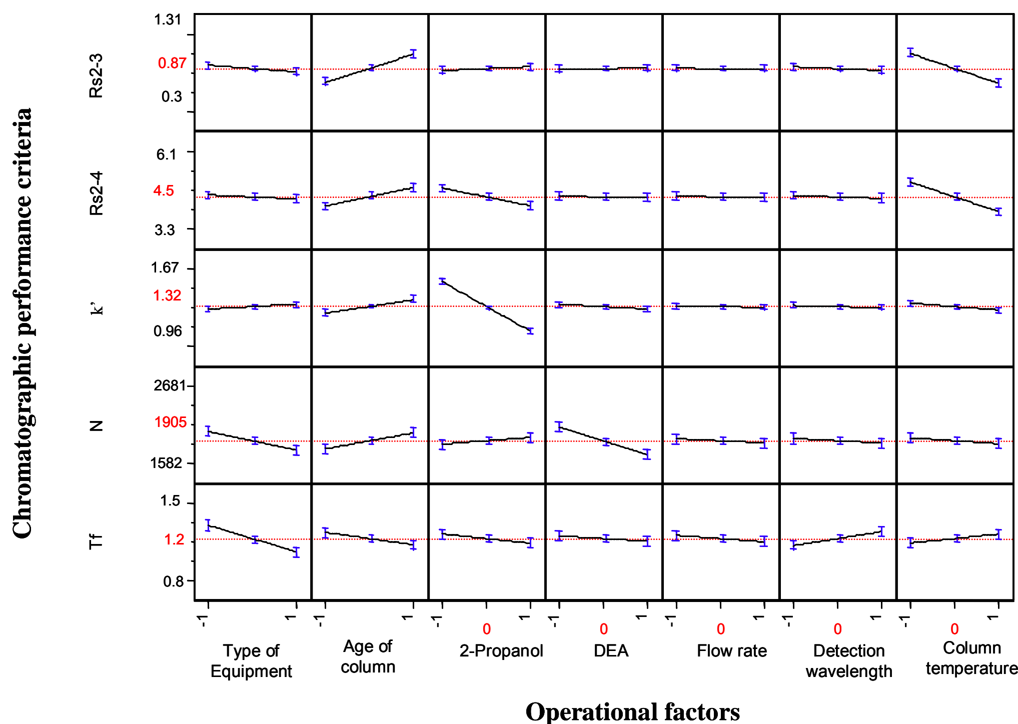


Figure 14 : Prediction profiles of the chromatographic performance criteria (qualitative responses)

Concerning the quantitative operating factors, it was noticed that when the proportion of 2-propanol increases in the mobile phase from 3.0% to 4.0%, the retention factor of R-timolol peak (k') decreases as well as Rs2-4 value. However, the column efficiency (N) was increased. On the contrary, a decrease of this qualitative response was observed when the proportion of DEA was increased in the mobile phase from 0.05% to 0.15%. Concerning the column temperature, it was noticed a decrease of the two Rs values and an increase of T_f when the temperature was increased from 20 to 30°C. For the wavelength, only T_f was increased when the value of this factor was varied from 292 to 302 nm.

Consequently, all the operating factors affecting significantly the different SST-1 specifications had been taken into consideration during the transfer of the LC method. These factors are the age of the OD-H column, the proportion of 2-propanol in the mobile phase and the temperature of the column. Therefore, it was important to ensure that the influence of these operating factors could or not influence the quantisation of R-timolol in the two samples. Table 10 indicates that no factor affected significantly the R-timolol content at 5 % level (p-value < 0.05) except the column temperature for which a significant effect was noticed only for sample A2.

This effect was not observed for sample A1 and was considered as a random effect. Therefore, even if some effects of the operating factors were observed on the chromatographic performance criteria, these operating factors had no significant influence on the content of R-timolol, which demonstrated the robustness of the LC procedure.

5) Uncertainty assessment

Since the eight experiments elaborated by means of the Plackett-Burman design can be assimilated to “laboratories” in the inter-laboratory study, the following steps could be performed in order to assess uncertainty by adapting the ISO 5725-2 guide [47]:

a) test of the variance homogeneity,
b) detection test of outliers,
c) calculation of the variance estimates and
d) estimation of the different uncertainty components.

(a) Test of the variance homogeneity

The scrutiny of results for consistency was performed by applying numerical tests to detect statistical outliers [28, 47].
For this purpose, the Cochran’s statistical test was applied. For a given set of *p* standard deviations, s_i, all computed from the same number (*n*) of replicate results, the Cochran test statistic, *C*, is calculated as:

$$C = \frac{s_{\max}^2}{\sum_{i=1}^{p} s_i^2} \qquad \text{(Eq. 19)}$$

where s_{max} is the highest standard deviation estimated for one sample. When the calculated value of *C* is larger than the 1% critical value, the standard deviation (variance) is considered to be an outlier, while it is considered to be a straggler when the *C* value is smaller than the 1% value but larger than the 5% one.

The calculated *C* values were 0.44 and 0.45 for the content of R-timolol in samples A1 and A2, respectively. They were found to be below the critical 5% ($C_{critical}$ = 0.52) and 1% ($C_{critical}$ = 0.62) values (degree of freedom (*dof*) = 8, 3). Consequently, the variances of the results obtained for the quantitative responses could be considered homogenous.

(b) Detection of outliers

Different Grubbs’tests [47] were applied in order to evaluate any mean results of the eight experiments which appears to differ unreasonably from the others. The following equations for the Grubbs’tests were used:

$$G_1 = \left(\bar{x} - x_1\right) / s \qquad \text{(Eq. 20)}$$

and

$$G_p = \left(x_p - \bar{x}\right) / s \qquad \text{(Eq. 21)}$$

with x_1 and x_p the smallest and the largest mean of the eight mean values respectively, $\bar{x}$ and s the general mean and the standard deviation, respectively, considering the eight experiments.

For the test of two largest outliers, the following equation was used:

$$G = s^2_{p-1,p} / s^2_0 \qquad \text{(Eq. 22)}$$

where $s_0^2 = \sum_{i=1}^{p}\left(x_i - \bar{x}\right)^2$, $s^2_{p-1,p} = \sum_{i=1}^{p-2}\left(x_i - \bar{x}_{p-1,p}\right)^2$ and $\bar{x}_{p-1,p} = \frac{1}{p-2}\sum_{i=1}^{p-2} x_i$, while for the test of two smallest outliers, the following equation was used :

$$G = s^2_{1,2} / s^2_0 \qquad \text{(Eq. 23)}$$

where $s^2_{1,2} = \sum_{i=3}^{p}\left(x_i - \bar{x}_{1,2}\right)^2$ and $\bar{x}_{1,2} = \frac{1}{p-2}\sum_{i=3}^{p} x_i$.

Since the G-values calculated for the different observed results (G_7 = 2.06 and G_6 = 1.11 for sample A1; G_3 = 1.07 and G_1 = 1.52 for sample A2) were found to be lower than the critical values of the tables at 5% ($G_{critical}$ = 2.27, *dof* = 8) and 1% ($G_{critical}$ = 2.13, *dof* = 8), it can be concluded that no outlier was detected in the eight mean values. Therefore, the results could be used for the analysis of variance.

(c) Calculation of the variance estimates

The variance estimates were calculated by means of the Anova table below. As can be seen in this Anova table, firstly the different mean squares were calculated, namely between-experiment mean square ($MS_{Experiments}$), between-series mean square (MS_{Series}) and between-replicate mean square ($MS_{Replicates}$) for the content of R-timolol in the two analyzed samples. Then, the estimates of the different variances could be calculated from these mean squares. These statistical calculations were performed by means of JMP software version 3.2 for Windows.

Sources of variability	**Mean squares**	**Estimated variances**
Experiments	$MS_{Experiments} = \frac{cg\sum\left(\bar{x}_i - \bar{x}\right)^2}{r-1}$	$MS_{Experiments} = s^2_{Replicates} + g\, s^2_{series} + c.g.s^2_{Experiments}$
Series	$MS_{Series} = \frac{g\sum\sum\left(\bar{x}_{ij} - \bar{x}_i\right)^2}{r(c-1)}$	$MS_{Series} = s^2_{Replicates} + g.s^2_{Series}$
Replicates	$MS_{Replicates} = \frac{\sum\sum\sum\left(\bar{x}_{ijk} - \bar{x}_{ij}\right)^2}{rc(g-1)}$	$MS_{Replicates} = s^2_{Replicates}$

As reported in Table 11, the between-experiment variance ($s^2_{Experiments}$) was found to be the lowest value for the R-timolol content in sample A1 while for the sample A2 the between-series variance (s^2_{Series}) could be considered as the lowest value.

Sources of variability	Rs2-3	Rs2-4	k'	N	T_f	Content of R-timolol in	
						Sample A1	**Sample A2**
Experiments	0.0813	0.5342	0.0628	77716	0.0298	$1.032\ 10^{-12}$	$0.17\ 10^{-4}$
Series	0.0056	0.0459	0.0017	12527	0.0056	$1.64\ 10^{-4}$	$0.03\ 10^{-4}$
Replicates	-	-	-	-	-	$3.31\ 10^{-4}$	$1.80\ 10^{-4}$
Total	0.0868	0.5801	0.0645	90243	0.0353	$4.95\ 10^{-4}$	$2.00\ 10^{-4}$

Table 11 : Estimation of the variance components

For the two samples, the highest variance was the replicates variance ($s^2_{Replicates}$). Consequently, the overall variability of the R-timolol content was mainly due to the replicates for which the contribution was estimated at 67 % and 90 % for samples A1 and A2, respectively. The different steps involved for sample preparation could contribute to the dispersion between the replicates such as the weighing, the sample dilution, ... Moreover, the balance, pipettes and volumetric flasks could influence the variability of the results. The sample rack of the equipments was not thermostated and consequently evaporation or contraction of the dissolution solvent might not be avoided which could explain the variability in the R-timolol content.

(d) Estimation of the different uncertainty components

The variances for repeatability (s^2_r) and reproducibility (s^2_R) were calculated using the following equations:

$$s^2_r = s^2_{Replicates} + s^2_{Series} \quad \text{(Eq. 24)}$$

$$s^2_R = s^2_r + s^2_{Experiments} = s^2_{Replicates} + s^2_{Series} + s^2_{Experiments} \quad \text{(Eq. 25)}$$

The s^2_r values obtained were $4.95\ 10^{-4}$ and $1.83\ 10^{-4}$ for R-timolol content in samples A1 and A2, respectively. Compared to the respective variances indicated in Table 11, the ratio between the reproducibility and repeatability variances is about 1, which means that the overall variability is mainly due to the repeatability conditions. The overall variability, explained as s^2_R and corresponding to the sum of the different variance components, allowed to calculate the standard uncertainty u_x using the equation (8) :

$$u_{\bar{x}} = s_R = \sqrt{s^2_{Replicates} + s^2_{Series} + s^2_{Experiments}} \quad \text{(Eq. 26)}$$

The standard uncertainty is u_x = s_R = 2.22 10^{-2} and 1.41 10^{-2} for R-timolol content in samples A1 and A2, respectively, and the expanded uncertainty becomes U_x = $2u_x$ = 4.44 10^{-2} and 2.82 10^{-2} for R-timolol content in samples A1 and A2, respectively, using a coverage factor k = 2.

Finally, for a single result x obtained running the experiments and considering the variance between experiments, the results for R-timolol content in samples A1 and A2 are x ± 0.044 and x ± 0.028, respectively.

The uncertainty values were found to be low and comparable for the two concentration levels of R-timolol.

Concerning the qualitative responses, it was observed that the overall variability was mainly due to the between-experiment variance (contribution > 85 %) and in lesser extent to the between-series variance (contribution < 15%). This high variability was expected since the operating conditions were changed in each experiment. However, except for the column efficiency (N), the uncertainty of the other qualitative responses was found to be less than 0.58, with an expanded uncertainty (U_x) less than 1.16. It was demonstrated that even if the operational conditions are modified within the experimental domain (Table 7), the expanded uncertainty remains small except for the column efficiency.

This study showed that the uncertainty values of both quantitative and qualitative responses were low in spite of the variations in the operating conditions. By assuming that a laboratory applies the optimized LC method in routine analysis, the results obtained for qualitative and quantitative responses will be very close to those observed in this study, except for the column efficiency. Nevertheless, the modifications in the operating conditions due to possible sources of errors should not be out of the tested experimental domain indicated in Table 7.

(e) Conclusion

The LC method developed for the determination of R-timolol and other related substances in S-timolol maleate samples remained unaffected by the variations of both qualitative and quantitative factors. Indeed, some significant effects of the factors were observed on the qualitative responses but did not influence significantly the content of R-timolol in the two samples analysed.

Since the 8 experiments of the Plackett-Burman design could be assimilated to laboratories in an interlaboratory exercise, different uncertainty components were evaluated using the data obtained from the robustness test. The observed uncertainty values were found to be relatively low. This robustness approach for the estimation of the uncertainty appears to be advantageous as no interlaboratory study is required.

IV. Interlaboratory study

The objective of the interlaboratory study related to timolol maleate was the evaluation of the sources of variability of the results and the assessment of the reproducibility uncertainty from interlaboratory test results.

Nine labs have accepted to participate to this interlaboratory study. A protocol was elaborated in order to allow the labs to perform identically the tests which comprised only the liquid chromatography aspects. A set-up (Figure 15) was made regrouping four parts namely the system suitability test 1 (SST-1) comprising the determination of the resolution values between R- / S-timolol and between isotimolol / R-timolol, SST-2 comprising the test of the signal-to-noise, the content of R-timolol in two different S-timolol maleate samples and the content of the other related substances in S-timolol maleate sample degraded at 190°C for 110 minutes. For the determination of the impurities content, two replicates per sample and two series per lab will be performed.

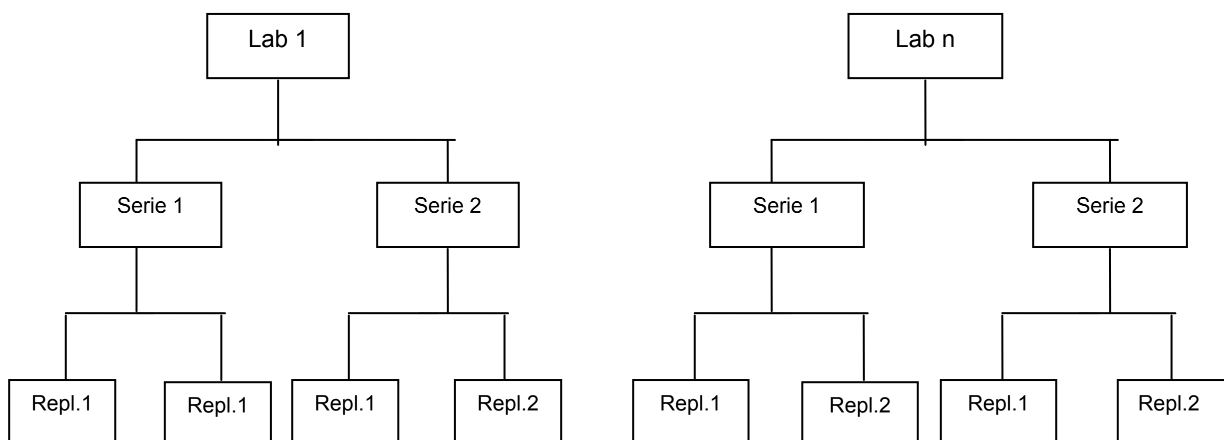


Figure 15 : Set-up of the interlaboratory study.

Before doing the determination of the impurities, the SST-1 and SST-2 should meet some requirements. These requirements are the minimum value of 4.0 and 1.3 to be reached for the resolution values between R- / S-timolol and isotimolol / R-timolol, respectively, a minimum value of 10 for the Signal-to-Noise ratio. The injection system of the LC equipment had to be qualified by obtaining a minimum value of 1.0 % for the relative standard deviation of peak area obtained after six injections of the same solution.

Firstly, a training round was performed according to ISO 5725-2 [47] in order to allow the labs to familiarize with the method and to evaluate the applicability of the protocol. The training comprised the execution of SST-1 and SST-2 and the determination of R-timolol content in two replicates of one S-timolol maleate sample, which means that the same sample is analysed two times, the same day and under repeatability conditions (Figure 16, r = 10 labs, c = 1 serie, g = 2 replicates).

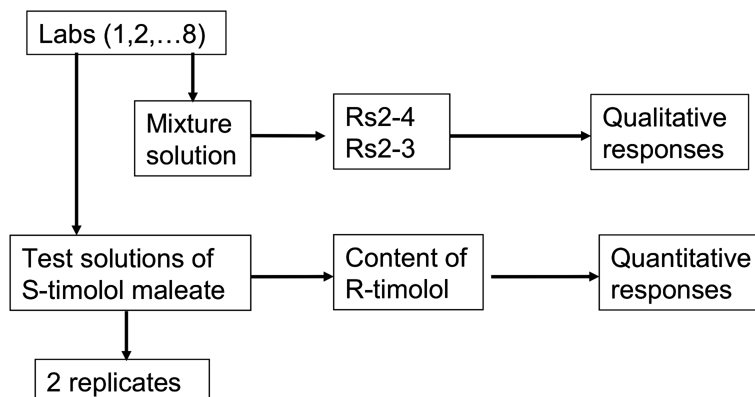


Figure 16 : Set-up of the training related to timolol maleate interlaboratory study.

The content of R-timolol maleate was determined according to the following set-up (Figure 17)

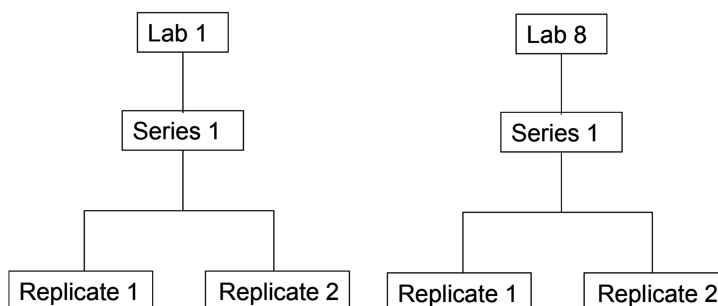


Figure 17 : Schematic design of the training execution for the determination of R-timolol content

As can be seen in Table 12, 3 types of LC equipment were used namely Agilent (4 labs), Waters (3 lab), LaChrom (1 lab) and Merck Hitachi (1 lab). Most of the different chiral columns used were new and were purchased from a unique manufacturer, Daicel Chemical Industries. Only three labs (Labs 1, 8 and 9) did not use the guard column which could not have any effect on the final results. After having equilibrated the LC system for minimum 3 hours, the SST-1 and SST-2 solutions were injected. The corresponding back pressures were below 50 bars which is the maximum value not to be exceeded since the chiral stationary phase is simply coated on silica particles. All the labs have signal-to-noise values above 10 which means that the detectors are suitable to quantify any peak which area is above the limit of quantisation that is 0.2%.

As can be seen in Table 13, all labs reached the minimum value of 4.0 between R- / S-timolol. Only one lab (Lab 2) modified the mobile phase composition in order to meet this requirement. Indeed, the proportion of 2-propanol was decreased to 2.0% in order to obtain a minimum value of 4.0, consequently, the analysis time was increased up to 22.9 min. (retention time of DMTDZ). The values for the critical peak pairs resolution was above 1.30 except for three labs (Labs 1, 7 and 9). The retention times of the different peak compounds were much closer except for Labs 2 and 4.

Labs	Type of LC equipment	Identification of the column (Type, Batch number, Manufacturer)	Age of the column	Use of guard column	Corresponding back pressure (bars)
1.	Agilent	Chiralcel OD-H, ODH0CE–BG 029 (Daicel Chemical Industries)	New	No	38
2.	Waters	Chiralcel OD-H, ODH0CE–CK 055 (Daicel Chemical Industries)	New	Yes	37
3.	Agilent	Chiralcel OD-H, ODH0CE–BC 081 (Daicel Chemical Industries)	New	Yes	39
4.	Agilent	Chiralcel OD-H, ODH0CE-CJ 072 (Daicel Chemical Industries)	New	Yes	34
5.	Merck Hitachi	Chiralcel OD-H, ODH0CE–CJ 068 (Daicel Chemical Industries)	New	Yes	32
6.	Agilent	Chiralcel OD-H, ODH0CE–CK 058 (Daicel Chemical Industries)	New	Yes	37
7.	LaChrom	Chiralcel OD-H, ODH0CE–CJ 011 (Daicel Chemical Industries)	New	Yes	28
8.	Waters	Chiralcel OD-H, ODH0CE–CJ 120 (Daicel Chemical Industries)	New	No	42
9.	Waters	Chiralcel OD-H, ODH0CE–IB 031 (Daicel Chemical Industries)	Old	No	35

Table 12 : Qualitative factors of the participating labs.

Labs	Change the mobile phase	S / N values	Rs2-4	Rs2-3	Retention times (min.)				
					DM	R-timolol	Isotimolol	S-timolol	DMTDZ
1	No	36.0	5.36	1.18	7.95	9.72	10.57	14.01	16.65
2	Yes	20.8	4.66	1.54	10.62	14.55	16.18	19.96	22.90
3	No	35.9	5.75	1.54	8.02	9.78	10.73	13.61	18.24
4	No	28.4	4.15	1.31	9.66	12.66	14.10	17.35	20.83
5	No	51.0	5.49	1.65	8.95	11.11	12.29	15.32	19.77
6	No	18.0	5.60	1.44	9.03	11.29	12.39	16.19	18.87
7	No	69.0	4.99	1.18	8.29	10.37	11.19	14.13	16.74
8	No	17.5	4.03	1.78	8.75	10.40	11.72	13.49	19.61
9	No	15.0	5.94	0.83	7.42	9.31	9.94	14.87	16.57
Mean values			5.11	1.38					
RSD			0.14 %	0.21 %					

Table 13 : Results of the qualitative responses

Concerning the determination of R-timolol maleate content, the repeatability of the LC system was checked and was acceptable for all labs (RSD below 1.0 %). The same sample was analysed by all labs

in two replicates. As indicated in Table 14, the content of R-timolol is fluctuating between 0.26 % and 0.38 %. However, the variability of the content of R-timolol in the two replicates were very low for all labs (RSD below 1.0%) except for Labs 5 and 8 (RSD above 5%) and in a lesser extent labs 6 7 and 9 (RSD between 1.0 and 2.0%). The general mean content of R-timolol maleate for all the nine labs was 0.279 % (RSD = 8.40 %).

Labs	**Repeatability (%) (n= 6 injections)**	**Content of R-timolol maleate**			
		Replicate 1	**Replicate 2**	**Mean**	**RSD (%)**
1	0.10	0.26 %	0.26 %	0.263 %	0.76
2	0.38	0.29 %	0.29 %	0.287 %	0.31
3	0.25	0.27 %	0.27 %	0.268 %	0.93
4	0.87	0.26 %	0.26 %	0.261 %	0.90
5	0.94	0.27 %	0.30 %	0.285 %	6.10
6	0.41	0.27 %	0.27 %	0.275 %	1.58
7	0.75	0.33 %	0.32 %	0.328 %	1.71
8	0.28	0.27 %	0.25 %	0.261 %	4.67
9	0.83	0.30 %	0.31 %	0.303 %	1.59 %
General mean content of R-timolol maleate (n = 9 labs)				**0.279 %**	
RSD (n = 9 labs)				**8.40 %**	

Table 14 : Results of the quantitative responses

The results of the training related to timolol interlaboratory study were compared to those of the robustness study. It concerned mainly the general mean content of R-timolol maleate in S-timolol maleate sample A since the samples had the same nominal declared content of R-timolol maleate. The comparison was performed by mean of the student test applying the formulas below and using the values in the table 15.

$$t_{calc}=\frac{\overline{x}_1-\overline{x}_2}{\sqrt{sp^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad \text{(Eq. 27)}$$

where sp^2 is the pooled variance calculated as following :

$$sp^2=\frac{n_1 \cdot s_1^2+n_2 \cdot s_2^2}{n_1+n_2-2} \quad \text{(Eq. 28)}$$

	Training test related to timolol interlaboratory study	Robustness test
Total number of replicates	$n_1 = 16$	$n_2 = 72$
General mean (%)	$\overline{X}_1 = 0.2807$ %	$\overline{X}_2 = 0.2797$ %
General variance	$s^2_1 = 5.21\ 10^{-4}$	$s^2_2 = 1.89\ 10^{-4}$

Table 15 : Comparison of the training and robustness test results

The homosedasticity test was performed in order to check variance homogeneity.
The *C* value calculated using the equation 19 (*C* = 0.585) was found to be below the C_{table} values at 5% and 1% (0.638 and 0.754, respectively) (*dof* = 9, 2). Consequently, the variances of the results obtained for the quantitative responses could be considered equal and the *t*-test can be used by mean of the Eq. 27 and 28.

The calculated *t* value was 0.24 and was found to be lower than the t_{table} value (t_{table} = 1.99; *dof* = 86, 0.975). Thus, the general means of the two studies can be considered equal. Consequently, the results presented are reliable and the labs are competent and experienced for the inter laboratory study.

3.2.3.4 Interlaboratory on a CE method for the analysis of metacyclin

3.2.3.4.1 Introduction

Previous studies examined uncertainty in LC methods. Two other interlaboratory studies were carried out in order to evaluate uncertainty in CE methods. As only one CE method is prescribed by European Pharmacopoeia and as this method is rather complex, it was decided to validate some new methods in an interlaboratory study. A CE method for the analysis of clindamycin was developed. However, this method uses a mixed micellar system and therefore was considered too complex to be the first CE method examined in an interlaboratory study. A more simple CE method by capillary zone electrophoresis was developed for the analysis of metacycline. Metacycline is a semisynthetic antibiotic obtained from oxytetracycline. Besides the main component metacycline, following impurities can be found: 6-epidoxycycline (6-EDOX), 4-epimetacycline (EMTC), 2-acetyl-2-decarboxamidometacycline (ADMTC) and doxycycline (DOX). The method uses an uncoated fused silica capillary (39 cm total length, 31 cm effective length, 50 µm ID) and a background electrolyte consisting of 160 mM sodium carbonate and 1 mM EDTA (pH 10.35) – methanol (87:13 v/v). The capillary temperature is maintained at 15 °C and the voltage at 12 kV, leading to a current of approximately 100 µA. The detection is done at

254 nm and the sample is injected hydrodynamically for 4 seconds. A typical electropherogram is shown in Figure 18.

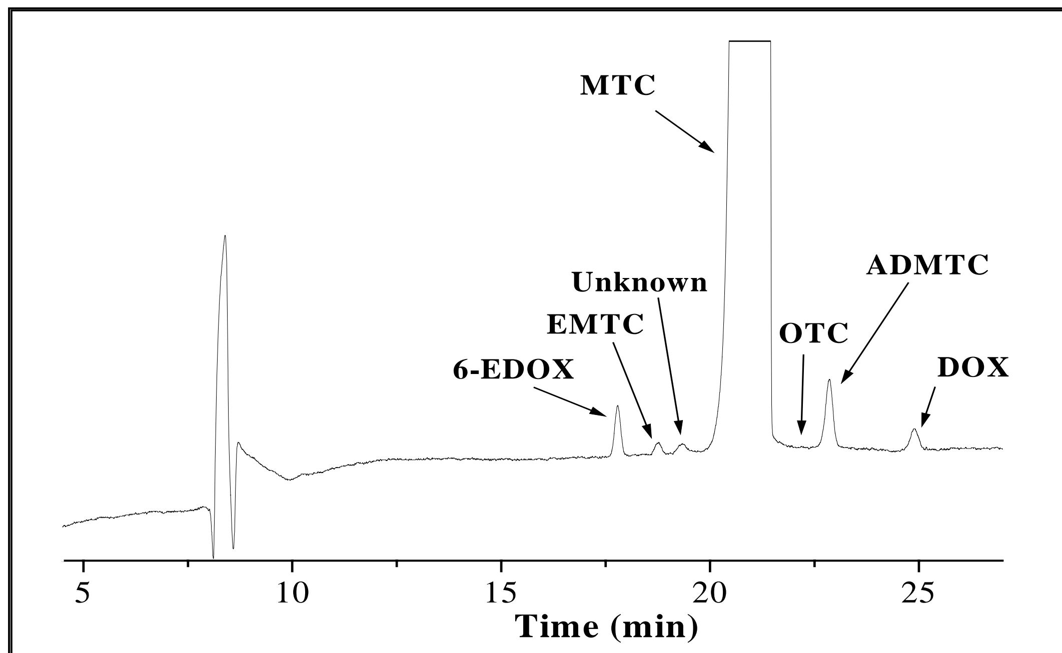


Figure 18. Typical electropherogram of a commercial sample of MTC.HCl (2.5 mg/ml).

The method was developed on a Spectraphoresis 1000 equipment (Thermo Separation Products, Fremont, CA, USA) and validated in one laboratory. Robustness, repeatability and linearity were examined. Epi-isochlortetracycline was used as an internal standard.

3.2.3.4.2 Set-up of the study

Ten labs participated in this interlaboratory study and followed an identical protocol. The set-up was the same as the one used in the penV study. Each lab performed two independent runs of experiments and in each run the same solution was measured twice (Figure 1, r = 10, c = 2, g = 2). One sample was analysed.

3.2.3.4.3 Results

Results of 7 labs were received. Many problems with the transfer of this method to other equipment were reported. The method was originally developed on a TSP 1000 equipment. During the interlaboratory study 4 labs used an Agilent equipment, 2 labs a Beckman equipment and 1 lab a TSP ultra equipment. An overview of the problems is given below.

I. Parameters according to the protocol

Longer migration times were observed on the Beckman equipment. The peak corresponding to MTC migrated at 33 min instead of 23 min, as observed on TSP and Agilent equipment. The internal standard (IS) migrates at 60 min instead of 40 min. The current on the Beckman equipment was found to be lower, more specifically 80 µA instead of 100 µA. These kind of problems were never observed on TSP and Agilent equipment. Similar migration times could be achieved on Beckman equipment by increasing the voltage from 12 kV to 17 kV. However the current at 17 kV was not the same for the two labs using the Beckman equipment: one lab mentioned a current of 150 µA and the other lab one of 128 µA. Selectivity changes for the pair EDOX-EMTC at 17 kV were observed as well. The cause of these observations is still unclear but the more efficient cooling in the Beckman equipment, which uses liquid cooling instead of air cooling, could play a role.

II. Selectivity problems

A poor selectivity between the main peak of MTC and the peak corresponding to ADMTC was observed on Agilent and Beckman equipment. This was due to tailing of the main peak. Some electropherograms even showed a splitting of the main peak. The reason of the tailing and splitting of the main peak, which was only observed when Agilent or Beckman equipment was used, remains unclear. When the same capillary and the same buffers of a system, giving typical electropherograms as shown in Figure 9 on a TSP equipment, were moved from the TSP equipment to a Beckman equipment, the main peak started to show tailing, the current dropped and the migration times became slower. This clearly shows that these problems are equipment related and do not depend on the capillary neither on the buffer.

One lab mentioned migration shifts of the main peak and IS during the analysis, but this was probably caused by methanol evaporation out of the vial. Buffer-vials with fresh buffer-solution should be used.

III. Poor repeatability

The variability on the results was high, even when an IS was used. However, the RSD on 6 injections, performed during the system suitability test, was in most cases below the maximum limit of 3 %. Content results for MTC are shown in Tables 16A-C. Contents were calculated in three different ways: (i) by area normalization, (ii) using a standard solution without IS and (iii) using a standard solution with IS.

The variability on the content can not be due to the injection alone, as the use of an IS does not improve the repeatability. It should be mentioned that the IS is not optimal, as it is migrating late in the electropherogram and some degradation of the IS can occur.

However it was the best IS which could be found. The integration of the tailing peak, with the ADMTC migrating on the tailing peak, is difficult and can contribute to the variability on the results as well.

Table 16A: content results of MTC by area normalization. S1, S2: series 1, series 2 ; M1, M2, measurement 1, measurement 2.

	S1		S2	
	M1	M2	M1	M2
Lab1:	95.99	96.00	93.16	92.57
Lab2:	95.82	95.88	93.71	98.71
Lab3:	97.70	97.49	97.69	97.58
Lab4:	96.85	97.19	97.10	----
Lab5:	94.64	94.69	94.40	94.64
Lab6:	97.57	98.31	----	----
Lab7:	93.95	94.05	94.44	93.94

Table 16B: content results of MTC using standard solution, without internal standard.

	S1		S2	
	M1	M2	M1	M2
Lab1:	90.10	92.50	92.86	102.02
Lab2:	82.81	80.99	108.10	---
Lab3:	91.81	97.41	90.57	123.35
Lab4:	81.81	82.02	87.09	87.24
Lab5:	88.39	89.33	92.42	87.66
Lab6:	91.49	82.34	----	----
Lab7:	86.93	97.23	76.87	85.61

Table 16C: content results of MTC using standard solution, with internal standard.

	S1		S2	
	M1	M2	M1	M2
Lab1:	87.83	85.75	89.09	87.14
Lab2:	84.87	81.09	----	----
Lab3:	94.72	87.24	88.26	94.47
Lab4:	87.15	87.46	89.49	85.77
Lab5:	86.98	88.36	91.41	88.58
Lab6:	88.15	89.50	----	----
Lab7:	89.02	93.33	93.69	93.66

3.2.3.4.4 Conclusion

Many problems with the transfer of the method to other CE equipment were observed during this interlaboratory study. Therefore the quantitative results can not be used in a variance analysis. The causes of the problems of transfer are still unclear and will be further investigated. A new interlaboratory study on a CE method should be performed, in order to investigate whether these problems are method related.

3.2.3.5 Interlaboratory on a LC method for the analysis of ketoprofen

In the purpose of the interlaboratory study related to capillary electrophoresis, another method was developed for the enantiomeric separation of ketoprofen. Several methods are described for the enantiomeric separation of ketoprofen by CE using cyclodextrins chiral selectors [7, 8, 22]. The objective of the optimisation the method was firstly to obtain an adequate resolution between the two enantiomers of ketoprofen in presence of sulfanilic acid used as an internal standard in an acceptable analysis time and secondly to validate the CE method for the enantiomeric purity testing.

3.2.3.5.1 Optimisation of the CE method

I. Electrophoretic conditions

The electrophoretic separations were carried out using an uncoated fused silica capillary (50µm ID) of 48.5 cm total length and 40.5 cm to the detector. A triethanolamine-phosphoric acid (TPA) buffer was made of 100 mM phosphoric acid adjusted to pH value with triethanolamine. The corresponding concentrations of the two cyclodextrins, trimethyl-β-cyclodextrin and sulfobutylether-β-cyclodextrin, sodium salt were added to this TPA buffer. A negative voltage was applied allowing the migration analytes at the anodic end of the capillary where the detection is performed spectrophotometrically at 214 nm. The solutions were introduced at the cathode by hydrodynamic injection mode for 4 seconds under 50 mbar of pressure. At the beginning of the working day or when using a new capillary, the conditioning was made successively with a 1N sodium hydroxide for 10 minutes at 60°C, a 0.1 N of sodium hydroxide for 10 minutes at 60°C and milli-Q water for 10 minutes at 30°C. Before each experiment, the capillary was conditioned with TPA buffer pH 2.5 followed b by TPA buffer pH 2.5 containing the CD’s.

II. Preliminary studies

Several preliminary studies were realised in order to evaluate the most significant on the electrophoretic criteria. These preliminary studies concerned the voltage, the temperature and the pH value evaluated in the range from –20 to –30 kV, from 20°C to 30°C and from 1.0 to 4.5, respectively. Finally, the value of pH was selected to 2.5, the applied voltage to –25 kV and the temperature of the capillary at 25°C. However, the factors that influenced significantly the electrophoretic criteria were the concentrations of the two CDs. Previous studies also oriented the selection of factors [1, 24].

III. Selection of the factors, the experimental design and the responses

On basis of the preliminary studies, two quantitative factors were selected namely the concentrations of TM-β-CD and SBE-β-CD in the TPA buffer. Their levels are indicated in Table 17.

Factors	**Low value (-1)**	**Central value (0)**	**High value (+1)**
Concentration of TM-β-CD (mM) [X_1]	10	20	30
Concentration of SBE-β-CD (mM) [X_2]	2	4	6

Table 17. Experimental domain

The selected responses were the three migration times (MT) at upslope half-height, apex and downslope half-height for each peak analyte as shown in Figure 20.

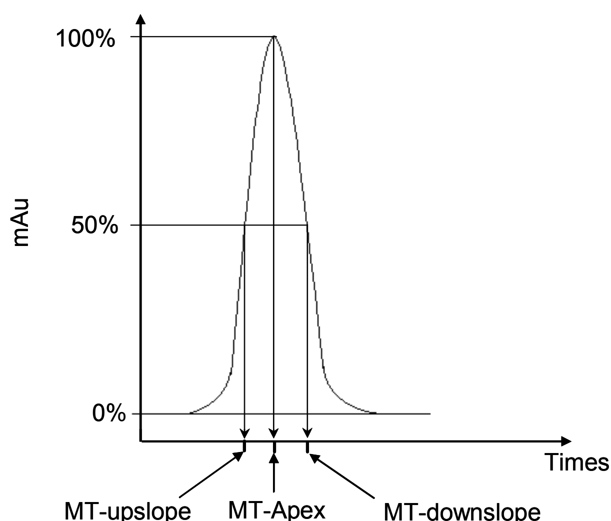


Figure 20. 3 migration times (MT) at upslope half-height, apex and downslope half-eight of a peak analyte

These migration times of the three compounds were modelled. The regression equation (equation 29) was used in order to evaluate the main effects (β_1 and β_2) the quadratic terms (β_{11} and β_{22}) and the first-order interactions (β_{12}) of the factors as envisaged in the model (Eq. 29).

$$MT_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2 + \varepsilon \qquad \text{(Eq. 29)}$$

A Central Composite design at centred face (CCF) was used to construct the experimental plan generating 10 experimental points as can be seen in Table 18 and to fit the different models.

Table 18. Experimental design

Experiments	TM-β-CD	SBE-β-CD
1	- 1	0
2	0	+ 1
3	0	0
4	-1	-1
5	+1	+1
6	+1	-1
7	+1	0
8	0	-1
9	0	0
10	-1	+1

IV. Evaluation of the effects of the factors

The quality of the fit of the model envisaged was assessed by R^2 which is the fraction of the response variation explained by the model and the plot of the residuals. All the value of R^2 obtained were above 0.99 for the migration times of the two enantiomers which means that each change observed in the MT of the two enantiomers is due to variation of the concentration of the two CDs in the TPA buffer. Concerning the sulfanilic acid, the R^2 value were around 0.85 meaning that the variation observed in the MTs is not totally due to the changes of CDs concentrations into the TPA buffer. This can be explained by the less affinity pattern of this non chiral compound towards the CDs chiral selector used in this study. Then, statistical calculations were performed by means of the JMP software version 3.2 for Windows in order to evaluate the influence of the factors on the selected response (Table 19).

	Migration times								
	Sulfanilic acid			R-ketoprofen			S-ketoprofen		
	Up	*Apex*	*Down*	*Up*	*Apex*	*Down*	*Up*	*Apex*	*Down*
Conc. TM	**0.01**	**0.01**	**0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Conc. SBE	0.66	0.67	0.67	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Conc. TM²	0.35	0.35	0.35	0.22	0.23	0.24	0.34	0.33	0.43
Conc. SBE²	0.91	0.90	0.88	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.02**
Conc. TM x Conc. SBE	0.31	0.31	0.31	**0.02**	**0.02**	**0.02**	**0.02**	**0.01**	**0.02**

Table 19. Statistical significance (p-value*) of the coefficients of the factors for the migration time at upslope half-height (Up), apex and downslope (Down) half-eight Experimental domain. The significant values at the 5% level are in bold-face type.

As can be seen in this table, the concentration of TM-β-CD has a significant effect on the migration times of the three analytes. However, the concentration of SBE-β-CD has a linear effect only on the migration times of the two enantiomers. Indeed, it was observed that when the TMCD concentration decreased, the migration times decreased while the opposite was observed for SBE-β-CD. No quadratic effect was observed on the migration times of the three compounds. An interaction effect of the two CDs was observed only on the migration times of the two enantiomers.

V. Prediction of the electrophoretic responses

The different migration times predicted from the model were used in order to estimate predicted electrophoretic responses namely the resolution (Rs), the minimum resolution (Rsmin) between peaks, the asymmetry factor (As), the separation efficiency and using the formulas below:

Resolution (European Pharmacopoeia)

$$Rs = 1.18 \times \left(\frac{MT_b - MT_a}{w_{(0.5)a} + w_{(0.5)b}} \right) \quad \text{(Eq. 30)}$$

$$\hat{A}s = \frac{w_{0.05}}{2d} \quad \text{(Eq. 31)}$$

$$\hat{R}smin = \min_{\substack{pics\ ij \\ i \neq j}} \left(\hat{R}s_{ij} \right) \quad \text{(Eq. 32)}$$

$$\hat{N} = 5.54 \left(\frac{\hat{M}_T}{w_{0.5}} \right)^2 \quad \text{(Eq. 33)}$$

Where MT_A and MT_B is the migration times of peaks A and B, respectively, $w_{(0.5)A}$ and $w_{(0.5)B}$ the width at half-height of peaks A and B, respectively, Rsmin the minimum resolution between peaks *i* and *j*, ..

The desirability function was defined for all the predicted electrophoretic responses. From these individual desirabilities, a combination of Rsmin (higher weighing), Max width, Min efficiency and Max symmetry as shown in Figure 21, allowed to obtain a robust area for the optimal concentrations (robust zone) which corresponds to 28 mM for TM-β-CD and 5 mM for SBE-β-CD. An electropherogram showing the separation obtained under these conditions is presented in Figure 22.

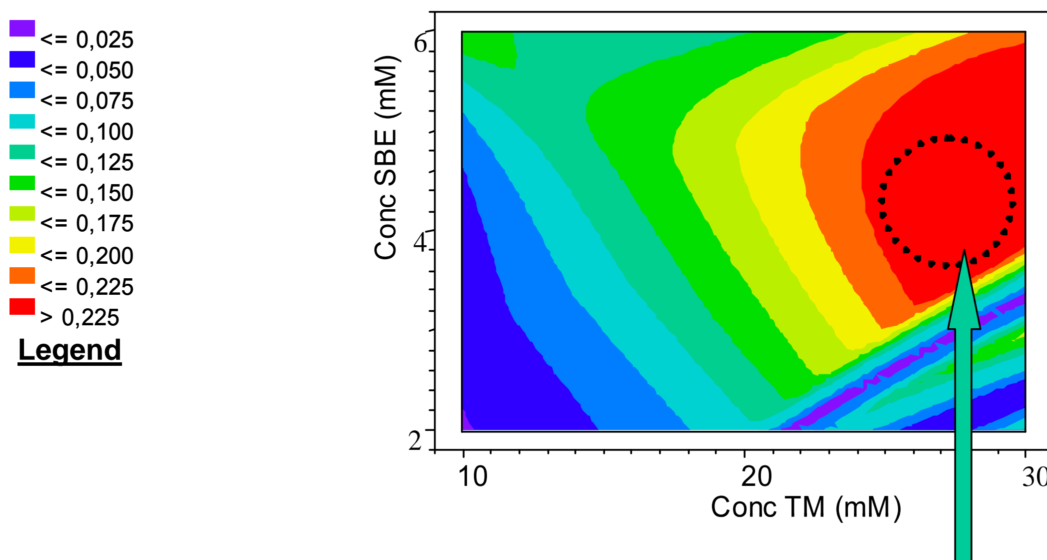


Figure 21. Combination of Rsmin (higher weighting) Max width, Min efficiency and Max symmetry

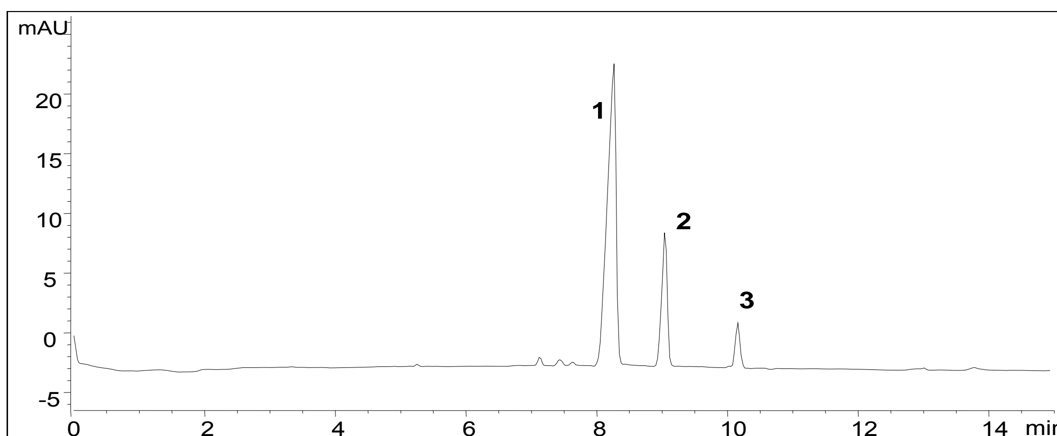


Figure 22. Enantionseparation of ketoprofen with an internal standard under the optimal conditions : TM-ß-CD : 28 mM and SBE-ß-CD : 5 mM. *Peaks and concentrations* : 1-. S-ketoprofen (50 µg.ml^{-1}), 2-. R-ketoprofen (25 µg.ml^{-1}) and 3-.sulfanilic acid (12.5 µg.ml^{-1}).

3.2.3.5.2 Validation of the CE method

The next step in the development of the CE method is the validation. This is still on performing. The method will be validated according to the novel approach considering the accuracy profile based on the 90% confidence interval of the total error (bias + standard deviation) (β-expectancy at 90%). This approach will be used to assess uncertainty from validation results [33].

3.2.3.6 Interlaboratory on a CE method for the analysis of amoxicillin

Preliminary experiments are taking place in order to perform an interlaboratory study on a CE method for the analysis of amoxicillin. The method uses Micellar Electrokinetic Chromatography (MECC) and is performed with an uncoated fused silica capillary (44 cm total length, 36 cm effective length, 50 µm ID). The background electrolyte consists of 70 mM sodium dihydrogen phosphate – 125 mM sodium dodecyl sulphate – 5 % acetonitril, adjusted to pH 6.0. The capillary temperature is maintained at 25 °C and a voltage of 15 kV is used. Amoxicillin is detected at 230 nm and the sample is injected hydrodynamically for 4 s. This method was originally developed on a TSP 1000 equipment and the transfer of this method to Beckman and Agilent equipment is currently under investigation.

4 DIFFUSION AND VALORISATION

4.1 GUIDELINE

The final aim of the project was to develop norms and guidelines for the practical estimation of uncertainty in analytical measurements obtained from methods applied in the chemical, pharmaceutical, agro-food and related industries. The following draft guideline, which is added as doc file, has been prepared :

"Guideline uncertainty.doc"

4.2 DIFFUSION THROUGH INTERNATIONAL CONGRESSES AND PUBLICATIONS

The results obtained in the frame of the project No. NM/12/23 were subject to large diffusion through international scientific congresses and publications in international scientific journals.

4.2.1 International congresses

R. D. Marini, P. Chiap, Ph. Hubert, J. Crommen
Separation and simultaneous determination of S-timolol maleate, its optical antipode and related substances by liquid chromatography using a cellulose based stationary phase.
10th Forum of the Belgian Society of Pharmaceutical Sciences (SBSP), Montréal (Canada) 24-27 May 2001 (Oral presentation).

R. D. Marini, P. Chiap, Ph. Hubert, J. Crommen.
Simultaneous LC determination of R-timolol and other closely related substances in timolol maleate using a cellulose based chiral stationaryphase.
XXVth International Symposium on High Performance Liquid Phase Separation and Related Techniques (HPLC 2001), Maastricht (Netherlands), 17 – 22 June 2001 (Poster).

R. D. Marini, P. Chiap, W. Dewe, B. Boulanger, Ph. Hubert, J. Crommen.
Development and validation of a simultaneous LC method for the determination of S-timolol maleate, its antipode and related substances using cellulose based chiral stationary phase.
VIIth International Symposium on Drug Analysis (Drug Analysis 2002), Bruges (Belgium), 21 – 25 April 2002 (Poster).

P. Dehouck, Y. Vander Heyden, J. Smeyers-Verbeke, D. L. Massart, J. Crommen, Ph. Hubert, R. D. Marini, O.S.N.M. Smeets, G. Decristoforo, W. Van de Wauw, J. De Beer, M. G. Quaglia, C. Stella, J.-L. Veuthey, O. Estevenon, A. Van Schepdael, E. Roets, J. Hoogmartens.
Determination of uncertainty in analytical measurements: Collaborative study of the analysis of phenoxymethylpenicillin.
VIIth International Symposium on Drug Analysis (Drug Analysis 2002), Bruges (Belgium), 21 – 25 April 2002 (Poster).

R. D. Marini, P. Chiap, B. Boulanger, Ph. Hubert, J. Crommen.
Assessment of uncertainty from robustness testing of a LC method developed for R-timolol and other closely related impurities (Oral presentation).
11th Forum of the Belgian Society of Pharmaceutical Sciences (SBSP 2003), Spa (Belgique), 08 – 09 May 2003.

P. Dehouck, Y. Vander Heyden, J. Smeyers-Verbeke, D. L. Massart, R. D. Marini, P. Chiap, Ph. Hubert, J. Crommen, W. Van de Wauw, J. De Beer, R. Cox, G. Mathieu, J. C. Reepmeyer, B. Voigt, O. Estevenon, A. Nicolas, A. Van Schepdael, E. Adams, J. Hoogmartens,
Determination of uncertainty in a liquid chromatographic method for erythromycin from Interlaboratory study results (Oral presentation).
11th Forum of the Belgian Society of Pharmaceutical Sciences (SBSP 2003), Spa (Belgique), 08 – 09 May 2003.

4.2.2 Publications in international scientific journals

E. Hund, D.L.Massart, J.Smeyers-Verbeke, Trends in Analytical Chemistry, 20 (8) 2001.

E. C. Gil, P. Dehouck, A. Van Schepdael, E. Roets, J. Hoogmartens,
Analysis of metacycline by capillary electrophoresis, Electrophoresis 22 (2001) 497-502.

P. Dehouck, A. Van Schepdael, E. Roets, J. Hoogmartens,
Analysis of clindamycin by micellar electrokinetic chromatography with a mixed micellar system,
Journal of Chromatography A 932 (2001) 145-152.

P. Dehouck, Y. Vander Heyden, J. Smeyers-Verbeke, D. L. Massart, J. Crommen, Ph. Hubert, R. D. Marini, O.S.N.M. Smeets, G. Decristoforo, W. Van de Wauw, J. De Beer, M. G. Quaglia, C. Stella, J.-L. Veuthey, O. Estevenon, A. Van Schepdael, E. Roets, J. Hoogmartens,
Determination of uncertainty in analytical measurements from collaborative study results on the analysis of a phenoxymethylpenicillin sample,
Analytica Chimica Acta 481 (2003) 261-272.

R. D. Marini, P. Chiap, W. Dewe, B. Boulanger, Ph. Hubert, J. Crommen,
Simultaneous LC determination of R-timolol and other closely related impurities in S-timolol maleate using a cellulose based chiral stationary phase : Application of an experimental design,

Journal of Separation Science 26, 809-817, (2003).

P. Dehouck, E. Roets, J. Hoogmartens,
Comparison of Two LC methods for the Analysis of Erythromycin,
Chromatographia, in press (2003).

P. Dehouck, Y. Vander Heyden, J. Smeyers-Verbeke, D. L. Massart, R. D. Marini, P. Chiap, Ph. Hubert, J. Crommen, W. Van de Wauw, J. De Beer, R. Cox, G. Mathieu, J. C. Reepmeyer, B. Voigt, O. Estevenon, A. Nicolas, A. Van Schepdael, E. Adams, J. Hoogmartens,
Interlaboratory study of a LC method for erythromycin: Determination of uncertainty,
Journal of Chromatography A, accepted for publication (2003).

R. D. Marini, P. Chiap, B. Boulanger, Ph. Hubert, J. Crommen,
Assessment of uncertainty from robustness testing of a LC method developed for R-timolol and other closely related impurities,
Submitted for publication in Analytica Chimica Acta (2003)

5 PERSPECTIVES AND OUTCOMES

The network of the project will try to obtain the acceptance of the draft guidelines in the ISO group and its application in the European Pharmacopoeia commission.

The network will try also to initiate a new commission of the “Société Française des Sciences Techniques” (SFSTP) on “Harmonization of quantitative analytical procedures”.

Finally, the network will try to define and assess a set of minimal guidelines for the determination of uncertainty in an analytical laboratory and its assessment in analytical reports.

International collaborations with pharmaceutical and chemical industries as well as University laboratories were made.

In a practical view point, a comparison of the uncertainty values estimated from the interlaboratory study and from the robustness testing will be done i.e. for to timolol maleate.

The network will try to perform another approach of uncertainty assessment according to the novel validation approach which considers the accuracy profile based on the 90% confidence interval of the total error (bias + standard deviation) (β-expectancy at 90%) [32].

6 REFERENCES

[1] Abushoffa A. M., Fillet M., Hubert Ph., Crommen J., J. Chromatogr. A, 948 (2002) 312 - 329.

[2] Analytical Methods Committee, Analyst 120 (1995) 2303 – 2308.

[3] Anglov T., Petersen I.M., Kristiansen J., Accred. Qual. Assur., 4 (1999), 504-510.

[4] Barwick V.J., J. Chromatogr. A, 849 (1999), 13-33.

[5] Barwick V.J. and Ellison S.L.R., VAM Project 3.2.1 "Development and Harmonisation of Measurement Uncertainty Principles", part d: Protocol for uncertainty evaluation from validation data, Version 5.1, January 2000 (http://www.lgc.co.uk/best/protocol.pdf).

[6] Barwick V.J. and Ellison S.L.R., VAM Project 3.2.1 "Development and Harmonisation of Measurement Uncertainty Principles", part d: Protocol for uncertainty evaluation from validation data, Version 5.1, January 2000 (http://www.caeal.ca/VAM%20uncertainty.pdf).

[7] Bechet I., Paques P., Fillet M., Hubert Ph., J. Crommen, Electrophoresis 15 (1994) 818 - 823.

[8] Bechet I., Fillet M., Hubert Ph., Crommen J., J. Pharm. Biomed. Anal. 13 (1995) 497-503.

[9] De Beer J.O., Baten P., Nsegyumva C., Smeyers-Verbecke J., J. Pharm. Biomed. Anal. 00 (2003) 1-45.

[10] BIPM, IEC, IFCC, ISO, IUPAP, OIML, Guide to the expression of uncertainty in measurement, ISO, Geneva, 1993.

[11] BCR report, EUR 18405 EN, Community Bureau of Reference, 1998.

[12] Caruso B.J., J. Manipulative Physiol. Ther, 19 (1996), 385-390.

[13] Chepkwony H.K., Dehouck P., Roets E., Hoogmartens J., Chromatographia 53 (2001) 159.

[14] Dickstein K., Hapnes R., Aarsland T., Am. J. Ophtalmol. 132 (2001) 626 – 631.

[15] Dybkaer R., Scand. J. Clin. Lab. Invest., 59 (1999), 579-583.

[16] Ellison S., Wegscheider W., Wiliams A., Anal. Chem., 69 (1997), 607A - 613A.

[17] EURACHEM /CITAC Guide "Quantifying uncertainty in analytical measurement", 2nd Edition, 2000.

[18] EURACHEM, Quantifying uncertainty in analytical measurement, 2nd Edition, EURACHEM, 2000.

[19] European Pharmacopoeia, 3rd ed., supplement 1998, European Pharmacopoeia, Strasbourg, 1998, pp. 504-505.

[20] European Pharmacopoeia, 4th Edition, Addendum 2002, Council of Europe, Strasbourg, France.

[21] Feinberg M., Spectra Analyse, 223 (2001), 19-24 [in French].

[22] Fillet M., Bechet I., Schomburg G., Hubert Ph., Crommen J., J. High Resolut. Chromatogr. 19 (1996) 669.

[23] Fillet M., Hubert Ph., Crommen J., Electrophoresis 18 (1997) 1013.

[24] Fillet M., Hubert Ph., Crommen J., J. Chromatogr. A, 875 (2000) 123 – 134.

[25] Frishman H.W., Furberg D.C., Friedwald T.W., N. Eng. J. Med. 310 (1984) 830 – 837.

[26] Guide for the expression of uncertainty in measurements, International Organization for Standardization, Geneva, 1993.

[27] Guide to the expression of uncertainty in measurement, 1st edition, 1993; corrected and reprinted 1995.

[28] Grubbs F. E., Technometrics 11 (1969) 1 – 21.

[29] "Harmonised guidelines for the use of recovery information in analytical measurement", technical report resulting from the symposium on harmonization of quality assurance systems for analytical laboratories, IUPAC, ISO, AOAC int. and EURACHEM, Orlando, FL, 4-5 September 1996.

[30] Hässelbarth W., Accred. Qual. Assur., 3 (1998), 418.

[31] Hibbert D.B., Jiang J. and Mulholland M.I., Anal. Chim. Acta, 443 (2001) 205-214.

[32] Horwitz W. and Albert R., Analyst, 122 (1997), 615-617.

[33] Hubert Ph., Nguyen-Huu J.J., Boulanger B., Chapuzet E., Chiap P., Cohen N., Compagnon P.A., Dewe W., Feinberg M., Lallier M., Laurentie M., Mercier N., Muzard M., Nivet C., Valat L., *STP Pharma Pratiques,* Vol. 13, N°3, mai/juin 2003.

[34] Hund E., Massart D.L. and Smeyers-Verbeke J., *Analytica Chimica Acta* 423 (2000) 145-165.

[35] Hund E., Vander Heyden Y., Haustein M., Massart D.L., Smeyers-Verbeke J., J. Chromatogr. A 874 (2000) 167 – 185.

[36] E. Hund, Y. Vander Heyden, M. Haustein, D.L. Massart, J. Smeyers-Verbeke, Anal. Chim. Acta 404 (2000) 257 – 271.

[37] Hund E., Massart D.L. and Smeyers-Verbeke J., Trends in analytical chemistry 20, 8 (2001) 394 - 406.

[38] Hund E., Massart D.L. and Smeyers-Verbeke J., Analytica Chimica Acta 480 (2003) 39 - 52.

[39] http://www.Vam.org.uk/terp/stats/mu/mu.html

[40] ICH Harmonised Tripartite Guideline prepared within the Third International; Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Text on Validation of Analytical Procedures : Methodology, 1994.

[41] ICH Topic Q2B Validation of analytical procedures: Methodology. Step 4, consensus guideline, the International conference on Harmonisation of Technical Requirements for registration of Pharmaceuticals for human use, 6 November 1996.

[42] International Standard, Accuracy (trueness and precision) of measurement methods and results, ISO 5725, Geneva, 1994.

[43] International Standard, Accuracy (trueness and precision) of measurement methods and results, Part 1: General principles and definitions, ISO 5725-1:1994, ISO, Geneva, 1994.

[44] International Standard, Accuracy (trueness and precision) of measurement methods and results, Part 3: Intermediate measures of the precision of a standard measurement method, ISO 5725-3:1994, ISO, Geneva, 1994.

[45] International Standard, Statistics – Vocabulary and symbols, ISO standard 3534-1:1993 (E/F), ISO, Geneva, 1993.

[46] ISO/IEC standard 17025, General requirements for the competence of testing and calibration laboratories, ISO, Geneva, 1999.

[47] ISO/IEC Standards 5725-2, Accuracy (trueness and precision) of measurement methods and results, 1994.

[48] Jimidar M., Van Ael W., De Smet M. and Cockaerts P., LC GC Vol. 15, 4 (2002) 230 – 242.

[49] Jülicher B., Gowik P., Uhlig S., Analyst, 124 (1999), 537 - 545.

[50] Küppers S., Accred. Qual. Assur., 3 (1998), 412 - 415.

[51] Kuttatharmmakul S., Massart D.L. and Smeyers-Verbeke J., Anal. Chim. Acta, 391(1999), 203 - 225.

[52] Linsinger T.P.J., Führer M., Kandler W. and Schuhmacher R., Analyst, 126 (2001), 211-216.

[53] Marini R. D., Chiap P., Boulanger B., Dewé W., Hubert Ph., Crommen J., *J. Sep. Science* 26, 809-817, (2003).

[54] Maroto A., Boqué R., Riu J. and Rius F.X., Anal. Chim. Acta, 440 (2001), 171-184.

[55] Maroto A., Riu J., Boqué R., Rius F.X., Anal. Chim. Acta 391 (1999) 173 – 185.

[56] NMKL, procedure No. 5, "Estimation and expression of measurement uncertainty in chemical analysis", Nordic Committee on Food Analysis, 1997.

[57] Plackett R.L., Burman J.P., Biometrika 33 (1946) 305 – 325.

[58] Quantifying Uncertainty in Analytical Measurement; Published on behalf of EURACHEM by the LGC, London, 1995.

[59] Quintana J., Martí I. and Ventura F., J. Chromatogr. A, 938 (2001), 3 – 13.

[60] Ramachandran R. and Rashmi, Analyst, 124 (1999), 1099 – 1103.

[61] Report on the FAO, IAEA, AOAC. int., IUPAC international workshop on "principles and practices of method validation", 4 to 6 November 1999, Budapest, Hungary.

[62] Rösslein M. and Wampfler B., "Evaluation of Uncertainty in Analytical Measurement" in Quality in Chemical Measurements – Training Concepts and Teaching Materials (Eds.: B. Neidhart, W. Wegscheider), Springer, Berlin Heidelberg, 2001, 43 - 63.

[63] B.N. Taylor and C.E. Kuyatt, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297, National Institute of Standards and Technology, Gaithersburg, MD, 1994.

[64] The International Collaborative Study Group, N. Eng. J. Med. 310 (1984) 9 – 15.

[65] United States Pharmacopeia, 24th Edition, National Formulary 18, United States Pharmacopeial Convention, 2002, Rockville, USA.

[66] Valcárcel M. and Ríos A., Trends Anal. Chem., 18 (1999), 68 – 75.

[67] Vander Heyden Y., Saevels J., Roets E., Hoogmartens J., Decolin D., Quaglia M.G., Van den Bossche W., Leemans R., Smeets O., van de Vaart F., Mason B., Taylor G.C., Underberg W.,

Bult A., Chiap P., Crommen J., De Beer J., Hansen S.H. and. Massart D.L, J. Chromatogr. A, 830 (1999), 3 – 28.

[68] Vander Heyden Y., Nijhuis A., Smeyers-Verbecke J., Vandeginste B.G.M., Massart D.L., J. Pharm. Biomed. Anal. 24 (2001) 723 – 753.

[69] Van der Veen A.M.H., Broos A.J.M., Alink A., Accred. Qual. Assur., 3 (1998), 462 – 467.

[70] Wardrop J., Ficker D., Franklin S., Gorski R. J., J. Pharm. Sci. 89 (2000) 1097 – 1105.

[71] Walsh M.C., Accred. Qual. Assur., 4 (1999), 365 – 368.

[72] Weng Naidong, de Beer J., Marcelis X., Derese P., McB J.H., Miller and Hoogmartens J., J. Pharm. Biomed. Anal., 10 (1992), 199 - 204.

[73] Wernimont G.T.: Use of statistics to develop and evaluate analytical methods, 1985, AOAC, Arlington, VA.

[74] Wood R., Nilsson A., Wallin H., Quality in the Food Analysis Laboratory, Royal Society of Chemistry Monographs, Cambridge, 1998.

[75] Yongxin Zhu, Moreno M.L., Porqueras E., Bourke E., Bruzzi A., Aletrari M, Kanari P., Partasidou D., Nienhuis J., Ferigo W., Robert J. L, McB J. H., Miller, Spieser J. M, Roets E., Hoogmartens J., J. Pharm. Biomed. Anal., 14 (1996), 1151 - 1156.

Publié en 2005 par la Politique scientifique fédérale

La responsabilité scientifique de ce rapport est assumée par les auteurs.

POLITIQUE
SCIENTIFIQUE FEDERALE

rue de la Science 8 ▪ B-1000 BRUXELLES
Tél. 02 238 34 11 ▪ Fax 02 230 59 12
www.belspo.be