UAB
Universitat Autònoma de Barcelona

Real Time Analysis of Medicines Supported by Quality Risk Management

Desirèe Moya Rodríguez

Universitat Autònoma
de Barcelona

Real Time Analysis of Medicines Supported by Quality Risk Management

Desirèe Moya Rodríguez

Ph.D. THESIS

Programa de Doctorat en Química

Director: Dr. Manel Alcalà Bernàrdez

Departament de Química

Facultat de Ciències

2023

Memòria presentada per **Desirèe Moya Rodríguez** aspirant al Grau de Doctor.

Desirèe Moya Rodríguez

Vist i plau:

Dr. Manel Alcalà Bernàrdez

Bellaterra, 16 de Maig de 2023.

Aquest projecte de recerca s'ha elaborat en col·laboració amb la indústria Laboratorios Menarini, S.A. en el marc del Pla de Doctorats Industrials, convocatòria DI 2019.

DOCTORATS INDUSTRIALS

Generalitat de Catalunya
Departament d'Empresa i Coneixement
Secretaria d'Universitats i Recerca

AGAUR Agència de Gestió d'Ajuts Universitaris i de Recerca

Responsable i tutor del projecte a la indústria:

Sr. Josep María González Bosch

GRUPO MENARINI

Badalona, 16 de Maig de 2023.

Agraïments

Aquesta tesi doctoral s'ha portat a terme gràcies a les següents ajudes institucionals i empreses:

Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) de la Secretaria d'Universitat i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya.

Ajuda de Doctorat Industria del projecte 2019DI009 amb títol "*Real Time Analysis of Medicines Supported by Quality Risk Management*".

Laboratorios MENARINI, S.A. Badalona, Barcelona.

Comenzaré por agradecer al Dr. Alcalà la propuesta de realizar el doctorado en el ámbito industrial junto al grupo de Quimiometría Aplicada de la Universitat Autònoma de Barcelona, haciendo una mención especial al Prof. Santiago Maspoch y al Prof. Jordi Coello. Junto a él, agradecer a Pilar Rajadel, al Sr. González y al Sr. Esquerdo la confianza depositada en mi al darme la oportunidad de iniciarme en esta aventura en Laboratorios Menarini, S.A.

En este trayecto he experimentado mis mejores y peores momentos. Miro hacia atrás, después de casi cuatro años, añadiendo los contratiempos de una pandemia vivida, y recuerdo cada elección, cada renuncia durante este tiempo. Familia y amigos han sido un pilar fundamental para llegar a este logro. Siento profundamente cada uno de los momentos en los que la situación ha podido con mi humor y ganas de disfrutar.

Mamà, papà, Noe; comprenderme, durante este tiempo, no ha sido tarea fácil para vosotros. Gracias por cuidarme y guiarme en cada etapa de mi vida. Vosotros sois mi ejemplo de superación, esfuerzo y constancia.

A ti Alexis, mi compañero de vida, gracias por tu paciencia y comprensión. Esta última etapa ha sido dura para ambos, pero estoy completamente segura que este será un año de éxitos. Juntos estamos formando un buen y bonito equipo.

Y qué decir de las personas que por casualidad encuentras en una aventura como esta. Compartir las experiencias de cada una de las personas que forman el grupo 'Kemometri' y 'Reines Analítiques' ha sido un placer. He tenido la gran suerte de conocer a una persona maravillosa, Judit. Una gran compañera de trabajo y batallas, pero sobre todo una grandísima AMIGA. Gracias por cada gesto de comprensión, apoyo y cariño durante este trayecto. Esto solo es el comienzo de todas las historias que vamos a compartir.

Mencionar que me siento plenamente agradecida y afortunada por haber coincidido con las personas que forman el grupo Menarini Badalona. Todos, desde el personal de limpieza, almacén, logística, mantenimiento, fabricación, ingeniería, operaciones, garantía, registros y mi querido personal de control de calidad; han aportado su mejor versión para ayudar cuando lo he necesitado. Y no podría dejar de mencionar al grupo Menarini Florencia por aportarme su conocimiento y duro trabajo durante este tiempo.

Y como dicen… todo esfuerzo lleva consigo una recompensa. Y SI, hoy puedo decir orgullosa que he logrado todo aquello que me propuse. ¡Qué gran satisfacción!

Gracias a todos por formar parte, *Desirèe*.

TABLE OF CONTENT

LIST OF ABBREVIATIONS AND ACRONYMS

ACS	Analytical Control Strategy
ANOVA	Analysis of Variance
API	Active Pharmaceutical Ingredient
AQbD	Analytical Quality by Design
ATP	Analytical Target Profile
CAA	Critical Analytical Attributes
cGMP	Current Good Manufacturing Practices
CI	Confidence Interval
CMA	Critical Material Attribute
CMP	Critical Method Parameter
CN	Critical Number
CPMP	Committee for Proprietary Medicinal Products
CPP	Critical Process Parameter
CQA	Critical Quality Attribute
CV	Cleaning Validation
CV%	Coefficient of Variation
DKP-T	Dexketoprofen Trometamol
DoE	Design of Experiment
EMA	European Medicines Agency
FDA	Food and Drug Administration
FMEA	Failure Mode Effects Analysis
GDO	Grades of Powder
HPLC	High-Performance Liquid Chromatography
ICH	International Conference on Harmonization
IR	Infrared
LAL	Lower Acceptance/Alert Limit
LMSA	Laboratorios Menarini, S.A.
LoD	Limit of Detection
LoQ	Limit of Quantification
LV	Latent Variable
LWL	Lower Warning Limit
MB	Manufacturing Batch
MD	Mahalanobis Distance

MODR	Method Operable Design Region
MSC	Multiple Scatter Correction
NIR	Near Infrared
NOC	Normal Operation Conditions
OOS	Out-of-Specification
OSD	Oral Solid Dosage Medication
PAT	Process Analytical Technology
PC	Principal Component
PCA	Principal Component Analysis
Ph. Eur.	European Pharmacopoeia
PLS	Partial Least Square (Regression)
QA	Quality Assurance
QC	Quality Control
QbD	Quality by Design
QTPP	Quality Target Product Profile
R	Correlation coefficient
R^2	Coefficient of determination
RID	Refractive Index Detector
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error of Cross-Validation
RMSEP	Root Mean Square Error of Prediction
RPN	Risk Priority Number
RSD	Relative Standard Deviation
RTRT	Real-Time Release Testing
SEL	Standard Error of Laboratory
SEP	Standard Error of Prediction
S-G	Savitzky-Golay
SM	Starting Material
SNV	Standard Vector Normalization
SOP	Standard Operation Procedure
SPC	Statistical Process Control
TOC	Total Organic Carbon
UAL	Upper Acceptance/Alert Limit
USP	United States Pharmacopoeia
UWL	Upper Warning Limit

SUMMARY

The complexity associated with the manufacturing processes of medicinal products and therefore the need for a better understanding of them to guarantee the quality, safety, and efficacy of manufactured products, requires the pharmaceutical industry to move towards advanced manufacturing. This shift in the manufacturing paradigm involves the combination of multivariate analysis along with advanced process analysis approaches to designing the quality of the product instead of evaluating the final quality of it at the end of the manufacturing process. To achieve this, the implementation of Near-Infrared Spectroscopy (from now on NIR), Raman spectroscopy, and Total Organic Carbon (from now on TOC) techniques has been proposed as an alternative to conventional analytical techniques.

Four lines of research have been proposed in this thesis considering the demands in each of the unit operations of the pharmaceutical manufacturing process, from the reception of the starting materials (from now on abbreviated as SM), the determination of the attributes in each unit operation, the validation of the cleaning of equipment after manufacturing, and optimizing the lifetime of a product:

1. The correct maintenance of an SM identification library has been guaranteed with the applicability of NIR spectroscopy. Moreover, the transferability between initial equivalent FT-NIR equipment from Bruker, following risk assessment and involving statistics and mathematical treatments, has been demonstrated.
    - Based on the workflow strategy for the identity of the set of the SM batches incoming in the sampling room, as a result of the detection of new sources of variability, the NIR spectral library based on the sub-cascading spectral strategy, involving 237 SM, has been properly calibrated and validated accordingly to the guideline requirements. The first level library has been calculated by spectral distance including a wide spectral range. The sub-cascading spectral libraries have been calculated involving the spectral distance in PCA (from now on Principal Component Analysis) space, more complex mathematical pre-treatments, and narrower spectral ranges have been necessary to guarantee the unequivocal qualification of those materials that generated confusion. As a result, six cascading levels have been created.
    - The transferability between NIR equipment from Bruker has been carried out considering the new version of the generated identification library. The implementation of the risk assessment based on the Failure Mode Effects Analysis (from now on FMEA) has been key to selecting, from the total of 237 SM, those considered critical for the evaluation of the equivalence between equipment. As a result, since no significant spectral differences have been found between the spectra acquired with both NIR

equipment for the materials categorized as critical, it is concluded with direct transferability.

2. Determination of critical quality attributes (from now on CQA) of solid-form medicines at different unit operations in the manufacturing process based on NIR and Raman methodologies has been carried out.
    - It has been possible to develop an identification and quantification procedure through offline NIR and Raman spectroscopy for a solid oral product in tablet form as a part of a corporate project. The experimental design and the measurement strategy have been crucial to guarantee the applicability of the procedures throughout the product's shelf-life.
    - It has been possible to develop an identification and quantification procedure through offline NIR spectroscopy for a product in solid oral form manufactured by direct mixing with a tendency to segregation, due to the different particle sizes of its components. To achieve this, the use of a specific measurement module with a sampler rotator, allows for an increase in the measurement surface, and obtaining more representative spectra of the sample has been crucial.
    - The correct performance of the determination of the CQAs identification, assay, and moisture content for a product in oral solid form has been demonstrated throughout the process life cycle, prior to the dosage unit operation, involving the NIR inline technology, allowing the real-time analysis. Moreover, the procedure for an offline NIR methodology has been successfully developed and validated for assay and content uniformity determination, if proceeding, for the batch product release.
    - Due to the satisfactory results, the offline NIR strategies developed are summited for registration as an alternative method for batch release in the authorized countries.
3. An alternative methodology has been developed and validated to correctly detect the presence of a selected contaminant after cleaning validation based on the TOC technique, following the analytical quality by design principle (from now on AQbD).
    - The ability of the TOC technique has been demonstrated to determine using a linear method, selective in the presence of organic carbon, accurate, precise, and robust, establishing the limits of detection and quantification, the presence of contaminants in cleaning water for the last rinse and swab samples, after pharmaceutical manufacturing, allowing its quantification from the calculated acceptable limits.
    - The transferability between TOC equipment from Sievers has been proven in terms of linearity, precision, and accuracy.

4. It has been possible to predict the shelf-life of specific pharmaceutical forms from historical long-term stability studies through the application of statistics, eliminating the presence of CQAs outside the specified limits.
    - For the liquid dosage form, it has been necessary to decrease the product shelf-life to 17 months removing the presence of assay-related out of specifications (from now on OOS).
    - For the solid dosage form, specific primary packaging materials should be selected to remove the presence of total impurities related to OOS at the authorized shelf-life (18 or 24 months depending on the packaging material).

OBJECTIVES

The main objective of this doctoral thesis is the development of feasible solutions based on process analytical technology (from now on PAT) using rapid analysis techniques involving the mathematical and statistical treatment of data with the use of chemometrics, to guarantee the quality, safety, and efficacy of medicinal products in each of the unit operations of the manufacturing process, allowing a better understanding of it through risk assessment, reducing the costs associated with the process.

NIR spectroscopy, Raman spectroscopy, and TOC are proposed as an alternative to the currently authorized and used conventional analytical techniques.

The following challenges to be achieved can be detailed below:

1. Guarantee the correct maintenance of an identification library that includes the set of SM that are received in the pharmaceutical industry with the applicability of NIR spectroscopy. In addition, to evaluate, following the risk assessment and involving statistical and mathematical treatments, the transferability strategy of the identification method between initial equivalent FT-NIR equipment.
2. Determination of CQAs for different solid forms of drugs at different unit operations of the manufacturing processes implementing offline and inline NIR and offline Raman methodologies.
3. Develop and validate an alternative methodology to correctly detect the presence of the selected contaminant by implementing TOC technology after cleaning validation, following the principle of AQbD.
4. Evaluate the optimization of the shelf-life of two drugs in different pharmaceutical forms through statistical applications, eliminating the presence of CQAs values OOS at the end of the authorized period.

PART I

INTRODUCTION

1. PHARMACEUTICAL INDUSTRY

This thesis has been developed within the framework of the industrial doctorate plan at Laboratorios Menarini, S.A (from now on LMSA) with the supervision of the Applied Chemometrics Research Group from the Universitat Autònoma of Barcelona (from now on GQA-UAB). The industrial doctorate plan is an initiative of the government of Catalonia managed with the support of the *Agencia de Gestió d'Ajuts Universitaris I de Recerca* (from now on AGAUR) and the *Consorci de Serveis Universitaris* of *Catalunya*. This thesis project was approved in 2019 as DI009.

Since 1972 the headquarters of Menarini Spain, one of the 18 manufacturing plants of the Menarini International Group, has been located in Badalona (Barcelona, Catalonia, Spain), where the manufacturing plant, the I+D+i center, and the central offices are located. LMSA, is one of the sites of the Menarini International Group, managing a plant of more than 11,000 m^2 for the manufacture of non-sterile drugs for oral and topical administration in solid (tablets, capsules, and sachets), semi-solid (cream, emulsion, gel, and ointment), and liquid forms (syrup). It is the only Menarini Group plant International with the necessary technology for manufacturing oral medicines in sachets, dispensed in a single dose. At the factory, located in the city of Badalona, more than 500 people have made possible the manufacture of 58 million units of medicines in 2021.

Menarini Spain comprises the following companies: A. Menarini Latin America, S.L.U (Lima, Perú); Menarini Consumer Healthcare, S.A.U. (Badalona, Catalonia, Spain); Guidotti Farma, S.L.U (Badalona, Catalonia, Spain); Retrain, S.A.U. (Badalona, Catalonia, Spain); and Laboratorios Fermon, S.L.U. (Badalona, Catalonia, Spain).

Currently, the company focuses on cardiovascular and respiratory risk, analgesia, male sexual health, and the digestive system. Other therapeutic areas such as oncology are recently investigated.

The Quality Unit integrates the departments of Quality Control (from now on QC), and Quality Assurance (from now on QA), and is responsible for directing the integrated management system as a fundamental support in all its activities, from I+D+i, through the selection of suppliers, the environment manufacturing, distribution, and promotion of its products.

The quality of medicines is a fundamental aspect of the Company ("*Quality comes first"*). This quality control process begins with the selection of suppliers of SM until the delivery end of medicines in pharmacies and hospitals, with special attention to the manufacturing process. The manufacturing plant that Menarini has in Badalona complies with the strictest standards of quality and is regularly inspected by the Spanish Health Authorities and those of other countries the marketing of medicines is required.

As a result of the objectives of continuous improvement, Menarini, for more than 20 years, has adopted measures to continue improving the quality of its processes and of the medicines it manufactures and markets. Proof of this is the adoption of improvements based on a greater scientific knowledge of its medicines and adequate risk management, as well as the adoption of product control methodologies during its processing, through the introduction of highly innovative and advanced development techniques involving the NIR and Raman Spectroscopies in the approach of PAT in collaboration with the GQA-UAB. In 2015, the start of the first industrial doctorate took place intending to promote the transfer of knowledge between the university and the company.

This change was motivated by two sources:

1. Regulatory authorities, mainly the European Medicines Agency (from now on EMA)[1] and the US Food and Drug Administration (from now on FDA)[2], emphasize that increasing efficiency, agility, flexibility, and the quality of the product achieved with the implementation of this new paradigm is key to ending the problem of drug supply (drug shortage)[3].
2. The new approach should make it possible to obtain high-quality products by reducing production costs. In addition, it must increase the flexibility of the processes so that they allow manufacturing different volumes of product and thus supply according to market demand. The new approach is in line with Industry 4.0, which is characterized by integrating digital technologies, including intelligent and autonomous systems through the use of the data generated involving machine learning. These will make it possible to optimize the value chain, reduce costs and save energy[4].

1.1.Manufacture of Pharmaceuticals

The largest volume of LMSA manufacturing is made up of the forms of oral solid medical products (from now on abbreviated as OSD) for single-dose sachets, tablets, or capsules.

The unit operations involved depend on the type of drug substances and the excipients, specific material properties assessing processability and meeting the administration requirements. The general goal is the obtention of a homogeneous blend, avoiding segregation effects[5].

Primary OSD manufacturing unit operations for the pharmaceutical products included in this thesis are detailed in Figure 1. 1. The process begins with the first unit operation adding the SM to the manufacturing equipment. Formulation products are accurately weighed and dosed into the proper process container. Then the blending is carried out to obtain a homogeneous distribution of the components of the formulation. This operation can be performed more than once throughout the manufacturing process. Depending on the final product characteristic other unit operations could be

involved. The components are then combined to create granules with suitable characteristics for each product through the use of granulation and mixing operations. The compression unit operation creates the final dosage form for the desired tablet product. If it is a question of obtaining a product in the form of a solid oral tablet, the final step in the OSD manufacturing process is the coating, where a layer is applied to cover the core previously manufactured by compression[5]. In the end, the packaging operations are involved to fill and label the bulk product, achieving the finished product.

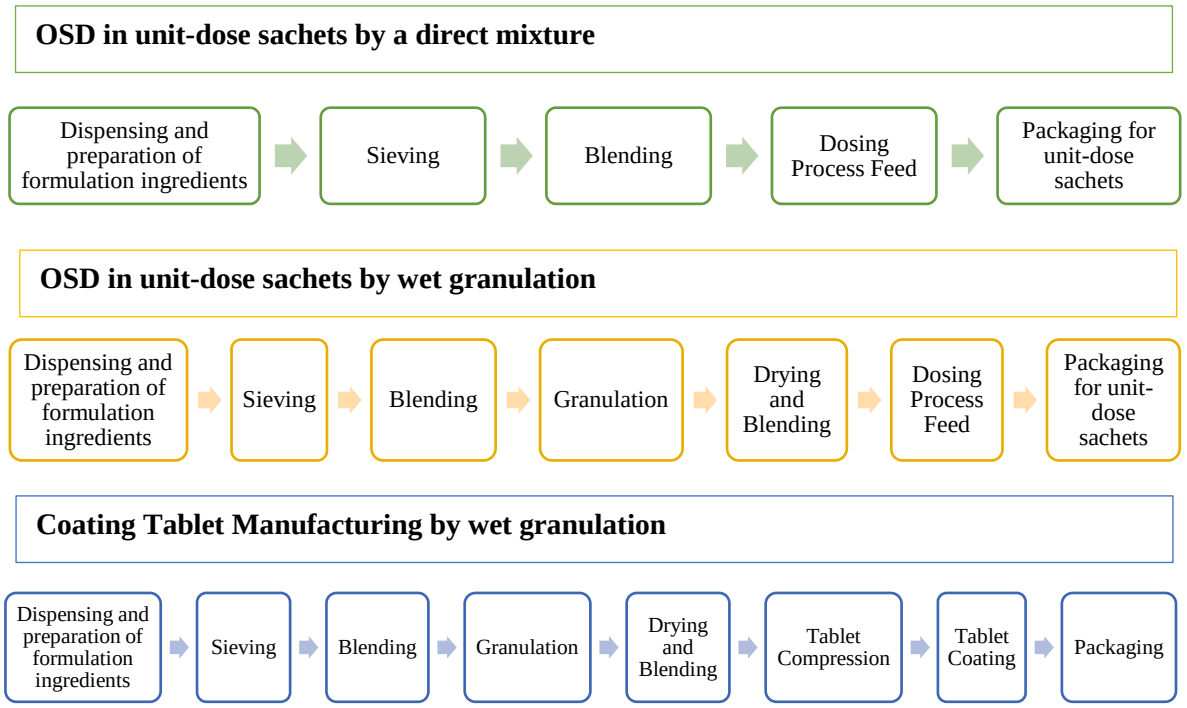


Figure 1. 1 Primary unit operations included in the manufacture of the studied OSD.

1.2.Product Lifecycle

The lifecycle management process provides a framework for defining the criteria for development of a procedure that meets the acceptance criteria[6].

Any procedure must be shown to be fit for its intended purpose before use including four main stages, training the implementation across all functions[7–10], from the purpose specification (defining the Quality Target Product Profile (from now on QTPP)), through design and development, the process qualification for the routine use, and the last stage including the on-going satisfactory performance (see Figure 1. 2):

- QTPP definition as a critical step including the desired properties and characteristics of a drug product ensuring its safety, efficacy, and quality.

- Procedure Design and Development: the criticalities of the process and the evaluated product are defined, including the Critical Process Parameters (from now on CPP), the Critical Materials Parameters (from now on CMP), and the CQA. The risk assessment could be involved to evaluate the criticalities. Moreover, the control strategy for the overall process and each unit operation is stabilized in this stage reducing the input variations and as a result the output impact[7].
- Process Qualification: process design evaluation is performed to ensure the reproducibility of the manufactured product based on current Good Manufacturing Practice (from now on cGMP) compliance[7,8,11], divided into two steps facility design and qualification of equipment and services, and qualification of process performance.
- Procedure Performance Verification: the state of control for the process during commercial manufacturing is ensured during this stage including the trend analysis for critical procedure parameters, and deviation management[7]. During the conduct of the procedure, performance verification is required continuous improvement to Stage 2. However, the possibility exists that a redevelopment or redesign is required in some instances necessitating a new Stage 1.

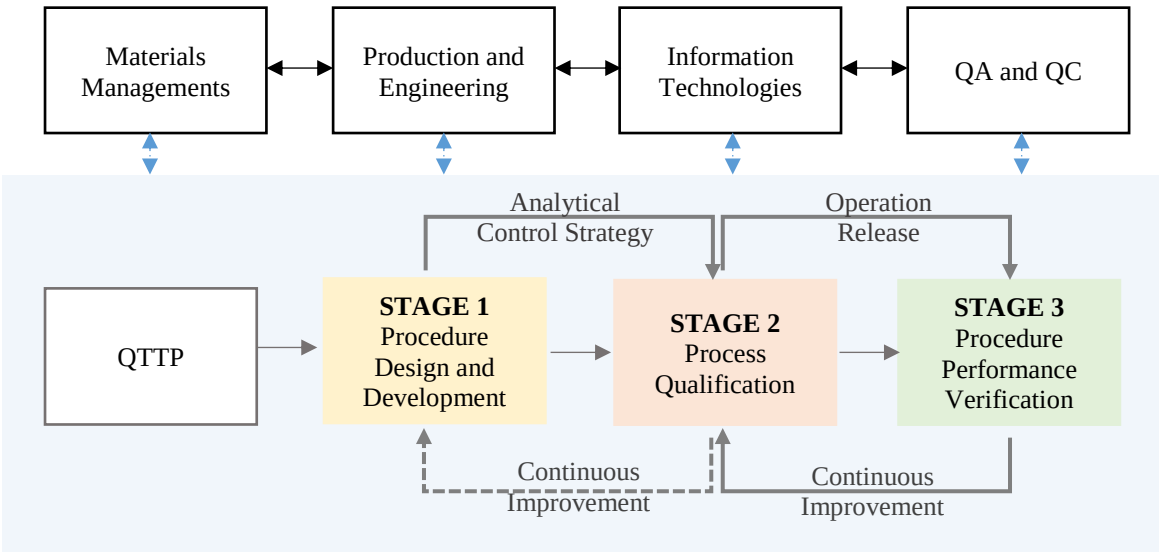


Figure 1. 2 Analytical Product Lifecycle.

2. QUALITY CONTROL AND STATISTICS FOR PHARMACEUTICAL PROCESSES EVALUATION

Before drug consumption, it must be analyzed and approved in verified pharmaceutical laboratories. QC, in a high-regulated industry, aims to verify the quality of the obtained product to proceed at the next pharmaceutical unit operation and as a result, release the manufacturing product by the regulations and specifications required for consumption[12,13].

Each phase requires following the literature established by the countries of authorization and consumption of the drug. The information and the activity, including failure laboratory or product results if proceed must be properly registered and identified[14].

The identification of assignable causes of variations in quality is based on the Statistical Process Control (from now on SPC) allowing to ensure a degree of uniformity in the manufactured and the obtention of the minimum industry standards for quality[15].

A process is presented as a system with a set of controllable and uncontrollable inputs (such as process variables, and SM properties provided by an external supplier or environmental factors, respectively) with a set of product outputs including measured quality characteristics (see Figure 2. 1).

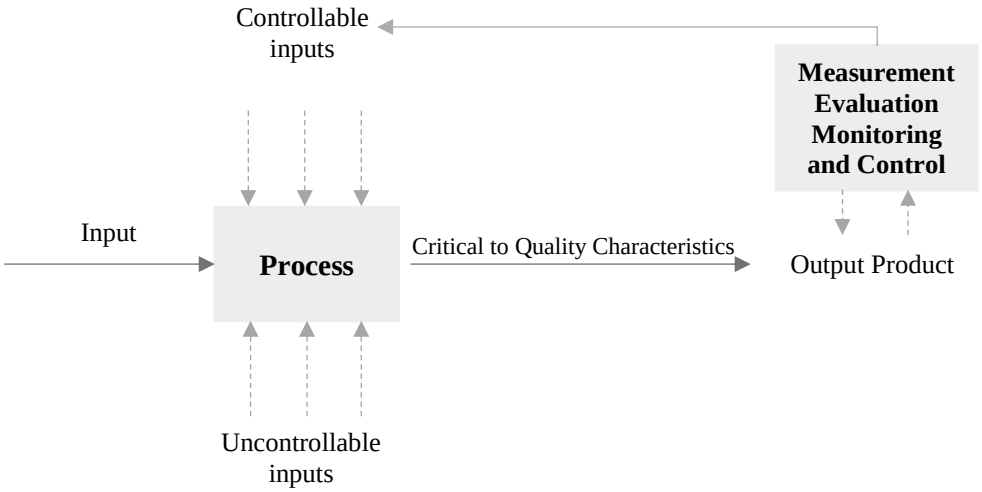


Figure 2. 1 Product evaluation quality scheme for a production process.

Cause-and-effect analysis, histograms, and scatter analysis such as graphical charts and graphs, can be included as SPC methods, helping to decode the statistics and data from quality control reports for an independent variable as a univariate analysis (describing in statistics a type of data based on a single characteristic or attribute)[15].

Nevertheless, the performance of the process may depend on more than one variable. The involvement of the univariate SPC includes several limitations, such as limited knowledge of the process, lack of context due to not considering external factors that may affect it, inability to detect complex patterns such as non-linear data relationships or interactions between variables, or the involved reactive approach due to detection of a runaway after the fact.

Hence, this is a traditional approach that assesses the quality of intermediate, bulk, and/or finished products at the end of a unit operation to reduce risk when moving on to the next. In addition, the process parameters are defined and fixed with their validation.

Based on what has been described, it is proposed to move towards a multivariate analysis by implementing the Multivariate Statistical Process Control (from now on MSPC) as a combination of advanced process analysis approaches to analyze complex processes[15].

3. ADVANCED PROCESS ANALYTICS

Today's pharmaceutical manufacturing needs to move forward to an advanced manufacturing process involving advanced analytical approaches. Hence, transitioning into Industry 3.0 being enabled to capture, and analyze high amounts of data (including computer and communication technologies enabling a higher degree of process and equipment automation) to finally transform into Industry 4.0 (to increase the agility, flexibility, efficiency, and quality of the industrial production processes[4] (see Figure 3. 1).

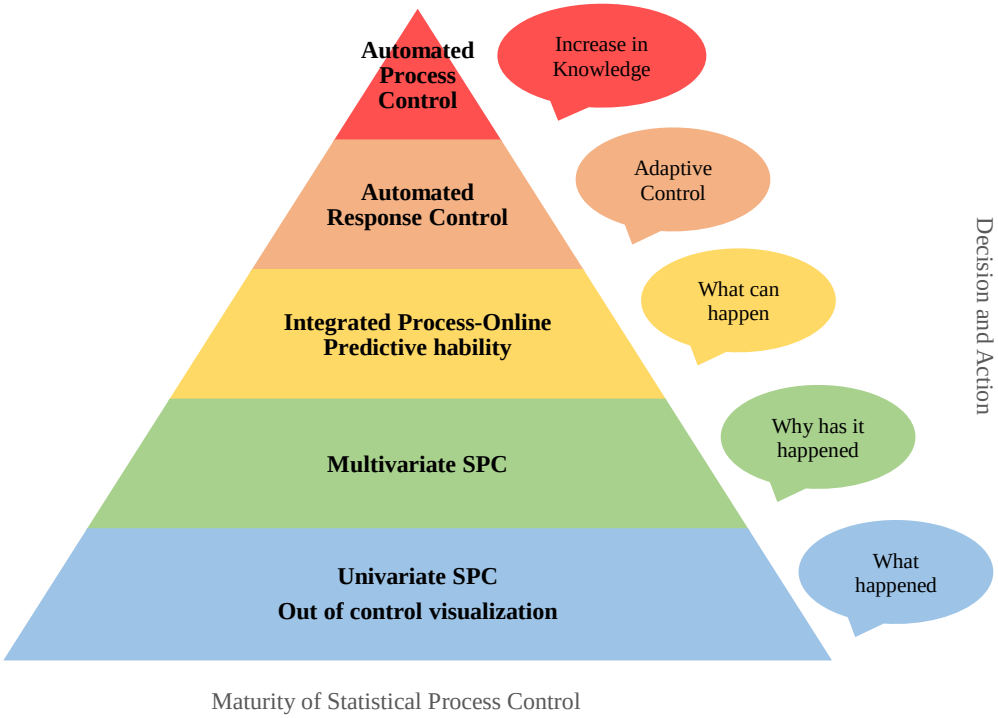


Figure 3. 1 Process Control evolution.

This new approach requires state-of-art technology in the development of processes including real-time, inline, and online sensors to generate actionable insight and predictions about process and unit operations through Quality by Design (QbD) and in the maximization of quality control with the PAT approach involving risk assessment as an improvement strategy in the different phases of drug production, to achieve real-time release.

The aim is the design of the quality of the product instead of the evaluation of the final quality of it at the end of the manufacturing process[4].

The following sections describe the QbD and PAT approaches.

3.1.Quality by Design

QbD is a systematic and iterative approach applied to product development and manufacturing focused on designing and controlling the product and process to achieve the desired quality[16,17], improving process efficiency, increasing regulatory compliance, and reducing time and costs[6,10,18] This approach includes the steps described in Table 3. 1. the same steps can be adapted to carry out the AQbD applied to analytical procedures.

Table 3. 1 QbD for Manufacturing Process.

QbD for a Manufacturing Process	
QTPP	Product quality attributes, including the intended use, dosage form, and physicochemical characteristics.
Criticalities Identification for process development	CQA for a product, CPP for process, and CMAs for formulation components. Including if needed risk assessment tools.
Process Design	Establish the design space where the product can be manufactured meeting the quality specifications.
Control Strategy about the Process	Monitoring plan for the manufacturing process; including a set of control, on process parameters and attributes, facilities, and equipment operating conditions, product specifications, and the associated methodologies.
Process Qualification	Demonstrate the meeting specifications for the evaluated manufacturing process.
Continuous Improvement	Monitoring and improving the product and the process to ensure quality based on regulatory requirements (continuous process verification)[10,18,19].

As a result, QbD aims to improve the knowledge of the process, being able to build a quality system based on the past to set future regulatory expectations.

3.2.Process Analytical Technology

The goal of the PAT approach is to improve understanding and control of the manufacturing processes[5]. It is a system of tools, techniques, and strategies such as advanced analytical and statistical methods used to monitor and control product quality in real-time throughout the manufacturing process.

The PAT development commonly involves a multidisciplinary approach including a combination of different tools[2,20]:

- Process analyzers: including univariate sensors able to measure speed, temperature, or weight, and multivariate techniques such as NIR and Raman Spectroscopies.
- Multivariate tools involving chemometrics to transform the acquired data into information.
- Process control tools for monitoring processes to keep them in the desired state of quality.
- Tools for continuous improvement based on risk assessment.

The set of these tools can be optimally applied to a single unitary operation or the entire manufacturing process.

As a result, the implementation of the PAT approach allows to reduce of the product cycle time that involves real-time measurements and the process control system, reducing the drug shortage problem and the possibility of real-time batch release[3].

4. RISK ASSESSMENT IN THE PHARMACEUTICAL INDUSTRY

The Quality Risk Management principles should be applied through all the stages of the analytical lifecycle. Risk management is a regulatory requirement of good manufacturing practices[12,19,21,22] that pharmaceutical companies should include as part of a risk-based approach throughout the product life cycle to optimize their benefit-risk balance. Quality risk assessment in the pharmaceutical industry must be based on scientific knowledge of the associated product and linked to patient protection. In the context of analytical procedures, the pharmaceutical product is the reportable result, and the risks are evaluated about the impact on their quality.

The risk management process should include the main steps summarized in Figure 4. 1[19]. The first step is the definition of the problem or the related risk question, described as risk identification. Then the estimation of the associated risk and the hazards are analyzed. The next step includes the risk evaluation involving risk criteria. Then a decision to accept or reduce the evaluated risk is established in the risk control step. Finally, the risk review should establish monitoring of the included risks. Hence, the application of the mentioned procedure interactively and continuously will allow the evaluation, control, communication, and review of the risks to the quality of medicines and related processes[12,18,21].

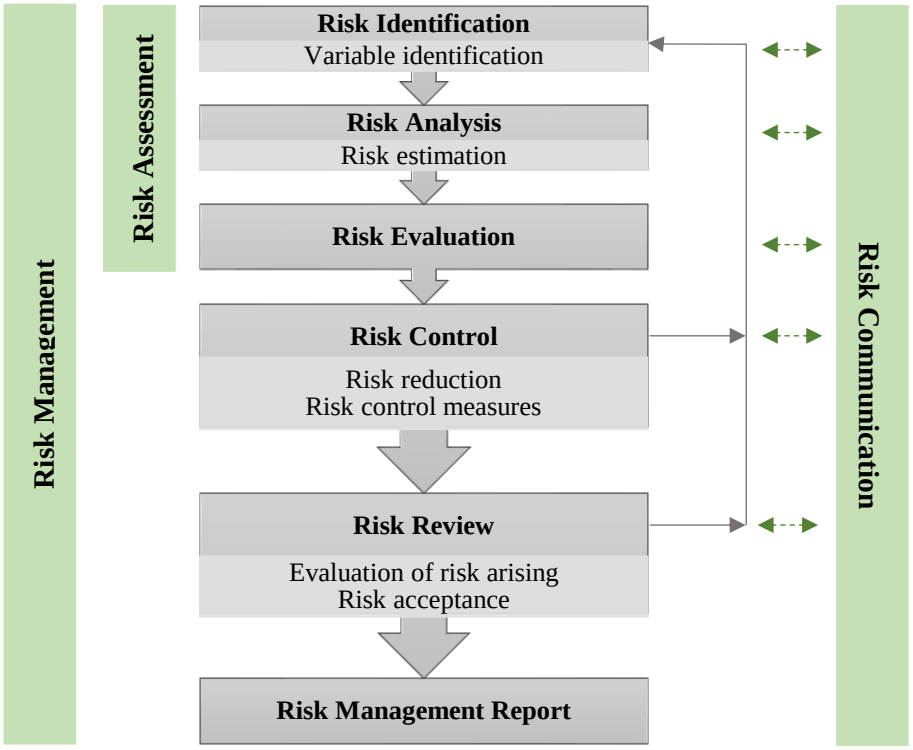


Figure 4. 1 Quality Risk Management Process Flow.

Different recognized risk management tools can be used[19]. The degree of rigor and formality applied must be a consequence of the complexity and/or criticality of the situation being analyzed, considering the principles described in the cGMP guide[12,21]. An overview of risk management tools involved in this thesis is included in this section.

4.1.Ishikawa Diagram

The Cause and Effect, Fishbone or Ishikawa Diagram was created in the 1940s by Kaoru Ishikawa. This risk management tool is commonly involved to detect and categorize potential causes of a studied product or quality problems[19].

It is typically created by drawing a horizontal line representing the problem or issue being investigated and then drawing several diagonal lines branching off from the central line. See Figure 4. 2.

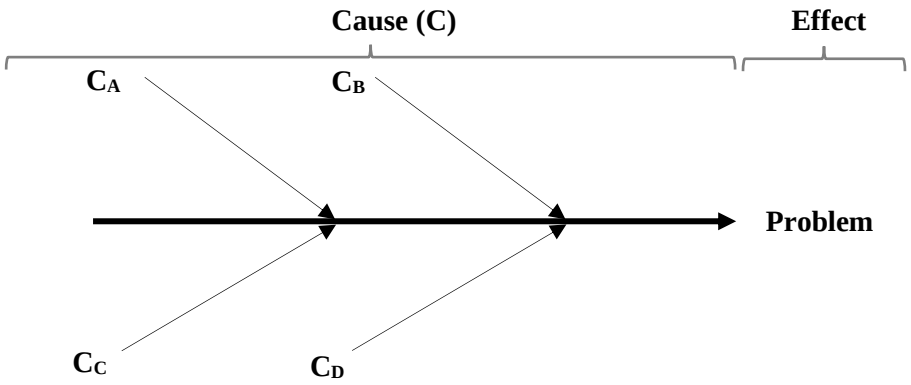


Figure 4. 2 Cause and Effect general scheme.

These "bones" represent different categories of potential causes (referred to as C), such as processes, people involved, equipment, or materials. Each bone could then further branch to identify specific causes within that category. The causes are often identified through brainstorming or root cause analysis. Once all potential causes have been identified, can be analyzed and prioritized based on their probability and impact.

The Ishikawa diagram is often used in conjunction with other quality management tools, such as FMEA.

Detailed information for applications is included in chapter 2.2.2.3., 2.2.3.4., and 3.4.2. in Part III.

4.2.Failure Mode Effects Analysis

FMEA is a systematic approach used to detect and evaluate possible failure modes and their effects on a product or process. The general scope of an FMEA includes establishing and preventing failures before they occur or mitigating their impact if they do occur[19,23].

FMEA involves several critical steps summarizes as the description of the process flow to be analyzed by specifying its sequence, the identification of the potential failure modes, the determination of their effects, and the evaluation of the severity of their effects, the probability of appearance of failures, and detectability of those effects. As a result, the risk priority number (from now on RPN) is calculated for each evaluated cause, as the product of the severity, the probability of appearance, and the probability of detection, which is used to prioritize which failures to address first. Based on the calculated value, corrective measures must be identified and implemented to eliminate the risk or reduce its level to limits considered acceptable[23,24].

Detailed information for applications is included in chapter 1.2.3.1., and 4.2.1 in Part III.

5. BIBLIOGRAPHY

(1) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. January 20, 2012. www.ema.europa.eu.

(2) Food and Drug Administration. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. 2004. http://www.fda.gov/cvm/guidance/published.html.

(3) Panzitta, M.; Ponti, M.; Bruno, G.; Cois, G.; D'Arpino, A.; Minghetti, P.; Mendicino, F. R.; Perioli, L.; Ricci, M. The Strategic Relevance of Manufacturing Technology: An Overall Quality Concept to Promote Innovation Preventing Drug Shortage. *Int J Pharm* **2017**, *516* (1–2), 144–157. https://doi.org/10.1016/j.ijpharm.2016.11.018.

(4) Arden, N. S.; Fisher, A. C.; Tyner, K.; Yu, L. X.; Lee, S. L.; Kopcha, M. Industry 4.0 for Pharmaceutical Manufacturing: Preparing for the Smart Factories of the Future. *Int J Pharm* **2021**, *602*. https://doi.org/10.1016/j.ijpharm.2021.120554.

(5) Sacher, S.; Khinast, J. G. An Overview of Pharmaceutical Manufacturing for Solid Dosage Forms. In *Methods in Pharmacology and Toxicology*; Humana Press Inc., 2016; Vol. 32, pp 311–383. https://doi.org/10.1007/978-1-4939-2996-2_10.

(6) International Council for Harmonisation. ICH Guideline Q11 on Development and Manufacture of Drug Substances. *EMA/CHMP/ICH/425213/2011*. February 11, 2013, pp 1–29.

(7) Food and Drug Administration. Guidance for Industry Process Validation: General Principles and Practices Guidance for Industry. January 2011, pp 1–19.

(8) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. 2014.

(9) European Medicines Agency. ICH Guideline Q12 on Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management. *EMA/CHMP/ICH/804273/2017*. March 4, 2020, pp 1–31. www.ema.europa.eu/contact.

(10) European Medicines Agency. ICH Guideline Q8(R2) on Pharmaceutical Development. *EMA/CHMP/ICH/167068/2004*. May 28, 2014, pp 1–28.

(11) Food and Drug Administration Department of Health and Human Services. Part 11, Electronic Records; Electronic Signatures. 2003.

(12) European Medicines Agency. Chapter 1: Pharmaceutical Quality System. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. European Commission Health and Consumers Directorate 2012, pp 1–7.

(13) European Medicines Agency. Chapter 6: Quality Control. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. 2014.

(14) European Medicines Agency. Chapter 4: Documentation. *Good Manufacturing Practice Medicinal Products for Human and Veterinary Use*. May 17, 2011, pp 1–9.

(15) European Pharmacopoeia 11.0. 5.28. Multivariate Statistical Process Control. January 1, 2023, pp 907–911.

(16) Yu, L. X.; Amidon, G.; Khan, M. A.; Hoag, S. W.; Polli, J.; Raju, G. K.; Woodcock, J. Understanding Pharmaceutical Quality by Design. *AAPS Journal* **2014**, *16* (4), 771–783. https://doi.org/10.1208/s12248-014-9598-3.

(17) Swarbrick, B.; Gadsby, M.; Wempen, F. *QbD & PAT*, Optimal Special.; Mengle, R., Ed.; John Wiley & Sons: Chichester, West Sussex, 2018.

(18) European Medicines Agency. ICH Guideline Q10 on Pharmaceutical Quality System. *EMA/CHMP/ICH/214732/2007*. May 28, 2014, pp 1–20.

(19) International Council for Harmonisation. ICH Guideline Q9 on Quality Risk Management. *EMA/CHMP/ICH/24235/2006*. September 2015.

(20) Igne, B.; Airiau, C.; Talwar, S.; Towns, E. Chemometrics in the Pharmaceutical Industry. In *Comprehensive Chemometrics*; Elsevier, 2020; Vol. 4, pp 33–68. https://doi.org/10.1016/b978-0-12-409547-2.14638-4.

(21) European Medicines Agency. Annex 20. Quality Risk Management. *Good Manufacturing Practice Medicinal Products for Human and Veterinary Use*. 2008, pp 1–25.

(22) Bakshi, I.; Puri, K.; Kaur, S. Applications of Quality Risk Management Tool in Pharmaceutical Industry: A Comprehensive Review. *Pharmaspire* **2018**, *10* (4), 148–152.

(23) Van Leeuwen, J. F.; Nauta, M. J.; de Kaste, D.; Odekerken-Rombouts, Y. M. C. F.; Oldenhof, M. T.; Vredenbregt, M. J.; Barends, D. M. Risk Analysis by FMEA as an Element of Analytical Validation. *J Pharm Biomed Anal* **2009**, *50* (5), 1085–1087. https://doi.org/10.1016/j.jpba.2009.06.049.

(24) Cogdill, R. P.; Anderson, C. A.; Drennen, J. K. Validation Column: Risk Analysis for Near Infrared Method Development. *NIR news* **2004**, *15* (5), 12–13.

PART II

DATA COLLECTION TOOLS AND MULTIVARIATE ANALYSIS TECHNIQUES

1. ANALYTICAL TECHNIQUES

Analytical chemistry as a fundamental part of the pharmaceutical industry provides quality and safety in the drug development process[1].

Recently the monographs for product analysis include a variety of analytical methods. Univariate and multivariate analytical techniques are involved such as High-Performance Liquid Chromatography (from now on HPLC)[2], Infrared Spectroscopy (from now on IR)[3], TOC[4], NIR Spectroscopy[5], and Raman Spectroscopy[6].

The following chapters describe the fundamentals and main applications in the pharmaceutical industry for the selected analytical techniques.

1.1.Near-Infrared Spectroscopy

NIR spectroscopy was discovered by William Herschel, who in 1800 first demonstrated the existence of the IR radiation, by examining the behaviour of the spectrum of solar radiation due to the heating effect. Despite this, it was not until the 1970s that Karl H. Norris involved the NIR spectroscopy in food analysis. During the 1900s, the first publications were born that reflected the high applicability of the technique[7].

Recently, the applications of NIR spectroscopy have increased due to its advantages over conventional analytical techniques. The spectra can be recorded for solid and liquid samples due to their low absorption coefficient, with no prior sample manipulation, providing spectra quickly, and predicting simultaneously physical and chemical parameters of the product of interest in a fast, non-destructive, or invasive way. Furthermore, it is environmentally friendly since it does not require the use of reagents. Added to these properties is the possibility of using the technique in atline, inline, online, or offline mode.

Nowadays NIR spectroscopy, involving the use of chemometrics as a tool for the mathematical and statistical treatment of the obtained data, is being used in different fields for routine control analyses including the pharmaceutical sector.

The results obtained with the NIR technique are comparable, even higher, in terms of accuracy concerning the conventional analytical techniques since the preparation of the sample to be analyzed is not required[7]. It should be noted the low sensitivity of the spectroscopic technique, compared to conventional techniques, when substances with significantly low content are determined (less than 0.1% (p/p))[8]. In addition, obtaining a NIR procedure capable of accurately, precisely, and robustly determining the parameter of interest requires the incorporation of a large set of samples to include the variability associated with the product of interest[7].

1.1.1. Fundamentals and Instrumentation of NIR Spectroscopy

The IR of the electromagnetic spectrum can be divided into three zones: NIR 780-2500 nm, mean infrared (from now on MIR) 2500-4.10^4 nm, and far infrared (from now on FIR) 4.10^4-10^6 nm.

NIR spectrum is a result of overtones (transitions between the fundamental vibrational state v= 0 and higher excited states v= ± 2, ± 3, etc.), and combinations of fundamental vibration bands (appear between 1900-2500 nm, and are produced on polyatomic molecules, where the interaction between the vibrational modes takes place simultaneously causing energy changes) appearing typically as broad and overlapped bands[9].

Absorption occurs when a molecule experiences a net change in the dipole moment as it rotates or vibrates. Hence, the interaction between the electric field and the molecule occurs. The transfer of energy causes the increase of the dipole moment and consequently, the absorption of the radiation takes place if an equivalence is produced between the molecular frequencies of radiation and vibration/rotation. Variations can be made in the length or in the angle of the links (stretching or blending, respectively) due to the aforementioned vibrations[9].

The dipole moment of a bond is related mainly to the difference in charge between the atoms that are involved and the distance between the atoms that form the bond. For this reason, hydrogen-containing bonds appear as the principal absorption bands in the NIR spectrum, O-H, C-H, N-H, and S-H groups. However, non-existent or weak bands are observed for the C=O, C-C, and C-Cl bonds. Moreover, the homonuclear species, such as N_2, O_2, or H_2 not absorbed in the NIR region[7,9,10].

The harmonic oscillator model describes molecular vibrations (Hook's law), where the characteristics of the vibration are close to those that would follow two masses joined by a spring. In the equilibrium position of the masses, the associated potential energy is zero, however, the modification of the distance between the masses increases the potential energy (see Equation 1. 1). The only transitions capable of causing a change in the dipole moment are those produced between consecutive energy levels, A*v*= ±1 (see Equation 1. 2 and Figure 1. 3).

$$Ev = \left(v + \frac{1}{2}\right) h\omega$$

h:$\frac{h}{2\pi}$; where h is the Plank constant.
v: vibrational state (Av=±1).
ω=$\sqrt{\frac{k}{m}}$; where k is the constant force and m the bonding atoms mass.

Equation 1. 1.

$$\Delta Ev = \Delta Ee = hv$$

E*v*: vibrational energy.
E*e*: emitted energy.
v: fundamental vibrational frequency.

Equation 1. 2.

However, the behavior of real molecules cannot be explained by this oscillator, since it does not consider Coulomb repulsion between atoms or bond dissociation. For this reason, NIR absorption is based on the premise that the behavior of molecules is based on the quantum mechanical model of an *anharmonic oscillator* (see Equation 1. 2 and Figure 1. 3). Even though the molecule in both models has the same behavior in the equilibrium position, the transitions could appear at higher vibrational levels, due to the energy difference between consecutive vibrational levels being less as they increase (Av= ±1, ±2, 3)[7,9–11].

$$\Delta Ev = hv[1-(2v+\Delta v+1)X]$$

X: anharmonicity constant

Equation 1. 3.

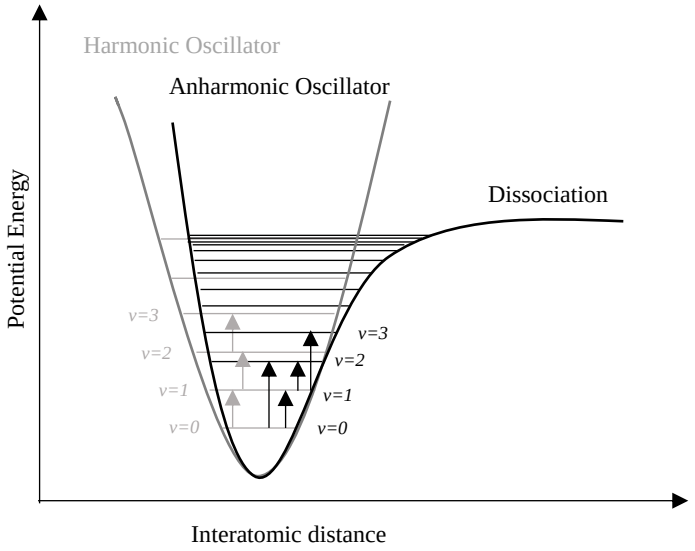


Figure 1. 3 Graphical representation for the harmonic and anharmonic oscillator.

The structure of a NIR spectrometer almost universally includes a tungsten halogen lamp with a quartz window that emits a constant and stable polychromatic light through the NIR region into a wavelength selector. An alternative to obtaining a restricted range of absorption is the Light Emitting Diodes. Inside it, a diffraction red converts this polychromatic light into discrete values of wavelengths. Different wavelength selectors are available depending on the final application[11]:

- Dispersive instrument: red monochromators (spectral scanning that refracts the light at different angles based on the wavelength), or polychromators (simultaneous wavelength detection with different channels).
- Non-dispersive instrument: filters (allowing the selection of a part of the spectrum), Acusto-Optic Tunable Filters (selects the wavelength based on radio frequency signals), or Fourier Transform (from now on FT) (modulates the radiation allowing simultaneously the codification of the wavelengths and the measurement of these).

When the light goes out from the wavelength selector, the light goes through a sample holder where it is reflected or goes into the sample and then goes to the detector. Commonly used detectors in NIR spectral region use dispositive composed of semiconductors such as photodiodes (Si, Ge, and *InGaAs*) or photoconductors (*PbSe*, *InAs*, and *InSb*)[7].

The adequacy of the spectral measurement is conditioned by the physical characteristics of the sample. Acquisition modes include reflectance, transmittance, and transflectance[11] (see Figure 1. 4). During this study, the reflectance mode was used.

In transmittance (from now on abbreviated as T) mode (for gas, liquid, semi-liquids, and solids, commonly for transparent samples), the incident beam (I_o) crosses the sample and arrives at the detector where the absorbance (A) is calculated by Lambert-Beer law (see Equation 1. 4)[12]; which is the fraction of radiation beam that is not absorbed by the sample ($I_{transmitted}$).

$$A = log\frac{1}{T} = log\frac{Io}{I_{transmitted}}$$

Equation 1. 4.

In reflectance mode (for solids and semisolids), the beam is reflected in the sample in two components: speculate and diffuse light. The second one is a result of the dispersion of the radiation in all directions. The last one interacts with the sample and it's detected ($I_{scattered}$) and expressed in apparent absorbance (A_{ap}) (see Equation 1. 5)[12].

$$Aap = log\frac{1}{R} = log\frac{I_{scattered}}{I_{reflectance}} = a' * c$$

a': the proportionality constant.

c: the concentration of absorbent analytes.

$I_{reflectance}$: the reference reflectance.

R: the relative reflectance.

Equation 1. 5.

Finally, in transflectance mode (for liquids and semi-liquids, for emulsions and turbid liquids commonly), the beam crosses the sample and is reflected in a reflector that is in contact with the sample and arrives at the detector ($I_{reflectance}$).

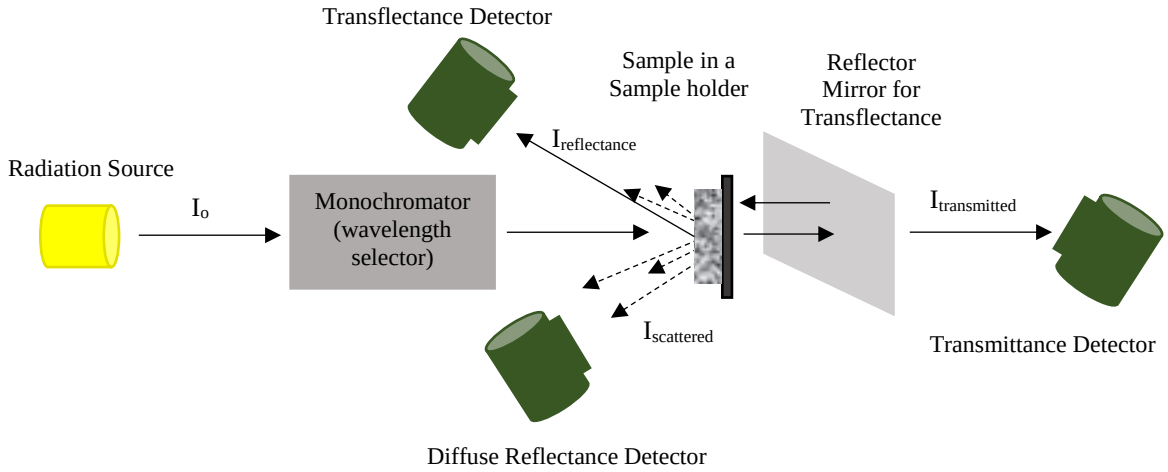


Figure 1. 4 General scheme of different acquisition modes for a NIRS.

1.1.2. Regulatory Environment in the Pharmaceutical Industry of NIR Spectroscopy

The CQA analysis in the pharmaceutical industry applied to SM[13] , intermediates, and finished products[14], are recently performed by the NIR method. Moreover, the advantages of this methodology encourage industries to move forward to a real process monitoring[15].

In this context, the specific guidelines published are detailed:

- *Guidance for Industry PAT- A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance* by the FDA regulation authority (September 2004)[16].
- *Guideline on the use of Near Infrared Spectroscopy by the Pharmaceutical Industry and the Data Requirements for New Submission and Variations* by the EMA (June 2014)[17].
- *Addendum to EMA/CHMP/CVMP/QWP/17760/2009 Rev 2: Defining the Scope of an NIRS Procedure* by the EMA (June 2014)[18].
- *Development and Submission of Near Infrared Analytical Procedures* by the FDA regulation authority (August 2021)[19].

1.2.Raman Spectroscopy

In 1928, Raman and Krishnan empirically discovered a new scattering effect called Raman scattering publishing an article in the journal Nature[20]. Inspired by the work done by Rayleigh, they carried out an experiment in which they demonstrated that the blue color of the sea was not a reflection of the blue sky, but that this color came from the scattering of sunlight by the interaction with water molecules.

In the 1930s and 1940s, the Raman technique was initially experimentally tested. The advances in instrumentation and the advantages that this technique presents over other conventional techniques, among them its versatility, since it can be used in opaque and coloured samples, also translucent and transparent, for semi-solid and solid liquids; the non-invasive analysis of the samples and the absence of a reference light path because it is defined with a scattering effect, have favoured its expansion in the pharmaceutical industry.

Raman considered as a complementary NIR technique is regulated by the same guidelines.

Some key applications of Raman spectroscopy include the determination of polymorphs[21], quality control[22,23], process control[24], product authentication[25], as well as the control of the useful life of a product[26].

1.2.1. Fundamentals and Instrumentation of Raman Spectroscopy

Chemically bonded ions and atoms are submitted to constant rotational and vibrational motions. These oscillations are carried out at certain frequencies that vary according to the dynamic behaviour of the bonds, and the molecular mass of the particles involved in them. The value of molecular energy will be associated with each of the movements of the molecule (vibrational and rotational).

The incident radiation on the oscillating dipoles induced in the molecule produces electromagnetic radiation that describes the light scattering phenomena. The external electric field induces the deformation of the electron cloud of the molecule (molecular polarizability) causing an induced dipole moment[27].

The laser light at a specific wavelength interacts with molecules with monochromatic radiation, causing the absorption of energy by the electrons, and modifying their vibrational state.

The electron tends to return to its initial vibrational state, releasing the absorbed energy in the form of photons.

If the amount of energy absorbed is equal to the energy emitted, the electron returns to the same vibrational state involving an elastic scattering (Scattering or Rayleigh effect).

The 1% of the scattering has different characteristics being inelastic (Raman scattering). This 1% finds the information about the molecule of interest, providing a unique molecular fingerprint for each one. This phenomenon takes place when the energy absorbed is different from the energy emitted by the photon and as a consequence the frequency of the photon, thus returning to a different vibrational state from the initial one. Considering the premise that the frequency variations observed in the Raman scattering phenomenon are equivalent to energy variations ($E= h\nu$, where h is the Plank constant).

In this case, there are two possible scenarios (see Figure 1. 5):

- The absorbed frequency is higher than the emitted one, the electron absorbs energy and returns to a higher vibrational state (v=m) than the initial one (v=0), emitting fewer photons than it has absorbed, remaining with excess energy (Stokes Raman Scattering).
- The electron in its initial state is already excited, starting from a higher vibrational state than the fundamental (v=m). When it comes off the energy returns to its ground state (v=0) so it emits more photons. In this case, the absorbed frequency is less than the emitted (Anti-Stokes Raman Scattering). Being the less common.

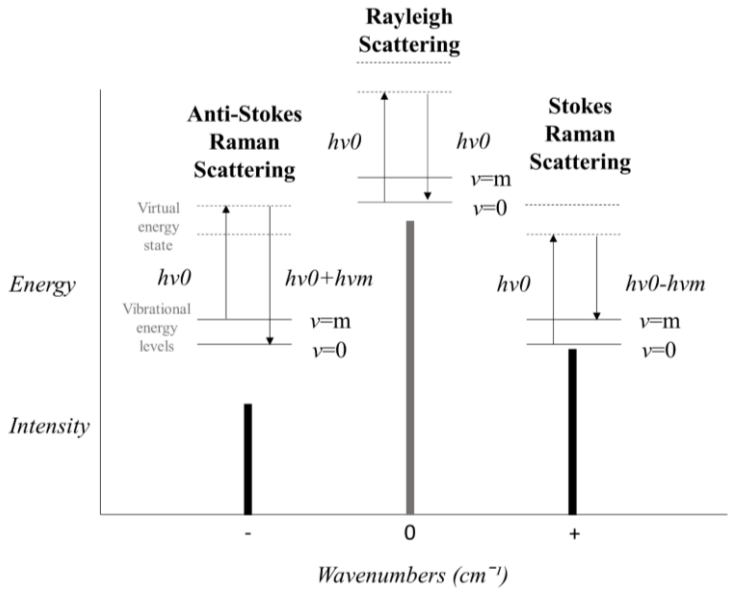


Figure 1. 5 Raman and Rayleigh scattering.

Based on the Maxwell-Boltzmann law of energy distribution, at room temperature, 1% of the molecules are in a vibrational state other than the lowest energy, and therefore, the probability that energy transfers that give rise to dispersion occur Raman Stokes is much larger than that of the Raman Anti-Stokes scattering. This difference in intensity means that work is usually done by measuring only the Stoke dispersion effect, showing the positive part of the abscissa axis.

Raman spectrometers contain a laser light source, focusing optics, and spectrograph(s) including scattering elements and a detector to capture the light[20].

Excitation in the UV, visible, and the NIR regions are available with different laser sources (commonly 785, 1064 nm, and 830 nm exceptionally)[28]. The working range could be presented from 20 to 4000 cm^{-1}.

The equipment involved in the experimental part is based on transmission Raman spectroscopy. The transmission mode is based on collecting the signal at the opposite of the laser illumination (scheme shown in Figure 1. 6). The laser travels through the entire sample due to the lack of light absorption.

If there is some fraction of the light achieving the detector without interaction with the tablet it causes a distortion in the spectra and model as it includes information not related to the sample.

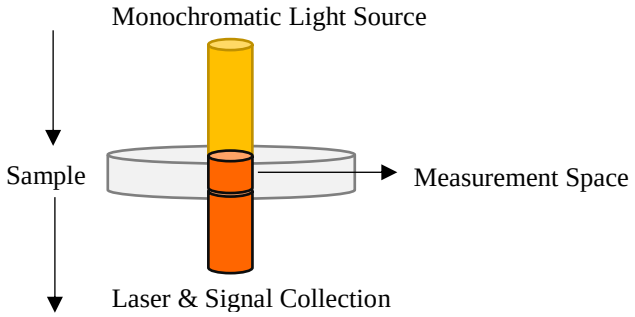


Figure 1. 6 Scheme for the transmission Raman spectroscopy.

The equipment provides a quick analysis of 30 seconds to 1 minute, commonly. The size of the illumination spot is possible in 2, 4, or 8 mm depending on the size of the sample to be analyzed. The power of the laser corresponds to 650 mW and the wavelength for it is 830 nm, allowing a balance between the sample signal and the fluorescence noise. The resolution of the spectrum is less than 8 cm^{-1}.

1.3. Total Organic Carbon

TOC is defined as a non-product-specific technique. Hence, it allows to determine the total amount of organic carbon present in a sample but does not allow to determine the amount present of specific compounds. The most common way of applying the TOC method is to assume that all residues detected are to the target residue. For the use of TOC, the target residue must have adequate aqueous solubility as this is considered the optimal matrix [29].

A short analysis of the sample, the low consumption of reagents during the analysis, and precise and independent quantification of the sample against the working matrix are advantages of the TOC analysis compared to other conventional techniques (i.e. HPLC).

1.3.1. Fundamentals and Instrumentation of TOC

TOC was introduced in the industry during the 90s by the included chapters at the pharmacopoeias[4,30].

Generally, TOC technology is based on the oxidation of organic molecules with UV light or by adding small volumes of a chemical oxidation solution. As a result, the measurement of the CO_2 generated is carried out [31].

The analyzer can be divided into the following main blocks:

- Sample injection system: online and offline measurements.
- Chemical reagent system: acid and oxidant reagents.
- UV-visible system: ultraviolet light for sample oxidation.
- CO_2 measurement system: transfer module for CO_2 separation.

For a pharmaceutical-grade TOC analyzer, once the sample is introduced into the analyzer (online or offline), the sample is acidified using a phosphoric solution (H_3PO_4 6 M) at a specific flow rate to reduce the sample's pH to a value of approximately 2, guaranteeing the precision of the values to be obtained. The sample is then divided into two separate but equal flow streams. One flow will be used for the determination of inorganic carbon (hereinafter IC), and the other for the determination of total carbon (hereinafter TC)[32].

The TC flow is exposed to ultraviolet light (185 and 254 nm) to oxidize it. If necessary, depending on the concentration of the target sample, full oxidation is achieved with the addition of a persulfate solution. In this way, the carbon present in the sample is converted into CO_2 following the equations below, from Equation 1. 6 to Equation 1. 9. It should be noted that the samples that do not require the addition of persulfate are those with a concentration lower than 1 ppm, obtaining total oxidation using only the hydroxyl radicals produced during the photolysis of water.

$$H_2O+hv\ (185\ nm) \rightarrow OH\cdot + H\cdot \quad \textit{Equation 1. 6.}$$

$$S_2O_8^{-2} + hv\ (254\ nm) \rightarrow 2SO_4\cdot^- \quad \textit{Equation 1. 7.}$$

$$SO_4\cdot^- + H_2O \rightarrow SO_4^- + OH\cdot \quad \textit{Equation 1. 8.}$$

$$\text{Organic compounds} + OH\cdot \rightarrow CO_2 + H_2O \quad \textit{Equation 1. 9.}$$

Complete oxidation is key in the measurement process. If the measurement of the TOC value is carried out when the oxidation is not complete, the presence of inorganic acids generated as intermediate species during the oxidation to CO_2 may increase the conductivity value obtained, which can give rise to high TOC values (obtaining thus false positives).

Then the CO_2 separation for both streams takes place via the transfer module. This is a conductometric design based on CO_2-selective membrane technology from the rest of the compounds that interfere in the sample. This technology is patented by SUEZ.

The membrane allows the sample to be separated from the analyzer on the deionized waterside, a closed circuit that contains two conductivity cells (one for each separated flow, TC and IC). This membrane allows CO_2 from the sample to selectively pass through the membrane into the deionized water. Interfering compounds and other oxidation by-products are retained on the sample side. So, after the reaction of CO_2 with water, carbonic acid is produced. This dissociates into hydrogen ions and bicarbonate ions that are collected in the measurement cells (see Equation 1. 10).

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \quad \textit{Equation 1. 10.}$$

The ion exchange resin is responsible for the removal of HCO_3^- and other ions. The concentration of TC and IC is determined by the conductivity resulting from the measurement of the CO_2 present in the TC and IC flows, respectively. The TOC value is calculated as an indirect measurement of the difference between TC and IC (see Equation 1. 11)[32], which is commonly reported in parts per billion.

$$TOC = TC - IC \quad \textit{Equation 1. 11.}$$

1.3.2. Regulatory Environment and Applications in the Pharmaceutical Industry of TOC

FDA reported four CQAs for water injection and purified water including inorganic and organic contaminants, endotoxins, and microbial. The control of these four parameters ensures the purity of the water used in all manufacturing areas.

There has been an increasing importance of TOC in detecting possible contaminants in water since the publication of 2.2.44. Ph. Eur. TOC Chapter on 2016 [4].

Pharmacopoeias describe TOC determinations at different water purities (Ph. Eur, USP, and Japanese). The involved procedures for qualifying the methodology used to analyze TOC and the interpretation of the results are described. The instrument must be properly calibrated and verified. The system's suitability has to be confirmed.

Hence, TOC analysis allows to obtain information on the degree of process control by detecting chemical compounds in water systems and process equipment [33].

1.4.Instrument Validation

Compliance with regulatory agencies must be ensured through the validation process for the use of the different analytical techniques in the pharmaceutical industry.

In addition to the correct validation of the model to be used (see more information in chapter 2. of PART II), validation of the software and the instrument must be carried out.

1.4.1. Software Validation

The involved software must be GMP compliant, considering the FDA 21 CFR Part 11[34] and the EU Annex 11 GMP[35].

It consists of objective evidence that the software specifications meet the intended objective over time. As a result, the reliability of the system is increasingly favouring the integrity of the data.

1.4.2. Instrument Qualification

Three main stages are included during the validation process of the instrument being good for practice compliant: the installation qualification (from now on abbreviated as IQ), the operation qualification (from now on abbreviated as OQ), and the performance qualification (from now on abbreviated as PQ). Descriptions are included in Table 1. 1 (USP 1058[36], FDA Quality System Regulation 21 CFR part 11[34], Annex 15 GMP[37]).

Table 1. 1 Stage for the instrument validation process.

Stage	Definition
IQ	Regarding the installation of hardware and software. Documentary evidence about its installation complying with the supplier's specifications and the security criteria in the established place.
OQ	Documentary evidence about the applicability of the system to the intended purpose. The correct operation of the equipment according to requirements. The data backup and user management are included.
PQ	Documentary evidence about the operating parameters for real operation, verifying process stability over time. Test equipment, including limits and calibration, is included.

1.5.Bibliography

(1) International Council for Harmonisation. ICH Guideline Q14 on Analytical Procedure Development. *EMA/CHMP/ICH/195040/2022*. March 31, 2022.

(2) European Pharmacopoeia. 2.2.29. Liquid Chromatography. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France, January 2023, pp 55–56.

(3) European Pharmacopoeia. 2.2.24. Absorption Spectrophotometry, Infrared. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France January 2023, pp 45–48.

(4) European Pharmacopoeia. 2.2.44. TOC in Water for Pharmaceutical Use. *Ph. Eur. 11th*. Strasbourg, France January 2023, pp 85–86.

(5) European Pharmacopoeia. 2.2.40. Near-Infrared Spectroscopy. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France, January 2023, pp 76–81.

(6) European Pharmacopoeia. 2.2.48. Raman Spectroscopy. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France, January 2023, pp 101–103.

(7) Blanco, M.; Villarroya, I. NIR Spectroscopy: A Rapid-Response Analytical Tool. *Trends in Analytical Chemistry* **2002**, *21* (4), 240–250.

(8) Jamrógiewicz, M. Application of the Near-Infrared Spectroscopy in the Pharmaceutical Technology. *J Pharm Biomed Anal* **2012**, *66*, 1–10. https://doi.org/10.1016/j.jpba.2012.03.009.

(9) Heinz W. Siesler. Basic Principles of Near-Infrared Spectroscopy. In *Handbook of Near-Infrared Analysis*; Donald A. Burns, Emil W. Ciurczak, Eds.; CRC Press, 2008; Vol. 35, pp 7–20.

(10) *Analytical Techniques in the Pharmaceutical Sciences*; Müllertz, A., Perrie, Y., Rades, T., Eds.; Advances in Delivery Science and Technology; Springer New York: New York, NY, 2016. https://doi.org/10.1007/978-1-4939-4029-5.

(11) Michael B. Simpson. Near-Infrared Spectroscopy for Process Analytical Technology: Theory, Technology and Implementation. In *Process Analytical Technology. Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*; Bakeev, K. A., Ed.; Wiley: USA, 2010; pp 107–156.

(12) Peter R. Griffiths; Donald J. Dahm. Continuum and Discontinuum Theories of Diffuse Reflection. In *Handbook of Near Infrared analysis*; Donald A. Burns, Emil W. Ciurczak, Eds.; CRP Press, 2008; Vol. 35, pp 21–64.

(13) Blanco, M.; Romero, M. A. Near-Infrared Libraries in the Pharmaceutical Industry: A Solution for Identity Confirmation. *Analyst* **2001**, *126* (12), 2212–2217. https://doi.org/10.1039/b105012p.

(14) Blanco, M.; Bautista, M.; Alcalà, M. API Determination by NIR Spectroscopy Across Pharmaceutical Production Process. *AAPS PharmSciTech* **2008**, *9* (4), 1130–1135. https://doi.org/10.1208/s12249-008-9156-3.

(15) Blanco Romía, M.; Alcalá Bernárdez, M. NIR Spectroscopy in Pharmaceutical Analysis: Off-Line and At-Line PAT Applications. In *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*; Bakeev, A. K., Ed.; John Wiley & Sons, 2010; pp 463–491.

(16) Food and Drug Administration. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Rockville, MD, September 2004, pp 1–19.

(17) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. London, UK, January 20, 2012.

(18) European Medicines Agency. Addendum to EMA/CHMP/CVMP/QWP/17760/2009 Rev 2: Defining the Scope of an NIRS Procedure. London, UK, June 5, 2014, pp 1–4.

(19) Food and Drug Administration. Development and Submission of Near Infrared Analytical Procedures Guidance for Industry. Silver Spring, MD, August 2021, pp 1–28.

(20) Paudel, A.; Raijada, D.; Rantanen, J. Raman Spectroscopy in Pharmaceutical Product Design. *Adv Drug Deliv Rev* **2015**, *89*, 3–20. https://doi.org/10.1016/j.addr.2015.04.003.

(21) Antonio, M.; Carneiro, R. L.; Maggio, R. M. A Comparative Approach of MIR, NIR and Raman Based Chemometric Strategies for Quantification of Form I of Meloxicam in Commercial Bulk Drug. *Microchemical Journal* **2022**, *180*. https://doi.org/10.1016/j.microc.2022.107575.

(22) Ohashi, R.; Fujii, A.; Fukui, K.; Koide, T.; Fukami, T. Non-Destructive Quantitative Analysis of Pharmaceutical Ointment by Transmission Raman Spectroscopy. *European Journal of Pharmaceutical Sciences* **2022**, *169*. https://doi.org/10.1016/j.ejps.2021.106095.

(23) Pino-Torres, C.; Maspoch, S.; Castillo-Felices, R.; Pérez-Rivera, M.; Aranda-Bustos, M.; Peña-Farfal, C. Evaluation of NIR and Raman Spectroscopies for the Quality Analytical Control of a Solid Pharmaceutical Formulation with Three Active Ingredients. *Microchemical Journal* **2020**, *154*. https://doi.org/10.1016/j.microc.2019.104576.

(24) Lomont, J. P.; Smith, J. P. In Situ Raman Spectroscopy for Real Time Detection of Cysteine. *Spectrochim Acta A Mol Biomol Spectrosc* **2022**, *274*. https://doi.org/10.1016/j.saa.2022.121068.

(25) Bloomfield, M.; Andrews, D.; Loeffen, P.; Tombling, C.; York, T.; Matousek, P. Non-Invasive Identification of Incoming Raw Pharmaceutical Materials Using Spatially Offset Raman Spectroscopy. *J Pharm Biomed Anal* **2013**, *76*, 65–69. https://doi.org/10.1016/j.jpba.2012.11.046.

(26) Liu, C.; Müller-Bötticher, L.; Liu, C.; Popp, J.; Fischer, D.; Cialla-May, D. Raman-Based Detection of Ciprofloxacin and Its Degradation in Pharmaceutical Formulations. *Talanta* **2022**, *250*. https://doi.org/10.1016/j.talanta.2022.123719.

(27) Peter Larkin. Chapter 2. Basic Principles. In *Infrared and Raman Spectroscopy. Principles and Spectral Interpretation*; Elsevier, 2011; pp 7–25.

(28) Peter, J. L. *IR and Raman Spectroscopy. Principles and Specral Interpretation*; Elsevier Inc.: USA, 2011.

(29) Parenteral Drug Association. *Points to Consider for Cleaning Validation. Nº 29.*; PDA, 2012.

(30) The United States Pharmacopoeia. 643. Total Organic Carbon; 2021; pp 266–267.

(31) Bisutti, I.; Hilke, I.; Raessler, M. Determination of Total Organic Carbon- An Overview of Current Methods. *TrAC - Trends in Analytical Chemistry* **2004**, *23* (10–11), 716–726.

(32) Martínez-Rosa, J. E. Method Validation, Qualification, and Performance Verification for Total Organic Carbon Analyzers. In *Practical Approaches to Method Validation and Essential Instrument Qualification*; Chung Chow Chan, Herman Lam, Xue Ming Zhang, Eds.; John Wiley & Sons, Inc., 2010; pp 299–325.

(33) Suez Water Technologies. *Total Organic Carbon for Cleaning Validation Programs; Laboratory, Online & at-Line Analysis.*; American Pharmaceutical Review, Ed.; 2020.

(34) Food and Drug Administration. Part 11, Electronic Records; Electronic Signatures. August 2003, pp 1–12.

(35) European Medicines Agency. Annex 11: Computerised Systems. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, June 30, 2011, pp 1–5.

(36) United States Pharmacopoeia. <1058> Analytical Instrument Qualification. USP-NF: Rockville, MD 2023.

(37) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 30, 2015, pp 1–16.

2. CHEMOMETRICS

The chemical discipline that involves the treatment of data extracted from chemical processes through the use of statistics and mathematics is chemometrics[1,2]. Likewise, the interaction of electronics and computing science with chemometrics, including analyst participation in decision-making, allows the success of the studies (see Figure 2. 2).

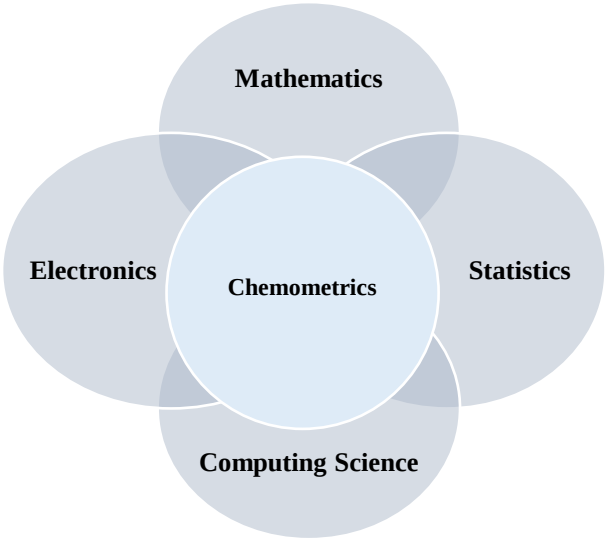


Figure 2. 2 Schematic representation of the interaction of chemometrics with other disciplines.

Chemometrics is involved as a statistical tool to establish a relationship between the studied samples and the predicted factors. As a result, improve the performance of the process and ensure the quality of the results, facilitating its interpretation, and justifying the conclusions.

The development of chemometrics was late due to the slow evolution of the instrumentation used. In the 1960s the development of new instrumentation was promoted. However, the extraction and treatment of the data were done manually. It was between the 1980s and 1990s, when, due to the popularization of the microcomputer, this discipline became more important. Appearing then the first specialized research journals in chemometrics: Chemometrics and Intelligent Laboratory Systems (1986) and Journal of Chemometrics (1987)[3].

The data is organized in the form of an object-variable matrix (rows and columns, respectively). Objects referred to experimental units to which variables are attributed {continuous numerical (countable number of values) or discontinuous (infinite number of values) or categorical} that describe their properties.

The study of the object-variable matrix is carried out following descriptive statistics, which provide a concise summary of the data, summarizing the data in numerical or graphical form, and inductive

statistics, which describe and make inferences about the population, demonstrate associations, and allow comparisons between observed characteristics.

The following sections describe the statistics commonly used during the application of chemometrics, as well as the mathematical pre-treatments involved and the models for the classification and quantification of the data.

2.1.Statistics Used in Chemometrics

2.1.1. Significance Tests

Significance tests are carried out to estimate the degree of difference between the results of the proposed analytical method and the real value, that is, the quantitative expression of random error, variations explained by chance that is inherently involved in each research process[3–5].

2.1.1.1.Paired t-tests

The paired *t*-test is applied to evaluate the difference between analyzed *n* paired samples (x_1, y_1 to x_n, y_n), containing the same number of samples each since it is known that could be significant variability between paired samples[3], following Equation 2 1. Moreover, the difference between the samples is unknown, hence the two-side test is selected.

$$t = \frac{\overline{d_d}\sqrt{n}}{s_d}$$

$\bar{d}$: difference values between the paired values.

s_d: standard deviation for the difference values between the paired values.

n= total number of samples.

n-1: degrees of freedom.

Equation 2 1.

The criterion is evaluated as follows: if the *t*-calculated is greater than the *t* tabulated, the null hypothesis is rejected (x_n=y_n) concluding in a positive bias from a method error with significantly different results between methods ($x_n \neq y_n$)[3]. Moreover, the probability is evaluated at a significant coefficient of α=0.05. Hence, if the evaluated probability is lower than the stipulated value the result is concluded as significant.

2.1.1.2.F-tests

The statistical *F*-test is applied to compare two variances following Equation 2 2, being the quotient between the highest and the lowest variance[3].

$$F = \frac{s_1^2}{s_2^2}$$

s_1^2/s_2^2: squares of the standard deviation for each evaluated value.

$n_1 - 1; n_2 - 1$: degrees of freedom for the numerator and denominator, respectively.

Equation 2 2.

The null hypothesis is defined as H_0: $s_1^2 = s_2^2$, and the alternative hypothesis is defined oppositely, H_i: $s_1^2 \neq s_2^2$. The null hypothesis is accepted if the calculated *F* is lower than the tabulated *F*[3].

2.1.1.3.Analysis of Variance

The analysis of variance (from now on ANOVA) is a statistical tool that allows the comparison between several samples means to establish if they are all the same (accepting the null hypothesis) or if at least one of them is different from the others (alternative hypothesis)[3].

To apply the ANOVA test, these three criteria must be met[3]:

- Independence: there should be no correlation between data series.
- Normality: the data distribution should be normal. Continuous distribution where the centre of the data is the average of these and the dispersion of them is explained by their standard deviation.
- Homogeneity: the variances of the data sets must be equal.

In addition, it is necessary to have at least one independent categorical variable, one-way ANOVA, (if there are two, defined as two-way ANOVA), and one continuous dependent variable.

See chapter 5.1.1.1. from Part III to see the related statistics.

2.2.Data Pre-treatment

Data pre-processing to remove noise or irrelevant information such as light scatter, baseline drift, background effects, or interferences[2,6–8]. See Figure 2. 3.

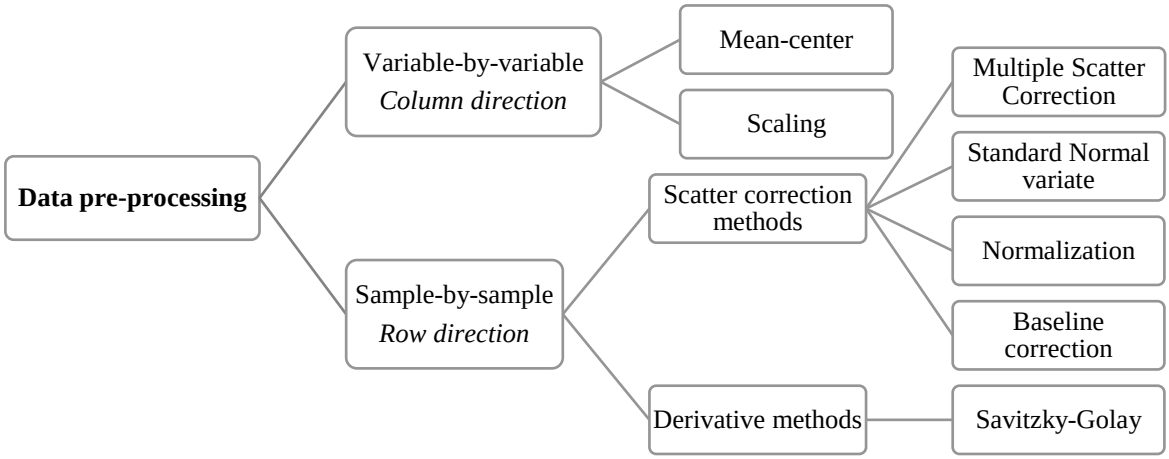


Figure 2. 3 Classification of data pre-processing.

Continuous dependent variables, applied in the row direction:

- Smoothing: probably the most common filter for eliminating spectral noise. It is based on the interpolation in small windows of a polynomial of a determined degree following the Savitzky-Golay (from now on abbreviated as S-G) methodology.
- Standard Normal Variate (from now on SNV): this pre-processing removes additive effects, the ones which add one constant magnitude to all the spectrum regarding to physical effect due to the particle size, roughness, shape, defined as scattering. It is applied to every single row without the modification of the spectral shape. For a given data matrix containing *M* spectra measured at *N* wavelengths the mathematical interpretation for each new pre-treated follows the Equation 2 3. See Figure 2. 4.

$$\hat{x}_m = \frac{(x_m - mean\ (x_m)}{std(x_m)}$$

Equation 2 3.

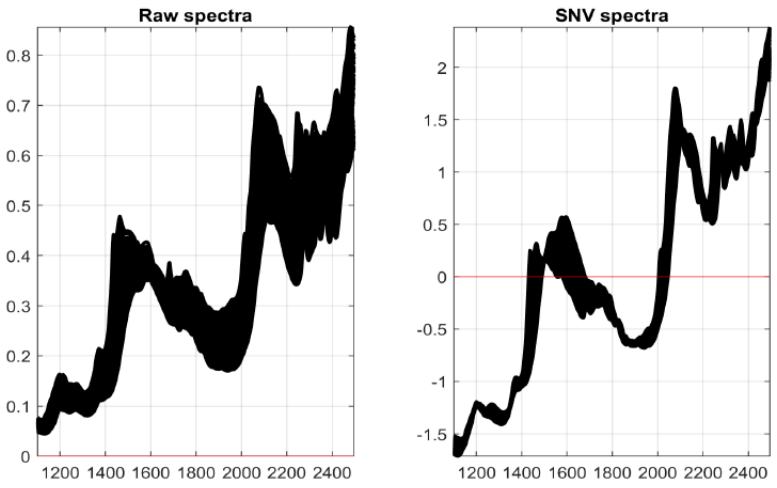


Figure 2. 4 SNV pre-processing.

- Multiplicative Scatter Correction (from now on MSC): this pre-processing removes constant effects in one spectrum based on scattering changing mainly the shape of the spectra. A reference spectrum is needed, being the mean, the median, or an external spectrum. Multiplicative effects are the ones adding a linear increasing magnitude with the wavelengths, which could be different for each spectrum, as a result of physical effects due to particle size, roughness, or shape, referred to as scattering. For a given data matrix (X) with M spectra measured at N wavelengths, for every spectrum the regression of x_m over the reference x_{ref}, calculating the slope and the offset. Then each spectrum is corrected ($\hat{x}_m$) following Equation 2 4[6]. See Figure 2. 5.

$$\hat{x}_m = \frac{(x_m - offset)}{slope}$$

Equation 2 4.

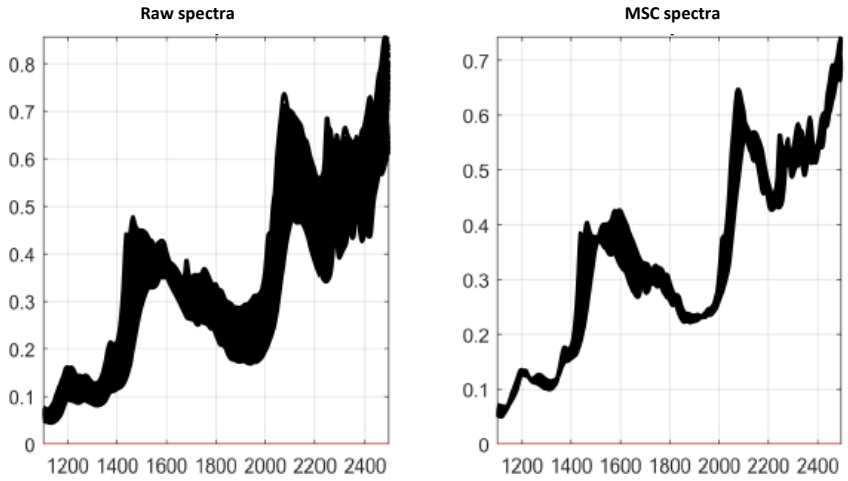


Figure 2. 5 MSC pre-processing.

- Derivatives (S-G methodology): this pre-processing allows noise minimization and spectra transformation. The optimization of three parameters is needed to achieve the optimum results, including the window size for the smoothing step, the polynomial degree, and the derivative degree to enlarge spectra differences[6]. See Figure 2. 6.

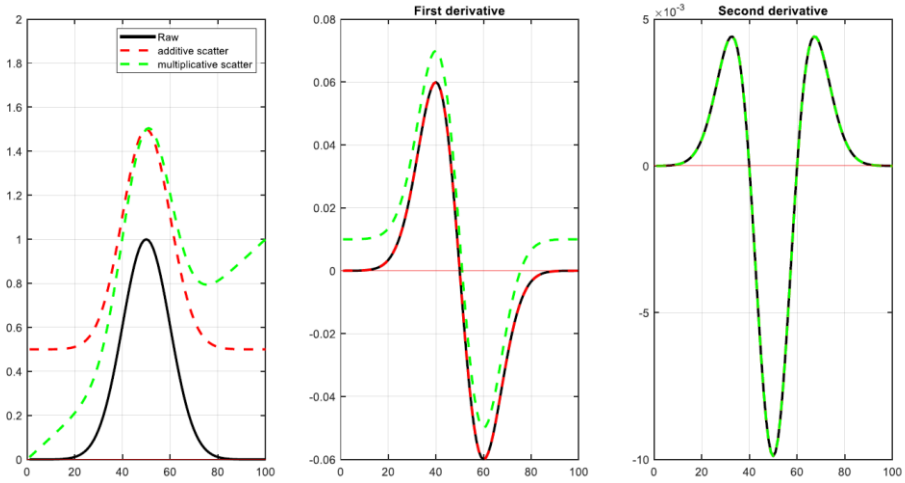


Figure 2. 6 Derivative pre-processing.

- Baseline correction: this processing allows the elimination of baseline offsets from data. Two possible methodologies are referred to: weighted skewed least squares involve an iterative approach determining the most probable points because of the baseline alone, iteratively fitting a baseline to each spectrum. The automatic Whittaker filter involves a piecewise method to remove data baseline offsets.

Continuous independent variables are applied in the column direction. They are used as normalization methods or together with normalization depending on the magnitude and the noise, including mean-center (Equation 2 5), scaling (Equation 2 6), and autoscaling (Equation 2 7)[6].

$$\hat{x}_i = x_i - \bar{x}$$

Equation 2 5.

$$\hat{x}_i = \frac{x_i}{std\ (x)}$$

Equation 2 6.

$$\hat{x}_i = \frac{x_i - \bar{x}}{std\ (x)}$$

Equation 2 7.

2.3.Pattern Recognition Methods

Pattern recognition methods as a critical part of chemometrics describing qualitative data analysis.

Most of the methods are based on the organization of several objects into groups according to similarities and mathematically expresses as correlation or distance calculations[1]:

- Correlation: based on the calculation of the correlation coefficient between two samples. It ranges between -1 and 1. The value of 1 indicates that there is a complete agreement between samples.
- Distances: expresses through a calculation of distances how different a sample is from another or a specific point in the workspace (see Equation 2 8).

$$D^2_{AB} = (X_A - X_B)^T \cdot \sum^{-1} \cdot (X_A - X_B)$$

Equation 2 8.

Being $\sum$ the dispersion matrix of the data. Depending on the value of this matrix, two types of distance are explained:

- $\sum^{-1}=1$; Distance defined as Euclidean distance. All groups have a spherical distance and there is no correlation between the variables.
- $\sum^{-1}$=inverse of the variance-covariance matrix; Distance defined as Mahalanobis distance (from now on MD). All groups have an ellipsoidal-shaped distance.

To establish the selection criteria of the statistical tools, it is necessary to previously choose whether the existence of specific categories or groups for the included objects is known (supervised pattern recognition methods) or if, on the contrary, the category membership is unknown (unsupervised pattern recognition methods)[1,9]. See Figure 2. 7.

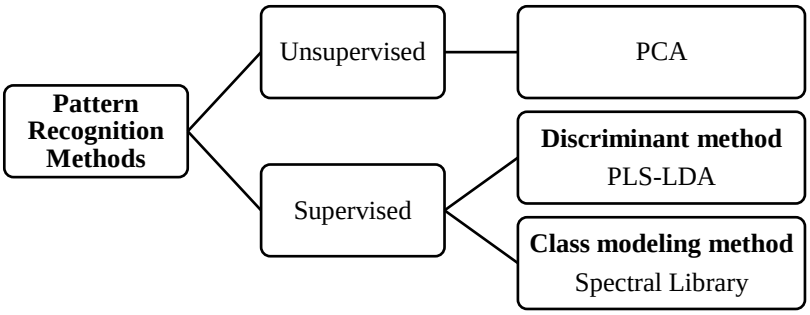


Figure 2. 7 Pattern Recognition Methods used in this thesis.

The discriminating method is referred to as multiclass classifiers where classes are modelled together, and a specific boundary between classes is determined. All data space is predicted. Class modeling methods are referred to as one-class classifiers: each class is modelled independently; a boundary around each class is determined separately. See Figure 2. 8.

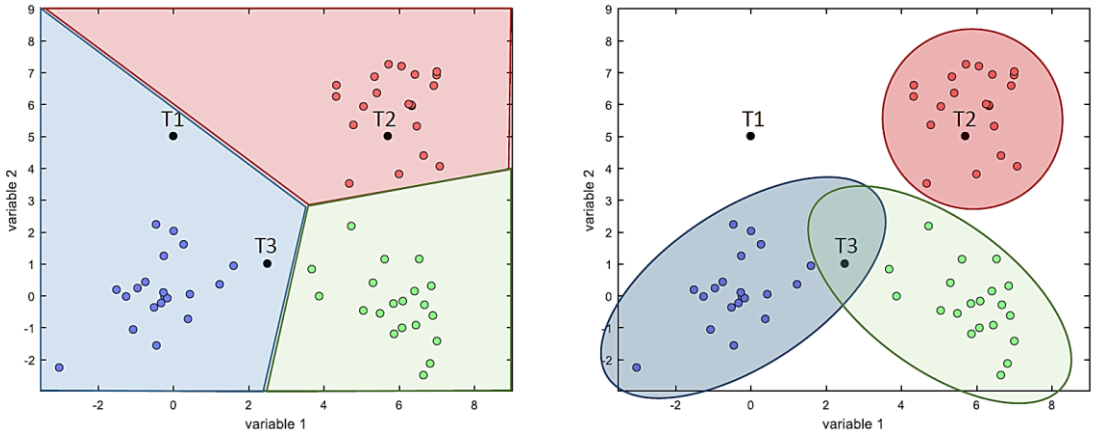


Figure 2. 8 Left) discriminant method, right) class modeling.

2.3.1. Descriptive Unsupervised Methods

The following chapter describes the unsupervised methods involved in this thesis.

2.3.1.1. Exploratory Analysis. Principal Component Analysis

Data exploration techniques are used to reveal the information contained.

Through PCA, new coordinate axes or principal components are established, as a linear combination of the original variables included, to show the spatial directions that offer the best view of the data. In this way, it is possible to explain relationships between the objects (*n*) and the variables (*m*) included in the data matrix. This information allows us to understand the factors that cause the inner relationship[2,3,6,7,10].

PCA analysis is a data exploration technique and therefore the components are searched ignoring the presence of possible categories.

For a total of *m* initial variables, it is possible to generate *m* components, one by one, making the variance explained by each one a maximum of information, maintaining orthogonality between them, and passing through the centre of the variable space (see Figure 2. 9). The main objective is to simplify and reduce the number of initial variables of the original dataset, preserving the maximum of information ($p \leq m$). Then each sample is plotted in the new *p*-components space[2,3,6,10].

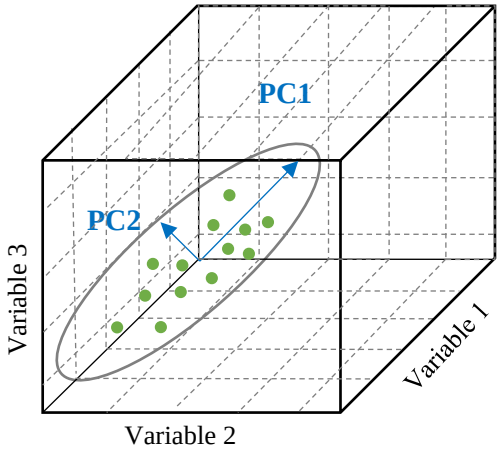
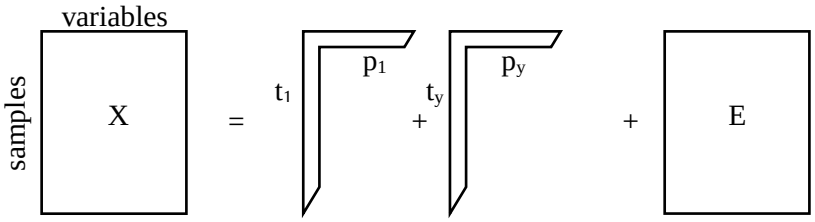


Figure 2. 9 Principal Components projection.

The first component (from now on abbreviated as PC1) will explain a given variance. The variability that is not explained by it will be divided into two, the variance explained (sum of squares of the distance of the points to the centroid of the component, divided by the number of points, figure dashed line) by the following component, and the residual variance (sum of squares of the distance of the points to the model, divided by the number of points, line of points perpendicular to the vector). Based on MD[2,3,6,7,10].

As a result, the PCA mathematical interpretation is shown in Equation 2 9[10].



$$X = t_1 p_1^T + t_2 p_2^T + \ldots t_y p_y^T + E$$

E: the amount of residual variance.

p: loading, describing the relationship between variables.

t: scores, describing the relationship between samples.

X: data matrix.

y: number of factors.

Equation 2 9.

The variance explained by the model (referred to as λ) constituted by these new vectors indicates the percentage of information contained in the model concerning the total information found in the data matrix, referred to as the cumulative explained variance for the first components (referred to as k)[3]. See Equation 2 10.

$$\frac{\sum_{p=1}^{k} \lambda_p}{\sum_{p=1}^{m} \lambda_p} x100\%$$

Equation 2 10.

The relation within the sample with the new axes is referred to as the scores, defined as the distance of the projected sample to the centre of the new axes, and to the loadings, defined as the relationship between the old variables included in the data matrix and the new axes, the principal components.

2.3.1.1.1. Outlier Detection. Q and T^2 Statistics

Any observation not fitted to the model can be defined as an outlier.

Three kinds of outliers can be described[6]:

- *X-sample outlier*.
- *Y-sample outlier*.
- *X-variable outlier*, referred to the property of the sample.

Different statistics can be involved in outlier detection.

Q residual is a measure of the size of the part of each sample not explained by the model. Hence, one Q value as a scalar is obtained for each sample. As a result, small Q residual means that the specified sample is well explained by the model, and the opposite for large Q residuals[2].

Hotelling's T^2 is the magnitude of the variation of each sample concerning the centre of the model. Hence, one T^2 value as a scalar is obtained for each sample. As a result, a small T^2 value means that the specified sample is close to the centre of the data (assuming that Q is not large), and the opposite for large T^2 residuals[2].

As a result, outlier samples can have a large influence on the PCA model. Hence, to check for outliers have to consider stray samples on the scores plot, and samples with very high Q, T^2, or both parameters.

2.3.2. Predictive Supervised Methods

The following chapter describes the supervised methods involved in this thesis.

2.3.2.1.Spectral Libraries

Different library levels are distinguished including the main library and the sub-libraries to unequivocally identify and qualify all the groups included. The number of sub-libraries is not limited[11–13].

For those materials which cannot be identified in sub-libraries can be assigned to specific classes if the class identity is enough.

The threshold value for each included group is selected. Moreover, the confidence level is selected commonly ≤95%. If the confidence level is set to 95%, then 5% of the total spectra will be identified as outliers[11].

Two different algorithms can be involved to develop the spectral library analysis based on Euclidean distance: the standard, where the spectral distance (from now on abbreviated as D) between test and calibration spectra is proportional to the area generated by them (see Equation 2 11), and the factorization methods, in which the spectra are presented as a linear combination of orthogonal factors (see Equation 2 12). The comparison between the test spectrum with the calibration is referred to as the spectral distance (the smaller the spectral distance, the more the spectra coincide).

$$D = \sqrt{\sum_{n} (a(n) - b(n))^2}$$

a: test spectrum.

b: calibration spectrum.

n: all included spectra.

Equation 2 11.

$$a = T_{1a} \cdot f_1 + T_{2a} \cdot f_2 + T_{ia} \cdot f_i$$

f: loadings as a factor spectrum.

T: scores coefficients.

Equation 2 12.

2.3.2.2.PLS-DA

The Partial Least Square Discriminant Analysis (from now on PLS-DA) is a discriminant method based on the PLS regression algorithm (which is described in chapter 2.4.1. of Part II).

PLS-DA aims to determine the category for a new observation as a linear regression between (the initial data matrix, for the predictor variables) and Y (the evaluated data matrix, for the predicted

variables). Is the application of PLS2 (multi-response) but instead of a quantitative response, a dummy matrix (transform the class vector in a matrix of zeros and ones) is described as a response that represents the I class membership[1,3] (see Figure 2. 10).

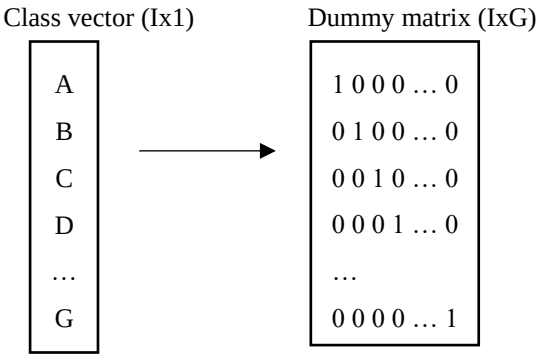


Figure 2. 10 Dummy matrix for PLS-DA.

2.3.2.3. Validation of a Classification Model

The validation of a classification model consists of establishing its recognition and prediction capabilities, as well as its stability and robustness[13].

The *confusion matrix*, defined as a square matrix with the dimensions of the number of categories, includes all the classification indices. Based on the defined matrix, the classification performance is evaluated. The recognition ability or the *non-error rate* (from now on NER) is evaluated as the percentage of objects in the training set that is correctly classified, considered the total included objects [3,14].

For the evaluation of the predictive capacity, the definition of an evaluation set is required, being objects whose belonging to the established groups or categories is known, but which are not included in the development data set. Predictability is the percentage of objects in the evaluation set correctly classified.

Finally, a model is stable if the removal of an object or its substitution does not change its recognition and prediction capabilities.

For binary classifications, only two classes are considered referred to as positive (from now on P) or negatives (from now on N). Hence, possible results are defined:[3,14]

- True positive (from now on TP): objects correctly classified as positives.
- True negative (from now on TN): objects correctly classified as negatives.
- False positive (from now on FP): object assigned as N and classified as P.
- False negative (from now on FN): object assigned as P and classified as N.

Specificity, sensitivity or true positive rate (from now on TPR), and precision or positive predictive value (from now on PPV) for the calculated binary model are mathematically calculated (see Equation 2 13, Equation 2 14, and Equation 2 15, respectively)[14], resulting in values between 0 and 1.

$$Specificity = \frac{TN}{FR + TN}$$

Equation 2 13.

$$Sensitivity/TPR = \frac{TP}{TP + FN}$$

Equation 2 14.

$$Precision/PPV = \frac{TP}{TP + FP}$$

Equation 2 15.

2.4.Quantification Tools

This chapter explains the theoretical background of a multivariate calibration technique.

The determination of a property Y of a data set from the set of measured parameters X is the goal of all quantitative analytical methods, which involve calibration and analysis of the data set Y to find its prediction. Initially, in the calibration stage, the correlation between X and Y is established. After making the calibration model for the measured parameters of X, the determination of the property of system Y of an unknown sample takes place[3]. See Figure 2. 11.

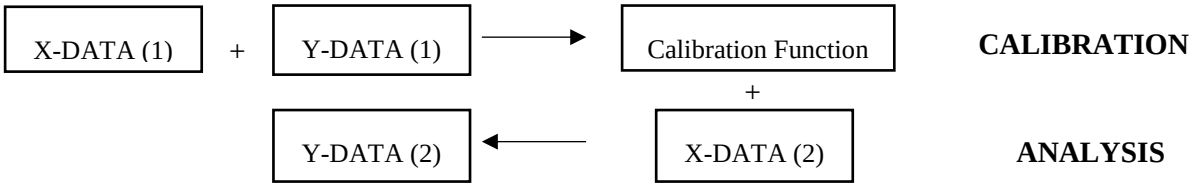


Figure 2. 11 Schematic procedure for the quantitative analysis.

There are two methods of setting up a calibration model: univariate (single variable) calibration, and multivariate calibration.

Univariate calibration is based on the correlation between a single variable in X and the reference value Y. The alternative, for the combination of a greater number of variables in X with the respective reference values Y, is multivariate calibration.

A good multivariate model can be built if the following conditions are met[3]:

- Data set Y contains as many variables as relevant sources of variance influence data set X.
- The variables included in the X data set contain variance correlated with all the variables in the Y block.

- The variables included in the data set X are poorly correlated with each other.
- The calibration set has been designed without correlation for the variables in the data set Y covering the calibration interval.

The following chapter describes the PLS Regression as a multivariate calibration method.

2.4.1. Partial Least Square Regression

PLS is related to Principal Component Regression (from now on PCR), capturing the maximum variance in X, and Multiple Linear Regression, capturing the maximum correlation between X and Y. As a result, PLS maximizes the covariance between X and Y, trying to assess both described premises[1–4,7,15].

As a result, the eigenvectors (factors or principal components) are generated reducing the data matrices. Hence, the original matrices are presented as the sum of products, scores vector with the loadings vector[2–4,15]. See Equation 2 16.

X data: $X=t_1p_1^T+ t_2p_2^T+t_3p_3^T+ \ldots t_Rp_R^T+F$

Y data: $Y=t_1q_1^T+ t_2q_2^T+t_3q_3^T+ \ldots t_Rq_R^T+G$

F: residual matrix of X.

G: residual matrix of Y.

p_i: loading vector of X.

q_i: loading vector of Y.

R: the number of factors.

t_i: scores vector.

T: the transpose matrix of the loadings.

Equation 2 16.

The PLS regression contains the mutually dependent factorization of the X and Y data. In the PLS method, identical score vectors for both data sets at the given factor numbers are assumed.

The PLS1 only takes the Y values of one analyte into account, all other data are interpreted as a disturbance. In the PLS2 algorithm, the Y values for all components in the system are considered for calibration.

2.4.2. Evaluation of Prediction Results

The predictive capacity of the calculated model must be evaluated. Such an evaluation is performed by predicting a certain number of samples with a known reference value with the chemometric model. The precision of the model is calculated considering the comparison of the reference values with the

predicted values. Two types of validation are possible: internal validation (referred to as cross-validation) and external validation (validation data set)[3,6].

Internal validation is characterized by using a set of calibration samples. The main stage requires the removal of certain samples from the calibration set and the generation of a new chemometric model with the samples that are still included in the set, determining the Y value of the removed samples. This is done iteratively until all calibration set samples have been removed from building this model. The predictive error of the model is calculated with the Root Mean Square Error of Cross Validation (from now on RMSECV)[2,3,6,7] that establishes the comparison of the reference value with the predicted value by the generated model, see Equation 2 17.

Another method to test the quality of the method is external validation. This method contains completely independent samples of the calibration set. The predictive error of the model is calculated with the Root Mean Square Error of Prediction (from now on RMSEP)[2,3,6,7] that establishes the comparison of the reference value with the predicted value by the generated model (see Equation 2 17.

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n_c}(\hat{y}_{ij} - y_{ij})^2}{n_c}}$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_p}(\hat{y}_{ij} - y_{ij})^2}{n_P}}$$

n_c: total number of samples for internal validation.

n_p: total number of samples for external validation.

$\hat{y}_{ij}$ measured value of sample *i*.

y_{ij}: predicted value of sample *i*.

Equation 2 17.

The number of components or latent variables (from now on LV) to be included must be determined to explain the greater variability associated with the parameter of interest based on the predictive model capacity.

Moreover, the Standard Error of Laboratory[16] (from now on SEL), referred to the reference method reproducibility, should be detailed and compared with the predicted equation (see Equation 2 18).

$$SEL = \sqrt{\frac{\sum_{i=1}^{n}(y_{1,i} - y_{2,i})^2}{n}}$$

n: number of included samples.

$y_{1/2}$: measured reference value of sample *i*, considering different laboratory conditions.

Equation 2 18.

2.5.Bibliography

(1) European Pharmacopoeia. 5.21. Chemometric Methods Applied to Analytical Data. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France January 1, 2023, pp 861–878.

(2) Roussel, S.; Preys, S.; Chauchard, F.; Lallemand, J. Multivariate Data Analysis (Chemometrics). In *Food Engineering Series*; Springer: New York, 2014; pp 7–59. https://doi.org/10.1007/978-1-4939-0311-5_2.

(3) Guillermo Ramis Ramos; M.ª Celia García Álvarez-Coque. *Quimiometría*; Síntesis, S.A.: Madrid, 2010.

(4) Jerome J.; Workman, Jr. NIR Spectroscopy Calibration Basics. In *Handbook of Near-Infrared Analysis*; Donald A. Burns, Emil W. Ciurczak, Eds.; CRC Press, 2008; Vol. 35, pp 123–149.

(5) Miller, J. N.; Miller, J. C. Significance Tests. In *Statistics and Chemometrics for Analytical Chemistry*; James N Miller, Jane C Miller, Eds.; Pearson; pp 37–74.

(6) Charles E. Miller. Chemometrics in Process Analytical Technology (PAT). In *Process Analytical Technology. Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*; Bakeev, K. A., Ed.; John Wiley & Sons, Ltd: USA, 2010; pp 353–434.

(7) Maspoch, S.; Coello, J. Calibración Multivariable. In *Temas Avanzados de Quimiometría*; Blanco, M., Cerdà, V., Eds.; Universitat de les Illes Balears: Illes Balears, 2007; pp 239–286.

(8) Rinnan, Å. Pre-Processing in Vibrational Spectroscopy-When, Why and How. *Analytical Methods* **2014**, *6* (18), 7124–7129. https://doi.org/10.1039/c3ay42270d.

(9) Blanco Romía, M.; Alcalá Bernárdez, M. NIR Spectroscopy in Pharmaceutical Analysis: Off-Line and At-Line PAT Applications. In *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*; Bakeev, A. K., Ed.; John Wiley & Sons, 2010; pp 463–491.

(10) Bro, R.; Smilde, A. K. Principal Component Analysis. *Analytical Methods* **2014**, *6* (9), 2812–2831. https://doi.org/10.1039/c3ay41907j.

(11) Miró Vera, A. Y.; Alcalà Bernàrdez, M. Near-Infrared Spectroscopy in Identification of Pharmaceutical Raw Materials. In *Encyclopedia of Analytical Chemistry*; John Wiley & Sons, Ltd, 2017; pp 1–19. https://doi.org/10.1002/9780470027318.a9619.

(12) Blanco, M.; Romero, M. A. Near-Infrared Libraries in the Pharmaceutical Industry: A Solution for Identity Confirmation. *Analyst* **2001**, *126* (12), 2212–2217. https://doi.org/10.1039/b105012p.

(13) Blanco, M.; Romero, M. A.; Alcalà, M. Strategies for Constructing the Calibration Set for a Near Infrared Spectroscopic Quantitation Method. *Talanta* **2004**, *64* (3), 597–602. https://doi.org/10.1016/j.talanta.2004.03.027.

(14) Ballabio, D.; Todeschini, R. Multivariate Classification for Qualitative Analysis. In *Infrared Spectroscopy for Food Quality Analysis and Control*; Elsevier Inc., 2009; pp 83–104. https://doi.org/10.1016/B978-0-12-374136-3.00004-3.

(15) Wold, S.; Sjostrom, M.; Eriksson, L.; Sweden˚˚, S. PLS-Regression: A Basic Tool of Chemometrics. *Chemometrics and Intelligent Laboratory Systems* **2001**, *58*, 109–130.

(16) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. London, UK, January 20, 2012.

3. MULTIVARIATE PROCEDURE DEVELOPMENT

The multivariate procedure development involves analytical chemistry, statistics, and data analysis to build a model based on multiple variables or factors. Those are used to analyze complex data sets to identify patterns and relationships among the selected variables.

The development of a multivariate procedure involves several steps, including the scope definition, data collection, and processing, model selection, choosing the appropriate multivariate model, model fitting, estimating the parameters of the selected model based on the processed data, and the model evaluation, assessing the quality of the model based on specific metrics such as the goodness-of-fit and the prediction accuracy.

The guidelines involved and the critical steps involved in developing a multivariate method in the pharmaceutical industry are detailed in this section.

3.1. Guidelines

The following sections are based on the application of the EMA and FDA guidelines for new submissions of multivariate methods to the regulatory agencies of the pharmaceutical industry. The guidelines included are: EMA "*Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submission"*[1], and the FDA guideline *"Development and Submission of Near Infrared Analytical Procedures Guidance for Industry"*[2].

In addition, the general text 5.21[3] of the European Pharmacopoeia dedicated to chemometric methods, chapter 2.2.40.[4], dedicated to the NIRS, and chapter 2.2.48.[5], dedicated to Raman spectroscopy, are considered.

3.2. Practical Procedure

There are specific steps involved during multivariate method development, which can be iterative and may require finetuning to obtain optimal results[1]. These are summarized in Figure 3. 2.

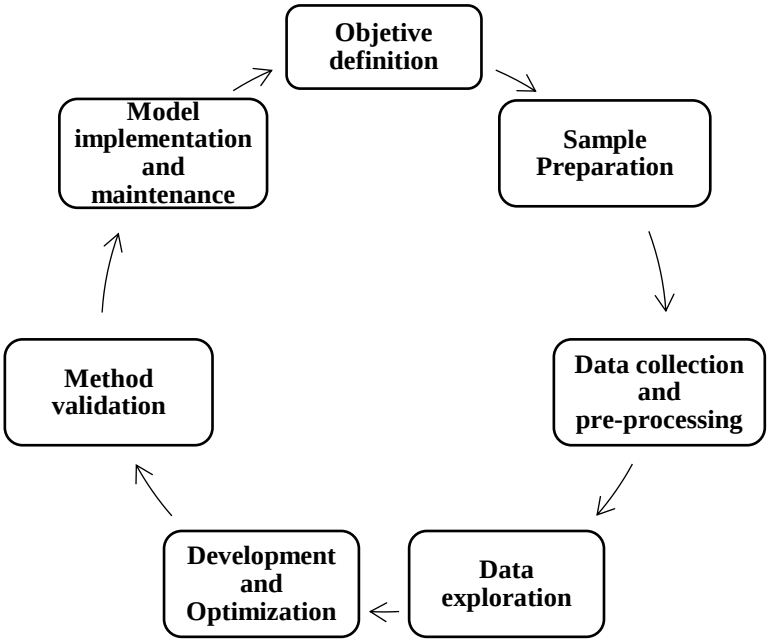


Figure 3. 2 Steps for multivariate model construction[1,6,7].

3.2.1. Objective Definition

The purpose of the multivariate analysis should be clearly stated, including the expected outcomes, goals, and any potential limitations or challenges. Hence the method, model, and procedure should be detailed[1].

- Method: summary details for the equipment involved, regarding critical parameters of measurement, including sample presentation regarding analysis strategy (online, inline atline, or offline measurements (see Figure 3. 3).

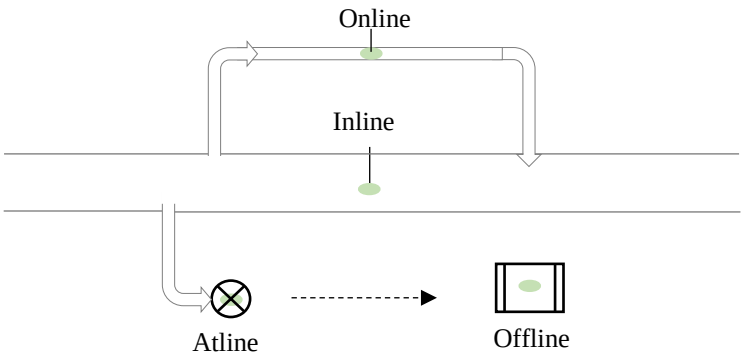


Figure 3. 3 Strategies for sample presentation during measurement.

- Model: the chemometrics involved to guarantee the correlation between the equipment signal and the analyte of interest.
- Procedure: the suitability of the selected technique for the intended purpose based on previous suitability studies, if considered (both method and model).

A risk assessment should be considered considering those factors that may interfere with the proper development of the procedure, based on the ICH Q9 guideline[8].

3.2.2. Sampling Preparation

The preparation of representative samples is considered a critical step to obtain a calibration model suitable for future predictions [9–11].

Three specific data sets are required[1,12]:

- *Calibration set*: a set of data used to calibrate the model. The preparation of representative calibration samples is considered a critical step since they must include the physicochemical variability of the product to be analyzed (for qualitative and quantitative methods, respectively). The three main strategies for the preparation of the involved samples in the calibration model are detailed below[9,11]:
    - *Laboratory samples*: direct mixing of the formulation components to expand the concentration range of the analyte of interest.
    - *Synthetic pilot plant samples*: pilot plant samples preparation to reproduce the manufacturing process.
    - *Under and overdosing production samples*: doping production samples with excipients and the active ingredient to expand the concentration range of the analyte of interest.

    The physical variability of the product must be covered by also including production samples in the calibration set[9].

- *Calibration test set*: a set of data used to evaluate the performance of the model, including model optimization and internal validation. If enough samples are available for calibration, these will be divided into a calibration set and optimization (calibration test set) through a manual strategy (score graph in the exploratory analysis) or mathematical algorithms such as Kennard-Stone (in the factorization space, select the most different samples among them dividing them into calibration and optimization). Otherwise, the same samples as those included in the calibration set are used in this case to perform the cross-validation
- *Validation test*: a set of independent new data not used in the model development process involved during the model performance evaluation. Hence, the test set should cover the calibration range of the multivariate model including lifecycle product and process variations[12].

3.2.3. Data Collection and Pre-processing

Data collection from relevant sources of variability is mandatory involving the qualified equipment based on Annexe 15 from the GMP regulation[13]. In addition, in this step, a mathematical method for data pre-processing is selected to clean the data from noise and transform it to be prepared for analysis, transforming the instrument signal into information[14,15]. See chapter 2.2. from Part II.

Of central importance in this context are the quality of the measurement parameters and the reference value which is normally determined by a reference analytical method.

This step may involve handling missing values and/or outliers[1].

3.2.4. Data Exploration

The most relevant variables for the analysis may be chosen. Hence, an exploratory analysis, based primarily on PCA analysis, may be involved to gain initial insights into the data, identify trends and patterns, and detect any potential issues.

This step is crucial for reducing the dimensionality of the data and improving model interpretability and performance in detecting possible outliers (see chapter 2.3.1. from Part II).

3.2.5. Development and Optimization

In multivariable classification methods, the workspace is initially selected, followed by the mathematical algorithm. Finally, the classificatory capacity of the model is evaluated, guaranteeing selectivity[1] (see chapter 2.3.2. from Part II).

In multivariate quantification methods for spectroscopic techniques, as long as the Beer-Lambert law is followed, absorbance spectra present a linear relationship between absorbance values and data concentration commonly involving the PLS algorithm (see chapter 2.4. from Part II).

3.2.6. Procedure Validation

Analytical method validation is one type of required validation in pharmaceutical manufacturing.

The suitability of the chosen data processing methods and the variables for the given task is evaluated during the validation step.

Internal and external validation may be involved as an optimization/cross-validation or test set validation, respectively. In the first step, the best model parameters are obtained by optimization or cross-validation, evaluating the RMSECV value (see chapter 2.4.2. from Part II). In the second step,

a completely external validation set is defined. The method is set up with the predefined parameters, and validated by a test set validation evaluating the RMSEP value (see chapter 2.4.2. from Part II).

The comparison between the SEP and the SEL should be included[1] (see chapter 2.4.2. from Part II).

External validation parameters are detailed in Figure 3. 4 and Figure 3. 5, based on the EMA[1] and the ICH guideline[16,17].

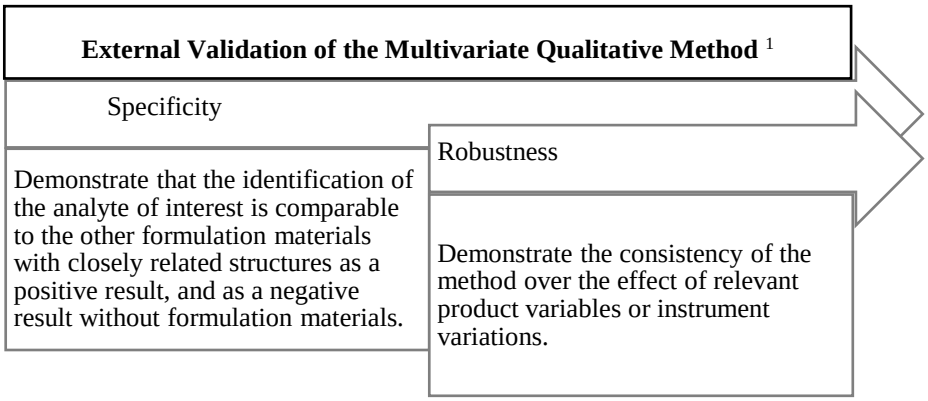
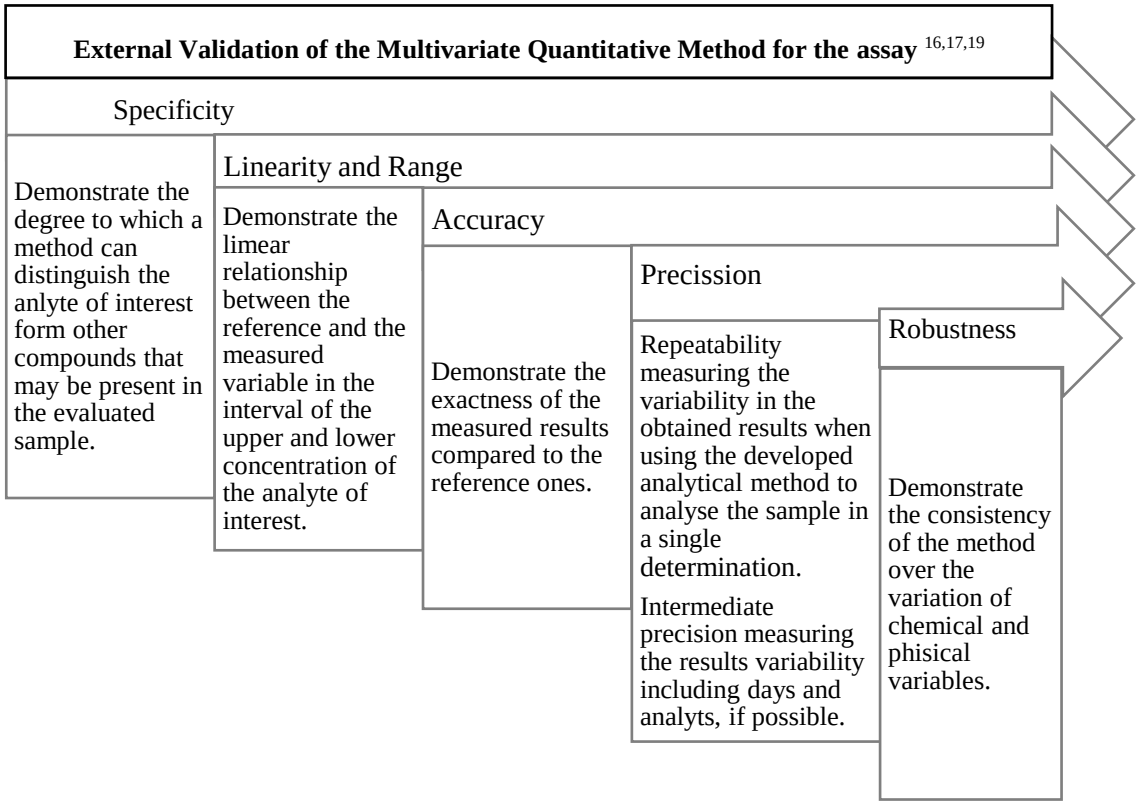


Figure 3. 4 External validation parameters for a qualitative multivariate method.



Limit of detection and quantification (defined as the lowest detected and quantified amount of the analyte, respectively) are not required.

Figure 3. 5 External validation parameters for a quantitative multivariate method.

3.2.6.1. Analysis Results

In this last step, the optimized multivariate method is used to analyze new samples. The goodness of prediction is evaluated involving different parameters such as the MD or the spectral residue (for qualitative or qualitative and quantitative multivariate methods, respectively).

- MD: the spectral structures between the calibration set and the analyzed sample are compared. The larger the difference between them, the larger the value of the MD. See chapter 2.3. from Part II for statistic calculation.
- Spectral residue: the residual is the difference between the reference data and the predicted model data applied to the spectral and concentration data. The residual matrix of the concentration data describes the part of the components of the calibration data set which cannot be explained by factorization (see Equation 3. 1).

$$Res_i = Y_i - Y_i^{prediction}$$

Equation 3. 1.

Moreover, for the analyte spectrum, the difference is calculated between the measured spectrum x_i and the expected spectrum from the factorization, s_i. See Equation 3. 2.

$$Spectral\ residue_i = \sqrt{\sum^{n}(x_i - s_i)^2}$$

n: number of samples.

Equation 3. 2.

Hence, the analysis of the results offers information about the value of the new sample and the outlier determination.

3.2.7. Procedure Implementation and Maintenance

The most important task is the selection of a representative calibration data set. Therefore, maintenance of the chemometric model should be involved once the procedure has been developed and validated for routine analysis[1,18]. The change may be covered by existing validation, in which case no further validation is required, or not, resulting in the re-validation of a new model[19]. See Figure 3. 6.

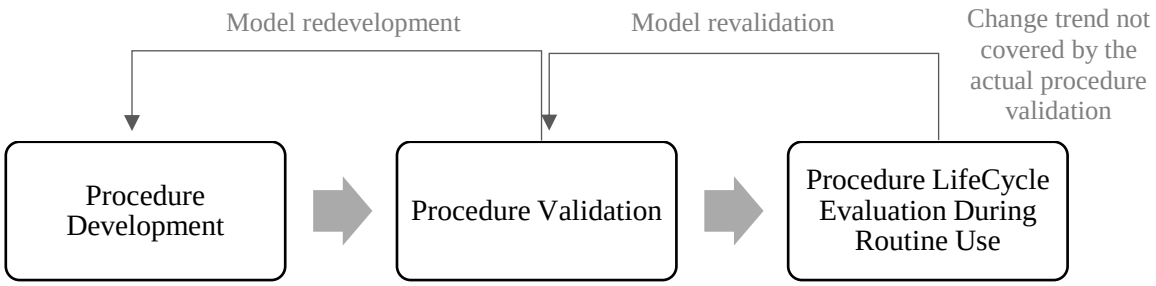


Figure 3. 6 Procedure lifecycle.

Commonly, residual plots are used to diagnose and improve the statistical model performance. Hence, it allows the identification of potential variations in the model, such as heteroscedasticity (the variance of the residuals is not constant across the range of the independent variables), non-linearity (the relationship between the independent and the dependent variables is not linear), and outliers (data points that lie far away from the other data points having a significant impact on the regression model).

3.3.Bibliography

(1) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. London, UK, January 20, 2012.

(2) Food and Drug Administration. Development and Submission of Near Infrared Analytical Procedures Guidance for Industry. Silver Spring, MD, August 2021, pp 1–28.

(3) European Pharmacopoeia. 5.21. Chemometric Methods Applied to Analytical Data. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France January 1, 2023, pp 861–878.

(4) European Pharmacopoeia. 2.2.40. Near-Infrared Spectroscopy. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France, January 2023, pp 76–81.

(5) European Pharmacopoeia. 2.2.48. Raman Spectroscopy. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France, January 2023, pp 101–103.

(6) Blanco, M.; Villarroya, I. NIR Spectroscopy: A Rapid-Response Analytical Tool. *Trends in Analytical Chemistry* **2002**, *21* (4), 240–250.

(7) Blanco Romía, M.; Alcalá Bernárdez, M. NIR Spectroscopy in Pharmaceutical Analysis: Off-Line and At-Line PAT Applications. In *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*; Bakeev, A. K., Ed.; John Wiley & Sons, 2010; pp 463–491.

(8) International Council for Harmonisation. ICH Guideline Q9 on Quality Risk Management. *EMA/CHMP/ICH/24235/2006*. September 2015.

(9) Blanco, M.; Romero, M. A.; Alcalà, M. Strategies for Constructing the Calibration Set for a Near Infrared Spectroscopic Quantitation Method. *Talanta* **2004**, *64* (3), 597–602. https://doi.org/10.1016/j.talanta.2004.03.027.

(10) Blanco, M.; Peguero, A. Influence of Physical Factors on the Accuracy of Calibration Models for NIR Spectroscopy. *J Pharm Biomed Anal* **2010**, *52* (1), 59–65. https://doi.org/10.1016/j.jpba.2009.12.009.

(11) Blanco, M.; Bautista, M.; Alcalá, M. Preparing Calibration Sets for Use in Pharmaceutical Analysis by NIR Spectroscopy. *J Pharm Sci* **2008**, *97* (3), 1236–1245. https://doi.org/10.1002/jps.21105.

(12) Romañach, R. J. Independence and Dependence in Calibration: A Discussion of FDA and EMA Guidelines. *Am Pharm Rev* **2016**.

(13) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 30, 2015, pp 1–16.

(14) Rinnan, Å.; Berg, F. van den; Engelsen, S. B. Review of the Most Common Pre-Processing Techniques for Near-Infrared Spectra. *TrAC - Trends in Analytical Chemistry*. November 2009, pp 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007.

(15) Faber, N. M. Multivariate Sensitivity for the Interpretation of the Effect of Spectral Pretreatment Methods on Near-Infrared Calibration Model Predictions. *Anal Chem* **1999**, *71* (3), 557–565. https://doi.org/10.1021/ac980415r.

(16) International Council for Harmonisation. ICH Guideline Q2(R2) on Validation of Analytical Procedures. *EMA/CHMP/ICH/82072/2006*. March 31, 2022.

(17) International Council for Harmonisation. ICH Q2 (R1). Validation of Analytical Procedures: Text and Methodology. *CPMP/ICH/381/95*. June 1995.

(18) Roussel, S.; Preys, S.; Chauchard, F.; Lallemand, J. Multivariate Data Analysis (Chemometrics). In *Food Engineering Series*; Springer: New York, 2014; pp 7–59. https://doi.org/10.1007/978-1-4939-0311-5_2.

(19) Oona McPolin. *Validation of Analytical Methods for Pharmaceutical Analysis*; Mourne Training Services: UK, 2009.

PART III

RESULTS AND DISCUSSION

Experimental Chapter 1. NIR Spectroscopy in Starting Material Identification

1. STARTING MATERIALS IDENTIFICATION

The verification of the correct SM reception, including the active substance (active pharmaceutical ingredient; from now on abbreviated as APIs) and excipients, by a discriminatory method, is mandatory [1–4] in the pharmaceutical industry to guarantee compliance with the specifications for quality product and safety. Once the identification was verified for each SM container, the final quality bath could be evaluated by a representative sample previously established in the sampling plan[4].

Package checking is the starting point for SM verification[5–8]. Then, the SMs are transported to a sampling room to prevent cross-contamination. Then the container is sampled following one of these strategies: 1) one sample per container received per batch, or 2) sampling the square root of the number of containers plus one ($\sqrt{N}+1$; where N is the number of containers per batch received). After sampling, the containers are properly and securely closed, avoiding contamination and stability problems, before leaving the sampling booth. Finally, the samples are handed over to quality control for identity verification analysis.

Commonly classical (including precipitation or colorimetric experiments) or instrumental (such as IR or HPLC analytical methods are involved. However, there requires expert knowledge of the technique on the part of the analyst, high time-consuming, or high-waste production[9]. NIR spectroscopy has been used for this purpose for some years[10]. Also, if you want to reduce the higher cost of service and lead time for SM compared to sampling, these rapid spectroscopic techniques allow you to move SM identity tests from the lab to the warehouse allowing the identification of the set of received batches for the set of SM except sucrose[6].

The identity testing for the SMs involving the spectroscopic techniques is evaluated by the spectral similarity between the measured spectrum and the reference library spectrum considering an established threshold[10]. The classification includes the identification and verification or qualification of the SM. Identification is determined by evaluating the greatest spectral similarity between the SM spectrum and each reference library material. As a result, the chances of the identification results are equal to the number of references SMs included in the spectral library. Finally, verification implies only the rejection or acceptance of the identity result.

During this section the applicability of NIR spectroscopy for the construction and maintenance of spectral libraries as an identity test in the pharmaceutical industry is detailed. In addition, the identification of the materials is guaranteed with any equipment that you want to use for this purpose.

1.1. Starting Materials and Methods

This section describes the different analyzed samples, the instruments involved in data acquisition, and the data processing.

1.1.1. Sample Description

A total of 237 SMs were included in the spectral library, including historical ones and newly identified batches, if necessary, with the accepted alternative reference method. The variability included in the spectral libraries depended mainly on the frequency of arrival of the materials at the sampling warehouse. This factor is directly related to the amount of SM used for the preparation of the formulation and the final volume of the finished product prepared with it. Different suppliers and batches were included.

1.1.2. Near-Infrared Method

Data acquisition was performed with a benchtop FT-NIR MPA, and an FT-NIR MPA II (from now abbreviated as MPA and MPA II, respectively) both fitted with a solid probe for solid samples and a transmission sample compartment for liquid samples. Key parameters for NIR spectra include the number of scans, set at 32 scans, and the resolution of the intervals over the selected wavelength range, all of them were previously optimized during the spectral library development (from 12500 to 4000 cm^{-1}).

1.1.3. Software and Data Processing

The OpusLab interface (Bruker Optics) was used for data acquisition, and the OPUS workspace including IDENT and QUANT packages were used for qualitative and quantitative methods, respectively, including the development and validation[11].

1.2. Maintenance of NIR Identification Spectral Libraries

The implementation of the NIR spectrometer in the sampling room, located in the materials reception warehouse, has allowed the identification and qualification of all the containers in which each batch of SM is divided, which is received at LMSA, quickly (in situ) and reliably, allowing any defective product to be discarded[6,12]. The strategy followed for the identification and qualification of SM batches with both techniques is shown in Figure 1. 7.

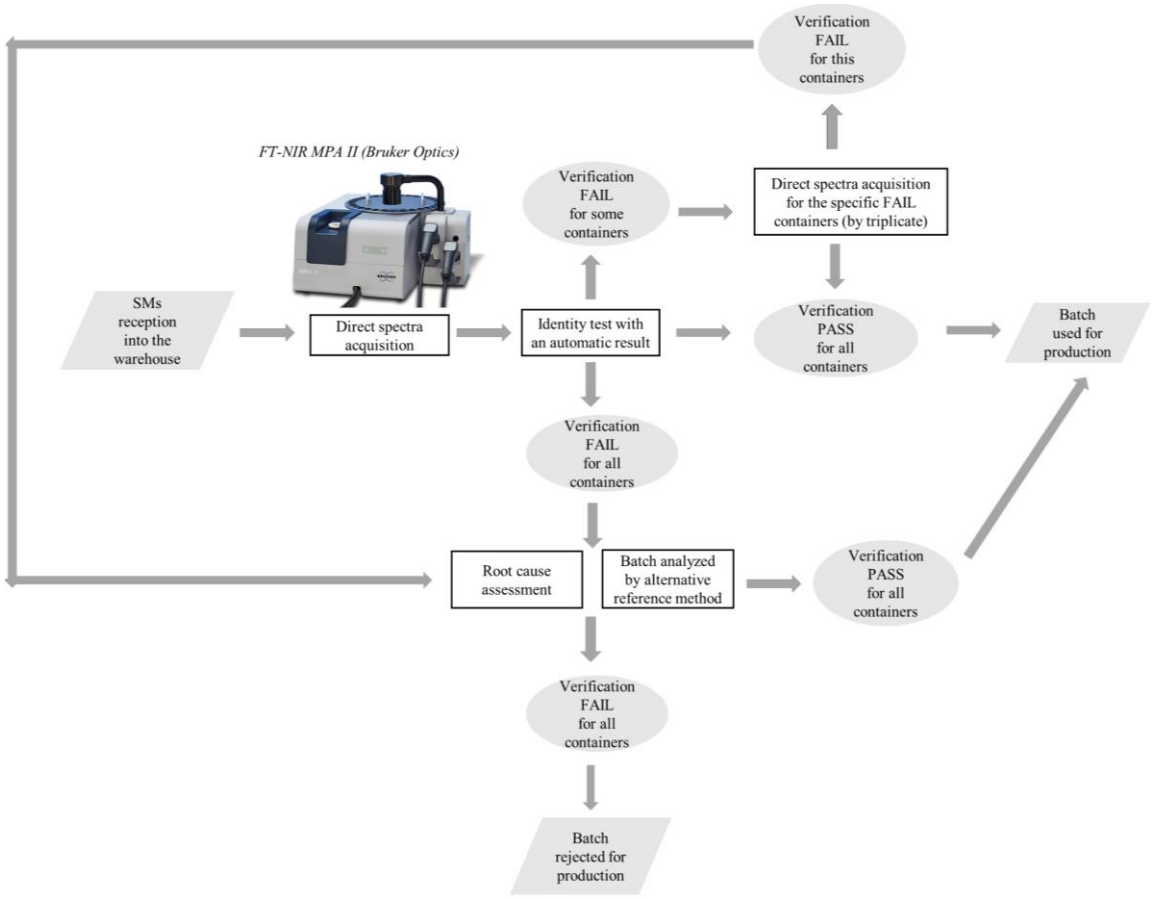


Figure 1. 7 Workflow for an identity SM batch.

A NIR spectral library has been calibrated and validated using the cascade identification strategy that allows the identification and qualification for all SM included[13]. This NIR library makes it possible to differentiate chemically identical products with different particle sizes, crystalline shapes, or polymer chain lengths.

Once a new source of variability is detected in the model prediction, library maintenance has to be developed to guarantee the usefulness of the library over time. The maintenance was developed after one year of the last version library implementation. This action was motivated by common situations in the pharmaceutical industry such as the need to include new batches of SM previously included in the library due to the greater variability associated with them, and the need to include new SM [10].

A description of the spectral library maintenance for the SMs identity test is detailed in the following sections.

1.2.1. NIRS Identification Library

The NIR library's maintenance was developed one year after the implementation of the last version of the library. During this period, the OOS obtained, meaning SMs batches were not correctly

identified by NIR but correctly identified by the alternative reference method, represented 5.5% of the overall number of SMs identified by NIR.

The objective was to include a total of 455 and 1344 new spectra for calibration and validation data sets, respectively, for a total of 36 SMs included in the previous version, and two new SMs.

1.2.1.1.Library Maintenance. Development and Validation

The identification spectral library was structured in a general library including all the SMs tested in a preliminary identification. This first library was developed with a wide spectral range where there was a significant variation in the NIR signal and a derivative spectral pre-treatment that applied instrumental noise smoothing before NIR signal derivation followed by vector normalization. The algorithm of this first identification was turned by distance, adjusting the identification threshold after careful analysis of the calibration and internal validation samples. There was no limit on the number of SMs that may be included in the general library.

During this first step of the calibration, the sub-cascading spectral library was involved as a qualification strategy for SMs misclassification resolving ambiguities including different classification levels[10]. There was no limit to the sub-library number, composition, and levels. The different spectral ranges were selected specifically for each identification level. Moreover, spectral pre-treatment was selected too (a combination of derivatives with normalization of the spectra). Both parameters were optimized to maximize the differences between the SMs. The discriminant algorithm used is known as factoring or PCA, allowing the highlighting of the subtle spectral difference between highly similar spectral classes. The selected combination must provide a complete resolution of spectral ambiguities. The identification criterion was established using measuring the distance in the PC space or residual variance. The threshold limit allows unequivocally qualifying each spectrum of the library solving and correcting the spectral ambiguities that may occur. Finally, groups that cannot be separated into further sub-libraries were assigned to common classes.

Each library/sub-library included a variable number of calibration spectra used for calibration and internal validation as positive samples. The internal validation showed that all the calibration spectra belong to each of their corresponding folders of each SM and there were no ambiguities in the calibration identification based on a hit quality value, resulting in a lower value than the threshold distance previously established[10].

Finally, an external validation of the library was developed including spectra not used during the calibration/ internal validation. The calculated distance had to be less than the established threshold for correct identification, a variable for each of the libraries/sub-libraries generated.

Two specific parameters were validated: specificity and robustness. The specificity was evaluated based on the selectivity parameter (from now abbreviated as Sel.), which compares one spectra group with the adjacent spectrum in the same group to see which clusters overlap. The selectivity compared SM1 with SM2 as detailed in Figure 1. 8.

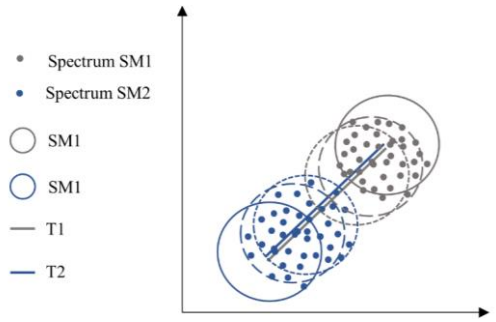


Figure 1. 8 Selectivity graphical calculation[14].

The selectivity calculation was based on the ratio of the distance between average spectra, and the sum of the threshold values for each cluster SM (being the values T1 and T2 in Figure 1. 8, referred to as the cluster ratio). As a result, three different situations can occur:

- Sel.>1: where the evaluated clusters SMs were completely separated.
- Sel.=1: where the evaluated clusters SMs were in contact.
- Sel.<1: where the evaluated clusters SMs were overlapped.

Robustness was evaluated by including the largest batches possible in the external validation, including the expected variability for each SM.

1.2.1.2.Library Maintenance. Results and Discussion

The general library included a total of 237 SMs. The algorithm involved was the factorization and the first derivative S-G with 17 points of window width, followed by vector normalization as data pre-treatment. The spectral range was specified between 9002 and 4420 cm^{-1}. The rest of the interval ranges (from 12500 to 9002 and from 4420 to 4000 cm^{-1}) were discarded because they did not present useful information to improve the capacity of the identification method. The threshold was established with a confidence level of around 97 percent. As a result, a total of 100 SMs were identified unequivocally.

Thereafter, a series of cascade sub-libraries were set for the rest of the 137 SMs. 18 sub-libraries were assigned for the second level, 13 for the third level, 8 for the fourth level, 5 for the fifth level, and 2 for the sixth level. At each of the specific levels, a total of 77, 30, 14, 12, and 4 SMs were identified without ambiguities, respectively. In the polysaccharides sub-library (starches and celluloses), long hydrocarbon chains, sucrose, lactose, and lacquers were included.

Finally, the construction of 6 specific classes was necessary due to the impossibility of resolving ambiguities between the materials included in each class using sub-libraries (groups of SMs that are not equivalent but cannot be separated without ambiguities).

The general structure for the spectral library is shown in Figure 1. 9.

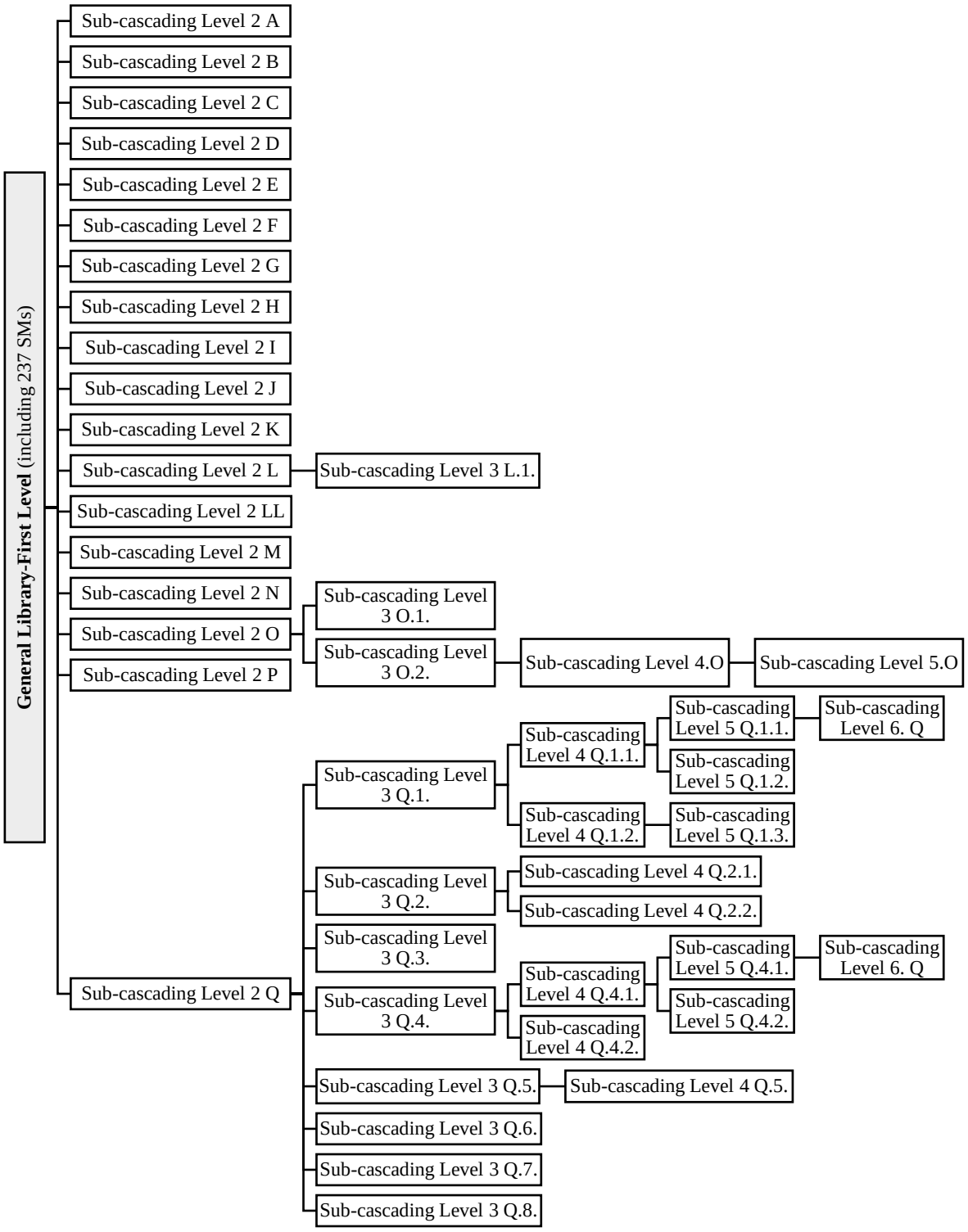


Figure 1. 9 Whole library tree for SMs NIR identity test.

As an example, the sub-cascading spectral library, included in the second level assigned as O in Figure 1. 9, and the class assignation was detailed for the Lacquer and the Hypromellose SMs.

The included compounds are detailed in Table 1. 2.

Table 1. 2 SMs description included in the sub-library level 2 O.

Component Description	Specifications
Lacquer Opadry OY-30923	Film-coated blue. Defined as a mixture of hypromellose, polyethylene glycol/macrogol, titanium dioxide, and blue carmine aluminium lake.
Lacquer Opadry OY-S-7163	Film-coated white. Defined as a mixture of hypromellose, polyethylene glycol/macrogol, titanium dioxide, talc, and simethicone emulsion.
Lacquer Opadry OY-S-33001	Film-coated orange. Defined as a mixture of hypromellose, polyethylene glycol/macrogol, titanium dioxide, talc, yellow aluminium lake, and erythrosine aluminium lake.
Lacquer Opadry OY-S-38913	Film-coated white. Defined as a mixture of hypromellose, polyethylene glycol/macrogol, titanium dioxide, and talc.
Lacquer Opadry 02F33175	Film-coated orange composed of a mixture of hypromellose, polyethylene glycol/macrogol, titanium dioxide, talc, and yellow aluminium lake.
Lacquer Opadry 02F38986	Film-coated white. Defined as a mixture of hypromellose, polyethylene glycol/macrogol, and titanium dioxide.
Lacquer Opadry Y-1-7000	Film-coated white. Defined as a mixture of hypromellose, titanium dioxide, and polyethylene glycol/macrogol.
Hypromellose 3cP[15]	Defined as hydroxypropyl methylcellulose with 3 centipoises (cP) of viscosity. A viscoelastic polymer that is used as an excipient due to its ability to be suspended as insoluble microscopical particles in a substance when in solution.
Hypromellose 5cP [15]	Defined as hydroxypropyl methylcellulose with 5 centipoises (cP) of viscosity. A viscoelastic polymer that is used as an excipient due to its ability to be suspended as insoluble microscopical particles in a substance when in solution.

Sub-cascading library Level 2 O

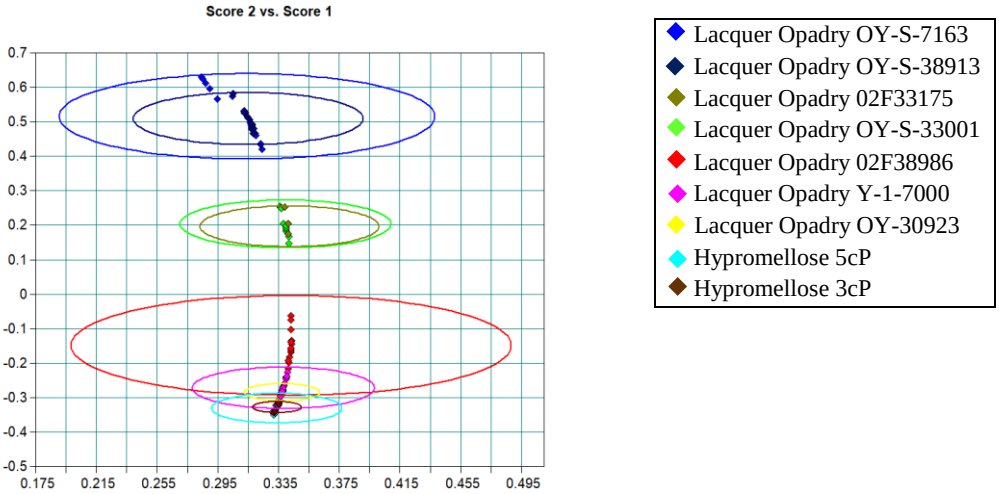
The scores plot for the presented sub-library, including the two first PCs is shown in Figure 1. 10. A cluster trend between the included SMs was obtained after applying the first derivative S-G with 25 points of widow width, followed by vector normalization, and selecting the specific wavelength range between 10854 and 10306, and 7868 and 6025 cm^{-1} for the second level sub-cascading library (see Figure 1. 10). For the rest of the sub-cascading libraries, different pre-treatments and wavelength ranges were selected to obtain an unequivocal qualification for the included SMs:

- Sub-cascading Level 3.O.1.: second derivative S-G including 21 points of window width, followed by vector normalization. Interval range from 8462 to 4088 cm^{-1}.
- Sub-cascading Level 3.O.2.: second derivative S-G including 17 points of window width, followed by vector normalization. Interval range from 8053 to 7791 cm^{-1}.

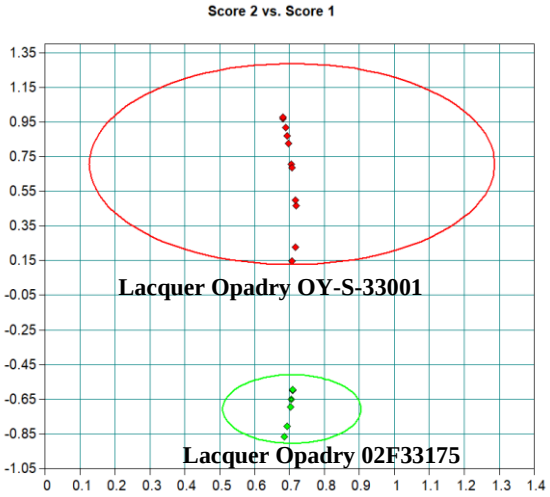
- Sub-cascading Level 4.O.: first derivative S-G including 9 points of window width, followed by vector normalization. Interval range from 7066 to 5909 cm^{-1}.
- Sub-cascading Level 5.O.: second derivative S-G including 25 points of window width, followed by vector normalization. Interval range from 12204 to 8964 cm^{-1}.

The creation of two classes was necessary for 2 specific lacquers (Lacquer Opadry OY-S-38913, and Lacquer Opadry OY-S-7163), and the hypromellose SMs due to their composition and low batch variability.

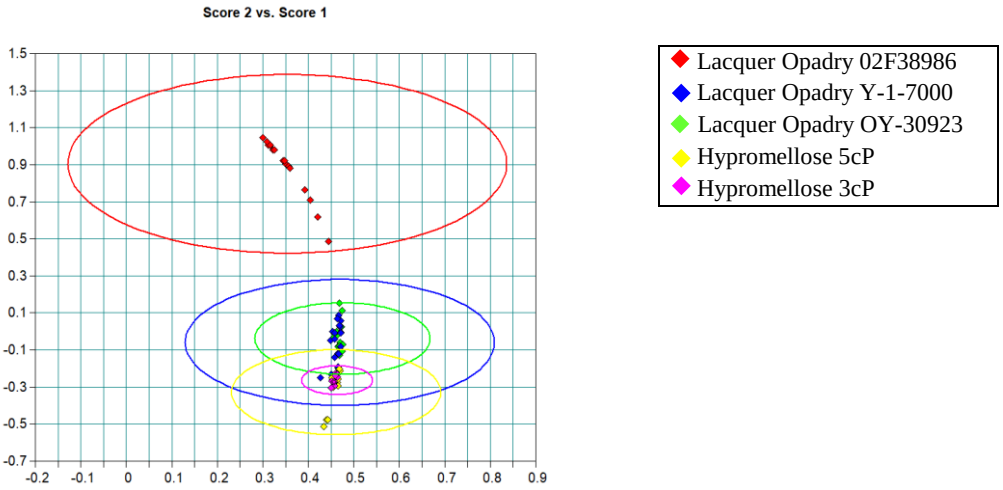
Sub-cascading Level 2 O.



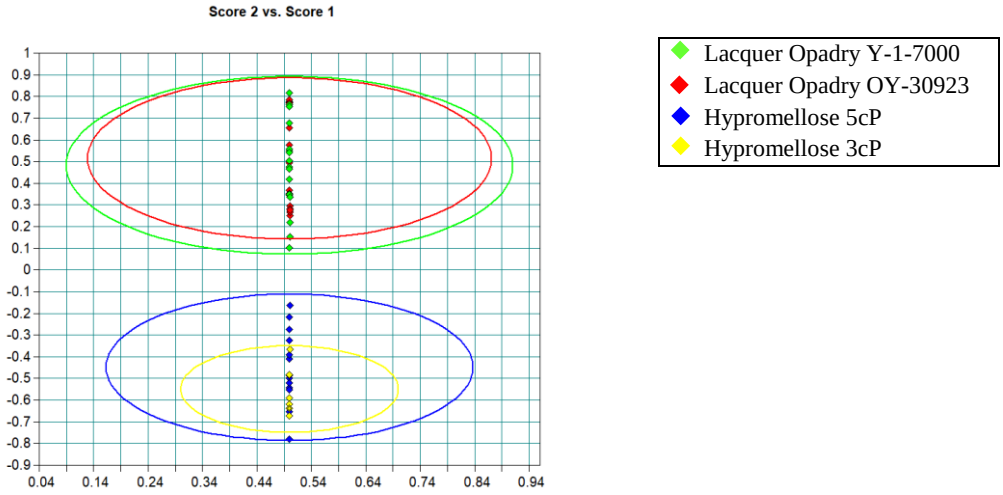
Sub-cascading Level 3. O.1.



Sub-cascading Level 3. O.2.



Sub-cascading Level 4. O.



Sub-cascading Level 5. O.

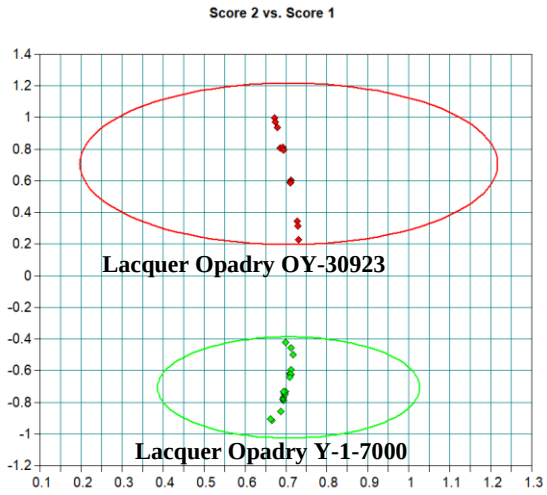


Figure 1. 10 Sub-cascading library for level 2 O.

Finally, the selectivity histogram for the specificity evaluated parameter is shown in Figure 1. 11.

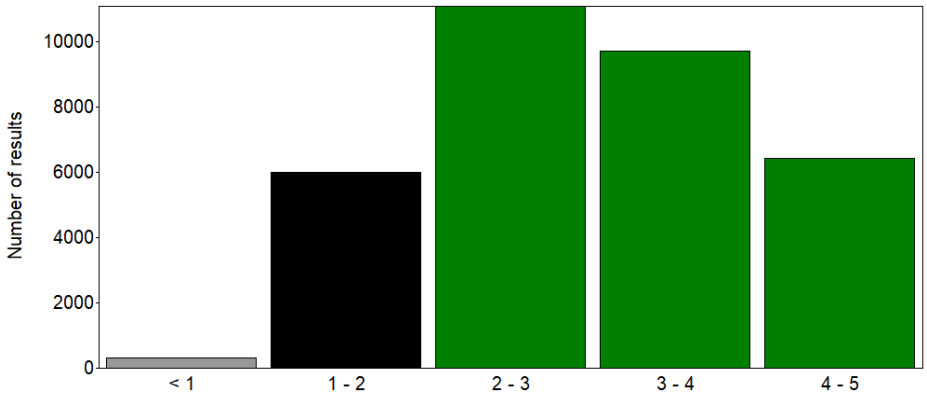


Figure 1. 11 Selectivity histogram for the NIR spectral library.

Despite the objective being to achieve a selective value higher than 1 for each SM, less than 1% of the selectivity evaluated for all the SMs presented a value lower than 1 (grey colour in Figure 1. 11). The groups of each SMs were overlapped but the spectra of the respective SMs were not in the overlap area (Figure 1. 8). The 12.5% represented a selectivity value equal to 1 (black colour in Figure 1. 11), and the rest of the calculations showed a selectivity greater than 1 (green colour in Figure 1. 11).

1.2.1.3.Risk Assessment. Transferability Study for the Identification Library Between NIRS Equipment

The NIR spectra for the same sample can be affected by three classes of factors: (1) changes in the physical and/or chemical properties of the samples, (2) changes in the environmental conditions where the instrument is located, thus affecting its components, and (3) changes in the instrumental response, which may appear due to the use of other equipment to acquire the data to be predicted in the already created procedure[16]. These spectral variations can lead to the failure of previously developed and validated multivariate qualitative or quantitative procedures. In some cases, the response of the new sample is not greatly affected, and the existing model can be used without having to make any corrections, concluding in direct transferability. For evaluating the spectral variations, a spectral equivalence study can be proposed implementing the exploratory analysis by reduction of the dimensionality of the variables, based on a PCA, for the spectral data acquired in parallel with the different NIR spectroscopic equipment. If, on the contrary, this does not happen, some strategies have to be used to eliminate the identification/prediction failure, resulting in the application of a calibration transfer of the model, choosing between the following strategies[16–19]:

- Model update, including spectra acquired with the second equipment to the library developed with the first equipment.

- Partial Direct Transfer (from now on PDS) standardization method where a succession of linear transformations of the spectrum acquired with the second equipment will cause its spectral response to coincide with that of the spectra acquired with the first equipment. As a result, the spectra of the products could be acquired with both equipment for a unique library.

If all this is not possible, it is worth considering a new development and validation of the procedure including this new variability.

LMSA has a spectral library for the identification of SMs, built with the spectra acquired with the MPA equipment, located in the sampling room of the facilities of the logistics operator, since 2015. This is periodically updated to include the set of factors listed in the previous paragraph, obtaining a robust library. Currently, it was intended to have a backup device for carrying out identifications, if necessary, at the facilities of LMSA, being the MPA II device, acquired in 2019, installed in the sampling room at LMSA. However, this had to be demonstrated experimentally and provide evidence that this was the case since differences in the precision of the wavelength, linearity, or bandwidth in the entire spectral range could make the instruments optically different [20]. The instrument specifications are shown in Table 1. 3.

Table 1. 3 Specification for the parameters of the MPA and MPA II instruments.

Instrument Parameters		**Specifications**	
Model		FT-NIR MPA Bruker	FT-NIR MPA II Bruker
Software version		OPUS version 7.2.	OPUS version 8.2
Light scattering principle		FT	FT
Resolution		2 cm^{-1}	Max. 2 cm^{-1}
Wavenumber Accuracy		>0.1 cm^{-1}	0.1 cm^{-1}
Wavenumber repeatability (RMS)		>0.04 cm^{-1}	0.006 cm^{-1}
Photometric Accuracy		0.1 %T	0.1 %T
Interferometer		High stability interferometer Rocksolid™ permanent alignment	High stability interferometer Rocksolid™ with gold-coated mirrors and permanent alignment
Scanner (i.e. moving interferometer mirror)		4 speeds: 5, 10, 20, and 40 kHz	4 speeds: 3.75, 7, 15, and 30 kHz
Sample Compartment and Fiber Optic Module (Solid Probe)	**Detector**	TE-InGaAs	TE-InGaAs
	Spectral range (cm^{-1})	12800-4000	11500-4000

This section aims to study the direct transferability of the SM identification library, the current version, between NIR spectroscopic equipment: from the MPA to MPA II. Risk assessment based

on the FMEA approach was used to select the critical SMs involved in the study, including the analysis of the spectra acquired in parallel, based on the PCA analysis, for the products of interest selected with each of the NIR spectroscopic equipment.

1.2.1.3.1. Risk Assessment Development

The risk analysis of the 237 SMs was assessed using the FMEA methodology, considering the description included in the ICH Q9 guideline[21]. The fundamentals included were: the evaluation of each critical variable from the identification library (CQAs), including the factor weight (from now on abbreviated as W) for each CQA, from the severity (from now abbreviated as S) of the error (referred to as the effect of the method variable on the CQAs of the method), being the error to obtain a no satisfactory identification with the library (method variable effect from method CQAs), the probability of the error occurrence (from now on abbreviated as P), and the detectability of the error (from now on abbreviated as D), being the ability to be detected during the experimental procedure.

Selected CQAs were considered to achieve a robust identification library. The selection was based on the recommendations described for the identification library construction in the provider software manual (OPUS version 8.2. from Bruker Optics) and on the EMA guideline related to using NIR spectroscopy in the pharmaceutical industry in the validation of identification method section[10] (see Table 1. 4).

Table 1. 4 CQAs for obtaining a robust identification library using NIRS.

CQA	Definition
Classification	The number of libraries/sub-library assigned in the effective version library. Reflects the difficulty to differentiate SMs due to the spectra similarity.
Selectivity	See chapter 1.2.1.1. For those SMs in which Sel.≤1 occurs, the probability (from now abbreviated as Prob.*) represented in the set of SMs was calculated. Reflects the spectral variability within and between the groups in the identification library.
Physical state	The physical state of the SM, that is related to the difficulty in spectral acquisition.
Variability	The number of batches included in the effective identification library. Reflects the spectral variability between included batches for each SM in the library.

The risk priority is defined by the RPN as the combination of S, P, and D, or Critical Number (from now on CN), as the combination of S and P, where the D contribution is the same for all the evaluated samples. The numerical values are expressed in Equation 1. 12.

$$\mathrm{RPN}= S * P * D;\ \mathrm{CN} = S * P$$

Equation 1. 12.

The different parameters of Equation 1. 12. are detailed in Table 1. 5 to Table 1. 7.

The total value for S was calculated as the sum of all S values for each CQA, multiplied by W assigned at each CQA, and divided by the sum of W (Equation 1. 13).

Table 1. 5 Critical criteria for S parameter.

S	Importance	Classification	Selectivity	Physical state	Variability
1	Low Impact	Group 1-2	External validation Sel.>1	Solid. Direct spectral acquisition to the material	Batch nº>20
3	Medium Impact	Group 3-4	External validation Sel.≤1 and %Prob.*≤1	Solid. Spectral acquisition through one plastic bag	20≥Batch nº≥5
9	High Impact	Group 5-6	External validation Sel.≤1 and %Prob.*>1	Semi-solid or liquid. Spectral acquisition with a certified glass vial	Batch nº<5

*%Prob.= nº of samples from which the SM of interest presents a Sel.≤1 divided by the total number of groups in the effective library, multiplied by 100.

$$S = \frac{\sum_{i=1}^{CQA} S_i * W_i}{\sum_{i=1}^{CQA} W_i}$$

Equation 1. 13.

Table 1. 6 Criteria for W associated with each CQAs.

W	Description
1	Low impact on the overall quality of the result to be obtained
3	Medium impact on the overall quality of the result to be obtained
5	High impact on the overall quality of the result to be obtained

Table 1. 7 Critical criteria for P and D parameters.

Punctuation	Frequency of occurrence (P)**	Probability of detection (D)
1	Highly improbable: theoretically possible but practically impossible. Not observed before	Very high: failure is determined each time immediately after or during the analytical process. Automatically deleted
2	Remote possibility: it rarely happens. They were observed in the past but there are good measures to avoid them	High: failure is determined sometime after or during the analytical process. Human action or decision is involved at this point
3	Occasional: It doesn't happen systematically. The measures are not entirely effective	Medium: the failure is not detected every time or only sometime after the complete completion of the analytical process
4	Probable: It happens at regular intervals. The measures to be taken are ineffective or non-existent	Low: the measurements are not appropriate to detect the failure
5	Frequent: No measures can be taken to prevent it	Very low: the detection is impossible
**This parameter was based mainly on the historical data for those SMs with was necessary to include new variability in the calibration set for the next library maintenance. For those without historical data, the probability was scored based on the criticality of S.		

The analysis presented as a potential error the possible spectral non-equivalence between the MPA and the MPA II equipment for certain SMs, giving the effect of obtaining false negatives and/or positives with the reference library created with the acquired spectra with the MPA equipment, using the spectra acquired with the MPA II equipment. It was assumed that the causes of the possible error could be: the transport of the sample from the reception in the logistics operator sampling room to the LMSA sampling room, as well as the variation in environmental conditions due to the different rooms where are installed the respective equipment, and/or lack of variability included in the current library. These were considered since the supplier of both pieces of equipment guarantees the equivalence between them.

The criticality criterion was based on the CN parameter, since D presented the same criticality value for the set of SMs evaluated, being 1. According to the risk matrix (see Table 1. 8), those that presented a value of CN≥10 were considered critical or pending to assess their criticality. These SMs were divided into two main groups. The first corresponds to the SM set that presented a value of P≥3, which reflects the lack of variability included in the previous version of the reference library, these being the ones that have historically given unsatisfactory identifications. For these, there was a corrective action already implemented, that implied the inclusion of the set of identified batches in an unsatisfactory way, due to a lack of variability, in the future maintenance of the identification library. The second corresponded to those SMs with a value of P=2, being those SMs for which there was no historical data of unsatisfactory identifications due to past maintenance. Moreover, they showed a critical value of S following the critical criteria of each CQA selected in Table 1. 9. Due to

the lack of corrective action, its criticality was considered and the transferability evaluation was developed.

Table 1. 8 Risk matrix for the FMEA analysis table.

Severity					
9	9	18	27	36	45
8	8	16	24	32	40
7	7	14	21	28	35
6	6	12	18	24	30
5	5	10	15	20	25
4	4	8	12	16	20
3	3	6	9	12	15
2	2	4	6	8	10
1	1	2	3	4	5
	1	**2**	**3**	**4**	**5**

Probability

Table 1. 9 Criticality situations for S parameter.

Criticality situations	Classification	Sel.	Physical state
A	1, 3 o 9	9	9
B	9 o 3	9	3

The FMEA table for the critical SMs is described in Table 1. 11.

Moreover, the set of critical SMs is described in Table 1. 10, including the date of receipt of the last batch and the NIR identification strategy. For hydroxypropyl cellulose and Opadry lacquer it was not possible to carry out the transferability study because, as can be seen in the table, it had not been obtained for years.

Table 1. 10 Critical SM listed pending to evaluate based on the FMEA analysis.

Definition	Date last batch	NIR identification
Hydroxypropyl cellulose L	2015	Liquid through a glass vial
Carbomer	2020	Solid through one plastic bag
Cellulose microcrystalline, type 102 and 90 µm	2019	Solid through one plastic bag
Dexketoprofen Trometamol	2020	Solid through one plastic bag
Gemfibrozil	2019	Solid through one plastic bag
Cetirizine dihydrochloride	2019	Solid through one plastic bag
Polysorbate 60	2018	Liquid through a glass vial
Lacquer Opadry 02F33175	2016	Solid through one plastic bag

Table 1. 11 FMEA analysis including the critical SMs.

CQAs from the identification method	**Library Classification**	**Selectivity**	**Physic state**	**Variability**	**S**	**P**	**D**	**RPN**
Factor Weight	**3**	**5**	**5**	**1**				
SM Definition	0: Null 9: High	0: Null 9: High	0: Null 9: High	0: Null 9: High		1: Low 5: High	1: Low 5: High	
Hydroxypropyl cellulose L	1	9	9	9	7.29	2	1	14.6
Cellulose microcrystalline, type 102	3	9	3	3	5.14	2	1	10.3
Carbomer	3	9	3	1	5.00	2	1	10.0
Cellulose microcrystalline, 90µm	3	9	3	3	5.14	2	1	10.3
Dexketoprofen Trometamol (1)	3	9	3	3	5.14	2	1	10.3
Dexketoprofen Trometamol (2)	3	9	3	1	5.00	2	1	10.0
Dexketoprofen Trometamol (3)	3	9	3	1	5.00	2	1	10.0
Gemfibrozil	3	9	3	1	5.00	2	1	10.0
Cetirizine dihydrochloride	3	9	3	3	5.14	2	1	10.3
Cetirizine dihydrochloride (Phardiag)	3	9	3	1	5.00	2	1	10.0
Polysorbate 60	1	9	9	9	7.29	2	1	14.6
Lacquer Opadry 02F33175	3	9	3	9	5.57	2	1	11.1
(1)(2)(3)-Different internal LMSA API codes including different suppliers.								

1.2.1.3.2. Results and Discussion

The evaluation of possible spectral differences between equipment was proposed with the implementation of exploratory analysis based on PCA, for the spectra acquired in parallel with both NIRS equipment. Before the exploratory analysis, the spectra acquired with both equipment (MPA and MPA II) were identified with the currently implemented identification library. If the identification was satisfactory and the formation of groups by NIRS equipment did not occur in the exploratory analysis, it was concluded that there was direct transferability. Furthermore, if both identifications were unsatisfactory, direct transferability was also concluded, evaluating the root cause of the failure. However, if the identification was satisfactory and the formation of groups was observed, the exploratory analysis of the spectra of the studied samples was carried out together with the set of spectra included in the calibration of the implemented library. If no spectral differences were observed, a direct transferability was concluded; otherwise, a calibration transfer would be performed using the transfer method offered by the PDS-based instrument software. If all this didn't work, a new spectral library should be made, including the spectral variability of the spectra acquired with the new MPA II.

Another scenario would be to obtain an unsatisfactory identification for the spectra obtained with the MPA II equipment; this would lead to carrying out a calibration transfer by PDS and the consequent spectral evaluation by PCA, as well as a new identification with the implemented library. If this were successful, the need for PDS to perform spectral transfer would be concluded; otherwise, a single spectral library should be populated. The decision tree detailed is shown in Figure 1. 12.

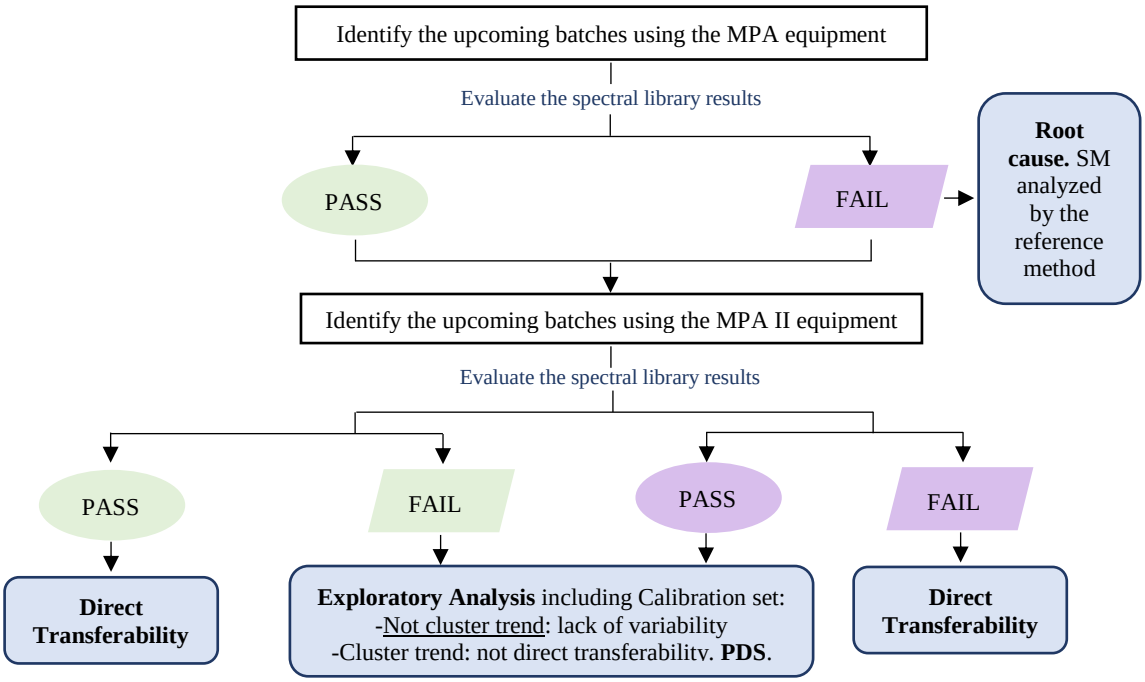


Figure 1. 12 Decision tree for preventive actions during the SM identification.

The transferability study for gemfibrozil material was presented as an example. The same spectral profile was visualized for the overall spectra acquired with the MPA and MPA II equipment. In addition, a cluster trend within the equipment, after applying the specific spectral pre-treatment used during the spectral library construction, was observed at the PCA analysis (see Figure 1. 13, and Figure 1. 14). No significant differences were observed with the calibration data set (see Figure 1. 15). Moreover, the identification of the effective spectral library version with the spectra acquired with both equipment (MPA and MPA II) was satisfactory.

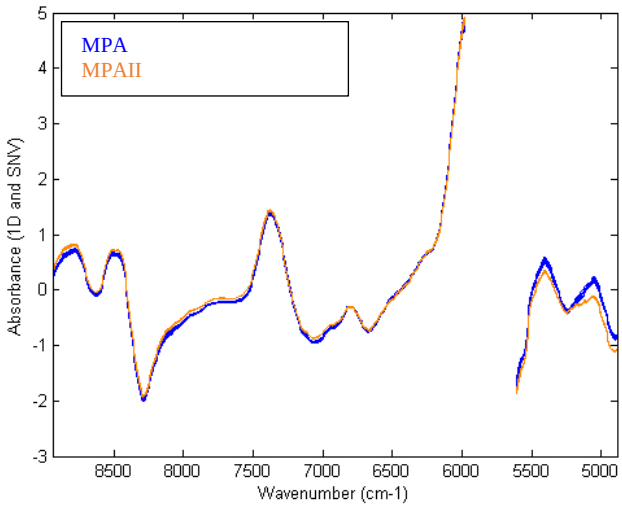


Figure 1. 13 Spectral data for the gemfibrozil acquired in diffuse reflectance mode with the MPA and the MPA II equipment. 1st derivative S-G (with 25 points of window width), followed by vector normalization was applied as data pre-treatment. Spectral range from 8941 to 5978 and 5608 to 4883 cm-1.

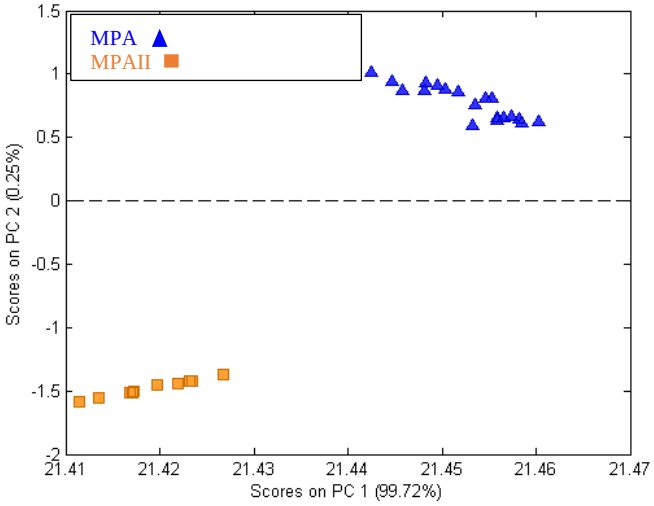


Figure 1. 14 Scores plot for PCA including 2 factors, PC1, and PC2) for the spectral data acquired in diffuse reflectance mode for the gemfibrozil material. 1st derivative S-G (with 25 points of window width) followed by vector normalization as data pre-treatment. Spectral range from 8941 to 5978 and 5608 to 4883 cm-1. 99.97% is included for all variability was included (99.72 and 0.25%, respectively).

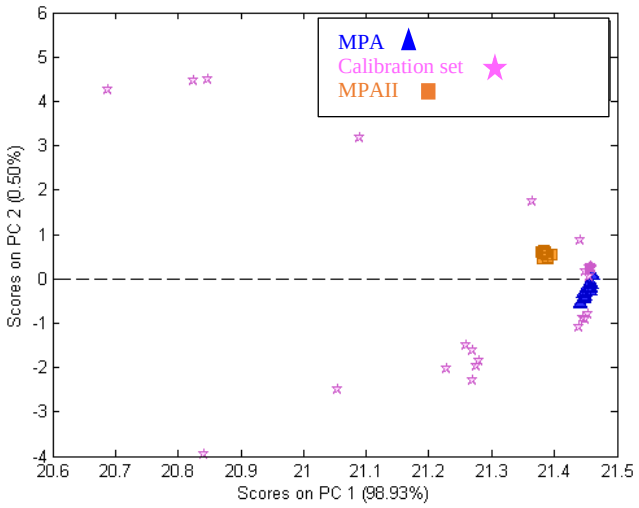


Figure 1. 15 Scores plot for PCA including 5 factors, from PC1 to PC5 for the spectral data acquired with diffuse reflectance for gemfibrozil material. 1st derivative S-G (with 25 points of window width) followed by vector normalization as data pre-treatment. Spectral range from 8941 to 5978 and 5608 to 4883 cm^{-1}. 99.99% for all variability was included (98.93, 0.5, 0.34, 0.14, and 0.08 %, respectively)

The graphical comparison including the spectra acquired with the MPA, MPA II, and the calibration data set included in the effective spectral library current version was presented for the rest of the critical SMs included in Table 1. 10 (see Figure 1. 16, Figure 1. 17, Figure 1. 18, Figure 1. 19, and Figure 1. 20).

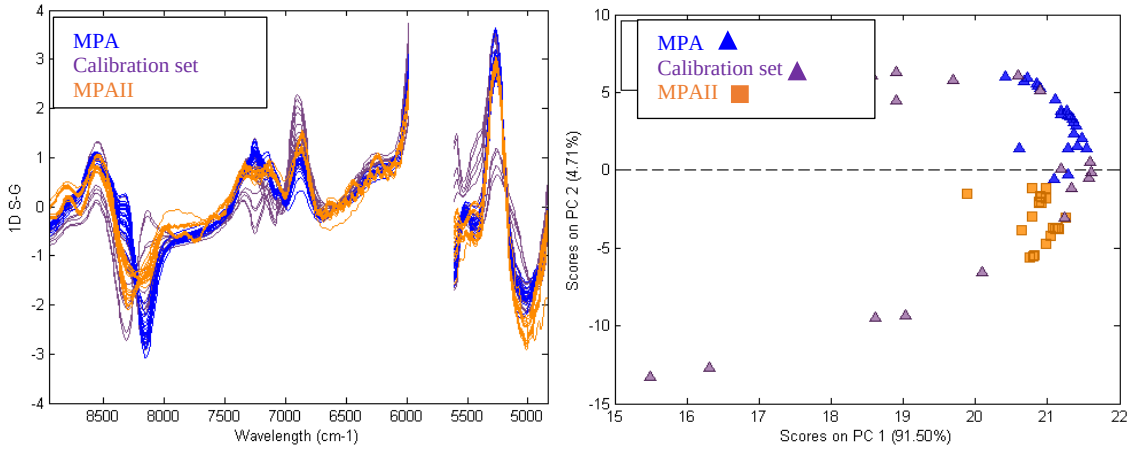


Figure 1. 16 Left) spectral data for the carbomer upcoming batch acquired in diffuse reflectance mode with the MPA and the MPA II. The spectra for the calibration set were included. Right) Scores plot for PCA including 2 factors PC1 and PC2, including the 96.21% for all variability was included (91.5 and 4.71, respectively). 1st derivative S-G (with 25 points of window width) and vector normalization as data pre-treatment, including the spectral range from 8941 to 5978 and from 5608 to 4883 cm^{-1}

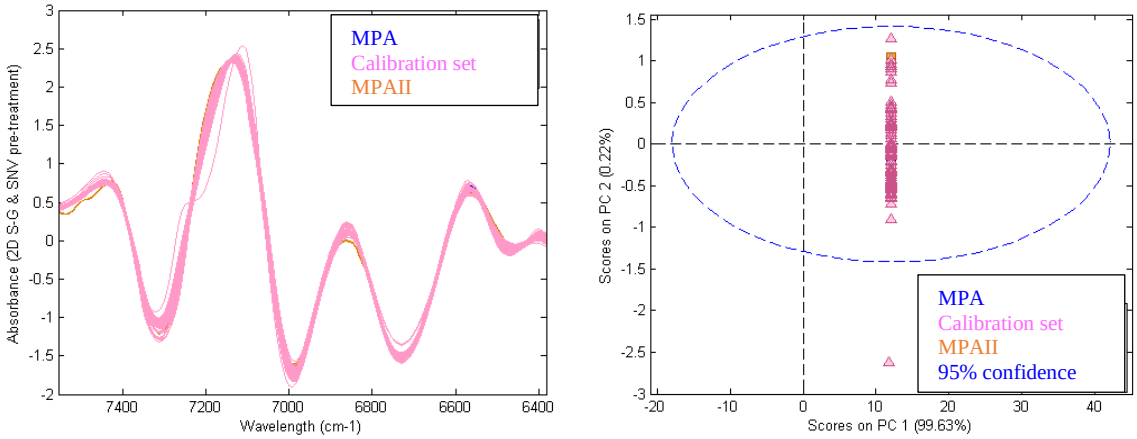


Figure 1. 17 Left) spectral data for the cellulose microcrystalline upcoming batch acquired in diffuse reflectance mode with the MPA and the MPA II. The spectra for the calibration set were included. Right) Scores plot for PCA including 2 factors PC1 and PC2 the 99.85% for all variability was included (99.63, and 0.22, respectively). 2nd derivative S-G (with 25 points of window width) and vector normalization as data pre-treatment, including the spectral range from 7552 to 6379 cm-1.

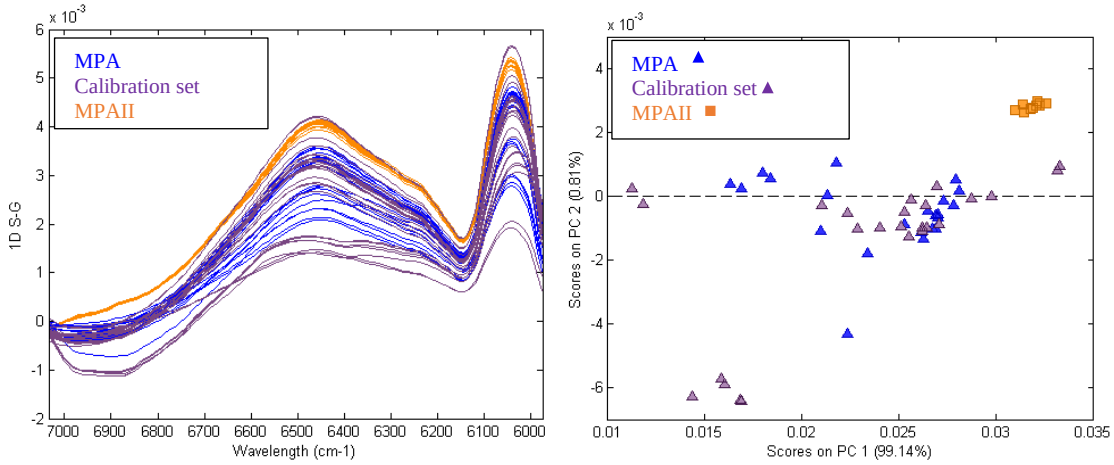


Figure 1. 18 Left) spectral data for the DKP-T upcoming batch acquired in diffuse reflectance mode with the MPA and the MPA II. The spectra for the calibration set were included. Right) Scores plot for PCA including 2 factors PC1 and PC2 the 99.95% for all variability was included (99.14, and 0.81, respectively). 1st derivative S-G (with 25 points of window width) and vector normalization as data pre-treatment, including the spectral range from 7035 to 5978 cm-1.

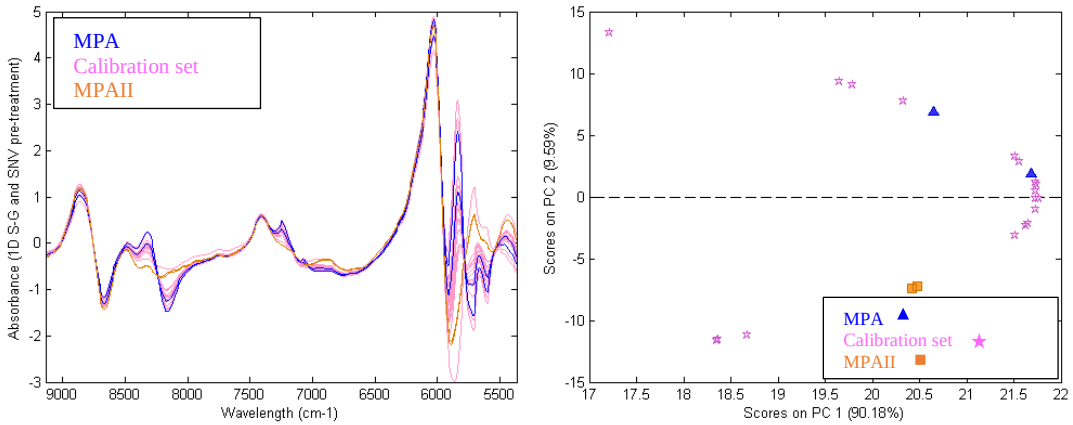


Figure 1. 19 Left) spectral data for the cetirizine upcoming batch acquired in diffuse reflectance mode with the MPA and the MPA II. The spectra for the calibration set were included. Right) Scores plot for PCA including 2 factors PC1 and PC2 the 99.77% for all variability was included (90.18 and 9.59, respectively). 1st derivative S-G (with 21 points of window width) and vector normalization as data pre-treatment, including the spectral range from 9141 to 5361 cm^{-1}.

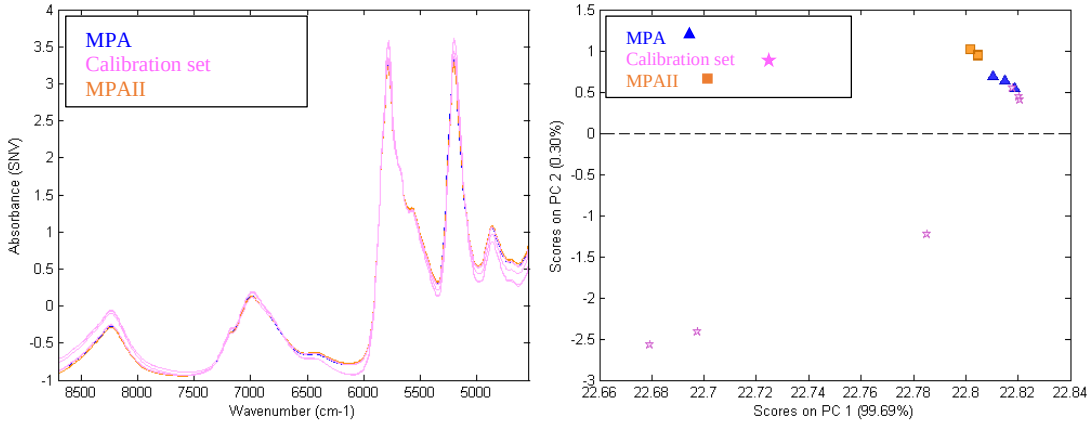


Figure 1. 20 Left) spectral data for the polysorbate upcoming batch acquired in diffuse reflectance mode with the MPA and the MPA II. The spectra for the calibration set were included. Right) Scores plot for PCA including 2 factors PC1 and PC2 the 99.99% for all variability was included (99.69 and 0.30, respectively). Vector normalization as data pre-treatment, including the spectral range from 8709 to 4528 cm^{-1}.

For all the defined critical SMs a direct transferability was concluded.

1.2.1.3.3. Verification of the Transferability Study

This study aims to verify the transferability between NIRS equipment (MPA and MPA II) of the identification library developed with the MPA equipment, once new library maintenance has been carried out and implemented. The frequency of the development of the present study was established after each maintenance of the library.

A total of 15 SMs were identified, including 26 different batches, using the latest library version, involving both NIRS equipment. Equivalent results were obtained in all studied cases (see Table 1. 12). The results presented were representative of 7% of the total SMs codes included in the latest version of the library detailed in this chapter. The variability included depends on the frequency of

arrival at the identification store between November (the month in which the new version of the library was implemented), and December 2022.

Table 1. 12 Set of batches identified with the latest version of the identification library involving the MPA and MPA II equipment.

Description	Num. Batches	MPA Identification	MPA II Identification
Cetirizine dihydrochloride	3	PASS	PASS
Citric acid, anhydrous	1	PASS	PASS
Dexketoprofen trometamol	7	PASS	PASS
Lactose, atomized	1	NOT PASS	NOT PASS
Lemon scent, ref. 03043	1	PASS	PASS
Macrogol cetostearyl ether 22	1	PASS	PASS
Neohesperidin dihydrochalcone	1	PASS	PASS
Nimesulide	1	PASS	PASS
Otolinium bromide	1	PASS	PASS
Sodium carmellose	1	PASS	PASS
Sodium starch glycolate type A	1	PASS	PASS
Starch, maize	1	PASS	PASS
Starch, pregelatinized	1	PASS	PASS
Sucrose, crystal	1	PASS	PASS
Sucrose, powder	4	PASS	PASS

The results of the identification, of the SMs identified with the latest version of the library, involving both NIRS equipment, were satisfactory. A non-conforming result was obtained with both NIR equipment for the lactose atomized material due to the lack of variability included in the identification library.

Transferability between NIR spectroscopic equipment was verified for the latest version of the identification library.

1.3. Conclusions

This section demonstrates the complete accomplishment of the correct maintenance for the Bruker NIRS identification library. All the evaluated spectra in the external validation for the library validation were correctly identified as belonging to the group of SM to which belongs. Furthermore, neither failure identification nor false positive results were obtained. Non-compliant identifications were given during the period of application of the presented maintenance library for two different types of SMs:

- SMs that had not been necessary to include new variability in the library maintenance presented; belonging to classes or including little variability in the calibration set.
- SMs that, even including new batches, the set of these included in the calibration set of the current library was less than 5.

The overall failed identifications represented 2% of the total of SMs identified during the period of application of the new maintenance presented (covering the period between October 2022 and May 2023).

Moreover, the direct transferability between NIRS equipment was demonstrated. The possible spectral differences were evaluated involving PCA as an exploratory analysis. It was assumed that the causes of the possible error could be: the transport of the sample from the reception in the logistics sampling booth to the LMSA sampling booth, as well as the variation in environmental conditions due to the different rooms where are the respective equipment and/or lack of variability included in the current library. The scores for the PCA should be grouped since there should be no separation between the data groups related to the instrumental origin. The initial hypothesis for those differences and a satisfactory spectral identification obtained demonstrated that the spectral variability included in the calibration set was sufficient to cover the variability shown in the new spectra acquired with the evaluated instrument.

Finally, the transferability between NIRS equipment was verified for the latest version of the identification library. A total of 15 SMs codes were identified, including 26 different batches. using the latest library version, with both NIRS equipment obtaining equivalent results in all cases. The results presented in this report are representative of 7% of the total SMs codes included in the latest version of the library.

1.4. Bibliography

(1) European Medicines Agency. ICH Q7 Good Manufacturing Practice for Active Pharmaceutical Ingredients Step 5. *CPMP/ICH/4106/00*. London, UK, November 2000.

(2) European Medicines Agency. ICH Guideline Q2(R2) on Validation of Analytical Procedures. *EMA/CHMP/ICH/82072/2006*. London, UK, March 31, 2022.

(3) Agency of Medicines and Medical Devices. Production. In *Guide to Good Manufacturing Practices Medicines for Human and Veterinary Use*; European Commission, 2016; pp 51–56.

(4) Agency of Medicines and Medical Devices. Sampling of Starting and Packaging Materials. In *Guide to Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*; European Commission, 2016; p 243.

(5) European Medicines Agency. Chapter 5: Production. *Good Manufacturing Practice Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, August 13, 2014, pp 1–12.

(6) European Medicines Agency. Annex 8: Sampling of Starting and Packaging Materials. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, October 20, 2010, pp 1–2.

(7) European Medicines Agency. Chapter 1: Pharmaceutical Quality System. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. European Commission Health and Consumers Directorate: Brussels, Belgium, December 14, 2012, pp 1–7.

(8) European Medicines Agency. Chapter 6: Quality Control. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 28, 2014, pp 1–8.

(9) Blanco, M.; Romero, M. A. Near-Infrared Libraries in the Pharmaceutical Industry: A Solution for Identity Confirmation. *Analyst* **2001**, *126* (12), 2212–2217. https://doi.org/10.1039/b105012p.

(10) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. London, UK, January 20, 2012.

(11) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 30, 2015, pp 1–16.

(12) Blanco, M.; Villarroya, I. *NIR Spectroscopy: A Rapid-Response Analytical Tool*; 2002; Vol. 21, pp 240–250.

(13) Miró Vera, A. Y.; Alcalà Bernàrdez, M. Near-Infrared Spectroscopy in Identification of Pharmaceutical Raw Materials. In *Encyclopedia of Analytical Chemistry*; John Wiley & Sons, Ltd, 2017; pp 1–19. https://doi.org/10.1002/9780470027318.a9619.

(14) Opus Spectroscopic Software. 7.6 Setup Identity Test Method: Validate. In *IDENT User Manual*; Bruker Optics, 2011; pp 104–115.

(15) National Center for Biotechnology Information. *Hydroxypropylmethylcellulose*. PubChem. https://pubchem.ncbi.nlm.nih.gov/compound/Hydroxypropylmethylcellulose (accessed 2023-01-22).

(16) Feudale, R. N.; Woody, N. A.; Tan, H.; Myles, A. J.; Brown, S. D.; Ferré, J. Transfer of Multivariate Calibration Models: A Review. *Chemometrics and Intelligent Laboratory Systems* **2002**, *64* (2), 181–192. https://doi.org/10.1016/S0169-7439(02)00085-0.

(17) Milanez, K. D. T. M.; Silva, A. C.; Paz, J. E. M.; Medeiros, E. P.; Pontes, M. J. C. Standardization of NIR Data to Identify Adulteration in Ethanol Fuel. *Microchemical Journal* **2016**, *124*, 121–126. https://doi.org/10.1016/j.microc.2015.08.013.

(18) Sun, L.; Hsiung, C.; Smith, V. Investigation of Direct Model Transferability Using Miniature Near-Infrared Spectrometers. *Molecules* **2019**, *24* (10). https://doi.org/10.3390/molecules24101997.

(19) Folch-Fortuny, A.; Vitale, R.; de Noord, O. E.; Ferrer, A. Calibration Transfer Between NIR Spectrometers: New Proposals and a Comparative Study. *J Chemom* **2017**, *31* (3). https://doi.org/10.1002/cem.2874.

(20) Ulmschneider, M.; Pénigault, E. Assessing the Transfer of Quantitative NIR Calibrations from a Spectrometer to Another One. *Analusis* **2000**, *28* (1), 83–87. https://doi.org/10.1051/analusis:2000101.

(21) European Medicines Agency. ICH Guideline Q9 on Quality Risk Management. *EMA/CHMP/ICH/24235/2006*. London, UK, September 2015.

Experimental Chapter 2. NIR and Raman Spectroscopy as Quality Control Tools

2. CQAs DETERMINATION BY SPECTROSCOPIC TECHNIQUES AS A QUALITY CONTROL TOOL

A pharmaceutical product has to meet the specifications before release. Proper control of pharmaceutical production involves determining CQAs that guarantee product quality through more efficient pharmaceutical manufacturing process[1,2].

This new approach has been recently achieved by effectively incorporating non-destructive and non-invasive, fast, and environmentally friendly techniques, such as NIR and Raman, as an alternative to mainly HPLC or UV-visible techniques at the QC laboratory[3,4]. Furthermore, the possibility of direct measurements that provide an accurate analysis without sample preparation or manipulation has positively impacted the expansion of spectroscopic techniques.

The active substance determination at different production process steps, for pharmaceuticals presented in different physical forms (including granules, powders (of variable granularity), core tablets, or coated tablets), provides the status process understanding confirming that a specific batch presents the required quality for release, avoiding the OOS formulations as soon as possible. As a result, the uniformity of dosage unit specification could be demonstrated if necessary[5].

This chapter aims to demonstrate the spectroscopic ability for quantitative analysis of the active substance at different production steps of two different pharmaceutical solid formulations. The content was evaluated after tablet coating (for Dexketoprofen Trometamol coated tablet (from now on DKP-T)) and after direct mixture for a powder formulation (for Sucramal unit-dose sachet product).

2.1. Instrumentation, Materials, and Methods

This chapter describes the instruments, equipment, and software involved during the development of the next projects.

2.1.1. V Blender

A V blender for the pilot plant scale from LMSA, V Mix-Lleal, was used to provide a fluid mixing process of solid/solid during laboratory sample preparation being as representative as possible of the manufacturing process.

The working volume was established at 50% of the total blender volume, being 16L or 8L. The dimensions were 195x65mm, and 152x87mm, respectively, for the loaded and unloaded diameters, of with a mixture velocity between 36-38 r.p.m. Figure 2. 12 shows a representative image of the V blender.

The CPP controlled was the blending time, established following the internal LMSA manufacturing guide for each formulation evaluated, between 10 to 15 min. At the end of the established time, a level sampling was performed to check the homogeneity of the prepared mix.

Figure 2. 12 V-blender model 16L (Lleal S.A.). Note: adapted from https://www.lleal.com/en/producto/mezclador-en-v/.

2.1.2. Tablet Compression Machine

At the pilot plant scale from LMSA, the rotary tablet press RONCHI AM 13/8 was used for DKP-T cores laboratory production. The equipment was supplied in the 8 punches version.

Figure 2. 13 Compression machine RONCHI AM 13/8. Note: adapted from https://www.equipnet.com/es/tableteadora-ronchi-am-13-8-listid-843636/.

The compression process was divided into three main steps (see Figure 2. 14):

-First step: the lower punch descends into the die, giving rise to a cavity in which the granules will flow by gravity from the hopper throughout the feeder system. The depth of the lower punch in the die will determine the volume of the compression chamber and therefore the final tablet weight. The tablet's thickness and size are also determined.

-Second step: an application of force, either by the upper punch or by both punches, exerting the necessary pressure on the particles to give rise to a tablet. This step determines the hardness of the tablet.

-Third step: the ascent of the upper and lower punch so that the tablet reaches the matrix and is ejected.

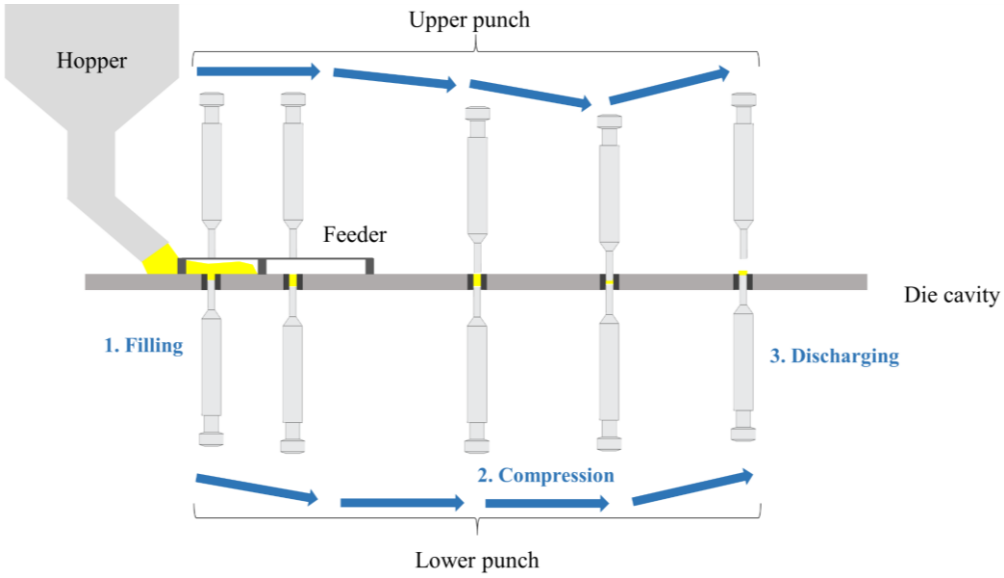


Figure 2. 14 Tablet compression flow. Note: adapted from https://www.saintytec.com/tablet-compression-machine-parts/.

The selected CPPs were manually adjusted to achieve the optimal physical core characteristics: volume of the compression chamber, compression force, and tablet velocity production which was established at the minimum level. The physical characteristics of the final DKP-T core evaluated are shown in Table 2. 1. The appearance of the core was also considered.

Table 2. 1. Specifications for the physical characteristics of the DKP-T cores.

Physical Characteristic	Specification
Weight (mg/core)	279-241
Thickness (mm)	3.0-3.4
Hardness (internal LMSA limit) (kp)	4.5-8.0

2.1.3. Coating Machine

The MyLab Automatic Coating System 6 L perforated pan was used for DKP-T tablet coating (see Figure 2. 15) at the A. Menarini Manufacturer Logistics and Services at Florence (from now on AMMLS FI), aimed at replicating the industrial scale manufacturing process as much as possible.

Figure 2. 15 MyLab automatic coating pan system. Note: adapted from https://ima.it/pharma/machine/mylab/.

The working volume was established at 50% of the drum capacity. The manufacturing method involved was adapted to the AMMLS galenic laboratory equipment and verified with two previous feasibility trials. The selected CPPs are reported in Table 2. 2.

Table 2. 2 CPP for coating process monitoring for DKP-T tablets.

CPP	Optimal Value
Coating suspension spraying time (min)	30
Drying time (min)	5
Air IN temp range (ºC)	45-55
Flow air IN (m^3/h)	200
Pan pressure (Pa)	-30
Pan speed (r.p.m)	3 (cores heating) 10-12 (spraying phase) 2 (cooling)
Atomization pressure (bar)	0.7
Spraying angle pressure (bar)	0.8
Cores temperature range (ºC)	38-40 20-25 (during cooling)
Pump flow range (mL/min)	5-15

The strategy followed for the spraying and coating process was established as follows:

- Coating suspension preparation following the internal LMSA methodology.
- Weight 200 cores before the coating process starts.
- Spray the proper coating suspension amount on the cores inside the pan, evaluating the CPPs.

- Monitor the coating process by weighing the coated tablets after 10 and 20 minutes of spraying.
- Weight 50 tablets at the end of the coating process, determining the final weight gain to evaluate some of the physical specifications of the final tablet (weight, thickness, and hardness). See Table 2. 3.

Table 2. 3. Specifications for the physical characteristics of the DKP-T tablets.

Physical Characteristic	**Specification**
Weight (mg/tablet)	260-270
Thickness (mm)	3.0-3.6
Hardness (internal LMSA limit) (kp)	4.0-7.5

2.1.4. Near-Infrared Spectroscopy

The FT-NIR MPA II benchtop spectrometer (Bruker Optics) located at the LMSA was used for data collection. The instrument is configured with an InGaAs photodiode detector. Spectra were acquired in reflectance mode. A wavelength range between 11550 and 3950 cm^{-1} was involved.

The instrument was daily qualified and maintained under a cGMP environment.

The spectra acquisition was followed using different strategies depending on the sample properties (for cores or tablets, and powder samples). See chapters 2.2.2.2.3.,2.2.3.3, and 2.3.2.3.

2.1.5. Raman Spectroscopy

The TRS100 benchtop (Agilent, Oxford) located at Agilent Technologies, Harwell, was used during a collaboration trip for cores and tablet data collection. The instrument was adapted with a specific tablet tray allowing the measurement of eighty samples per run (see Figure 2. 16). Transmission Raman technology was involved. The instrument presents a laser power of 650 nm and a laser wavelength of 830 nm specifically for pharmaceutical samples considering the balance between signal level, and competing fluorescence. The illumination spot size was established at 2 mm. The entire content was analyzed through ≥10 mm of sample thick. The interval range includes from 50 to 2500 cm^{-1}. See chapters 2.2.2.2.3., and 2.2.3.3.

The instrument was daily qualified and maintained under a cGMP environment using polystyrene, green glass, and a mercury argon light source as calibration standards.

Figure 2. 16 Left) TRS100 Raman instrument. Right) specific tablet accessory for DKP-T tablets.

2.1.6. Sample Preparation Strategy

A calibration and validation set were included for chemometric model construction[6]. Both have to include the variability sources that potentially affect the NIR and/or Raman spectra[7]. The controllable factors are the active substance content as well as the active substance manufacturer. Other uncontrollable factors can affect the spectra, such as water content, particle size, thickness, hardness, and coating performance. To take them into account, production samples were used and accurately selected to include these factors inside the expected variability and include as much variance as possible from the manufacturing process (including, as far as possible, different SM batches, expected physical variability, manufacturing process developed by different operators, and involved, if possible, different manufacturing machines).

The strategy for constructing the calibration and validation set consisted of manufacturing laboratory samples, obtained by doping the production samples by under and overdosing with a direct mixture of placebo and active substance, respectively[8]. The objective was to expand the active substance concentration from 70 to 130% relative to the nominal value including the acceptance specification intervals for the active substance content[9].

The mixture design strategy was used for placebo preparation. The objective was to expand the variability of the excipient around the acceptance interval for the specified concentration range. Different scenarios were obtained depending on the evaluated formulation. The prepared placebos were assigned randomly for laboratory sample preparation.

As far as possible, the SM involved in the preparation of the placebo included in the preparation of the calibration and validation laboratory samples were independent and were prepared on different days[6], depending on the frequency of arrival of the different batches of the included SM. In addition, the production batches included in each sample set were also independent.

The strategy followed for the preparation of the sample set of DKP-T tablets included, after the preparation of the laboratory mixture in the studied concentration range, the preparation of cores, and later of synthetic coated tablets with a pilot plant machine[8].

2.1.7. Software

The OPUS software version 8.2 was used for NIR instrument management, spectra acquisition, and chemometric modelling.

The Agilent ContentQC analysis was used as management software for the Raman instrument. Eigenvector Solo chemometrics engine was used as a development software for chemometric Raman modelling.

2.2. Develop a NIRS Procedure for Identifying and Qualifying a Coated Tablet Bulk Product and Quantifying the Assay for the Content Uniformity Test

The interest in the pharmaceutical industry in implementing rapid analytical techniques, to reach production in a PAT environment, is evident in the corporate project started at the end of 2017, which, given the knowledge acquired by LMSA, is the coordinator of the NIR implementation part. This project consists of three phases:

1. Implementation of NIR and Raman spectroscopies in the QC laboratories for the determination of the CQA active substance concentration for unit-dosage forms, for content uniformity analysis (considering tablets, capsules, or granules in sachets).
2. Implementation of NIR and Raman spectroscopies in production lines (PAT environment).
3. QbD and real-time release testing (from now on RTRT) implementation.

Initially, a technical and productive risk analysis was carried out, which served as the basis for selecting the three drugs with which the NIR and Raman methods will be developed for the routine determination of CQAs in the bulk product.

This study aims to summarize the development and validation of qualitative and quantitative analysis procedures for the selected pharmaceutical products involving NIR and Raman techniques.

2.2.1. Studies Previously Developed

Measurements were optimized and carried out for the candidate medicines, selected by the different manufacturing plants of the Menarini Group attached to the project (Badalona, Berlin, Dresden, Florence, and L'Aquila). These measurements were made by NIR in the analytical development facilities that Bruker has in Ettlingen (Germany), using a NIR MPA II equipment (Bruker Optics), led by the Italian QC department. As a result, three pharmaceutical products were selected as candidates: bilastine, atorvastatin tablets (both manufactured at the L'Aquila Menarini plant), and DKP-T tablets (manufactured at the LMSA plant).

Quantitative NIR and Raman procedures were properly developed for the Bilastine tablet bulk product[10]. Scientific advice was sent to regulatory authorities. They proposed to demonstrate scientific evidence of the robustness of the procedures presented.

2.2.2. Dexketoprofen Trometamol 25mg Tablet Product

This section contains the evaluation of the composition and the manufacturing process for the DKP-T 25mg coated tablet product. The effect of each evaluated factor on the NIR and Raman was determined using a risk analysis as a tool. Furthermore, a proper calibration and validation set are presented based on a DoE. Finally, the development, optimization, and validation of both NIR and Raman spectroscopic procedures are presented to properly identify the DKP-T 25mg coated tablet product and correctly quantify the CQA of interest (active substance content for content uniformity analysis).

2.2.2.1. Composition and Manufacturing Process

The DKP-T tablet 25 mg is a film-coated tablet with an active substance concentration of about 9.4% (w/w). It is a round biconvex tablet with equally scored faces. The composition of the cores and the coated tablets for DKP-T are shown in Table 2. 4. Two different suppliers are nowadays authorized for the active substance (from now abbreviated as A1 and A2) and only one authorized supplier for each included excipient.

Table 2. 4 DKP-T composition for cores and coated tablets.

	Description	**Content per tablet (mg)**	**Composition of manufacturing formula (mg)**
Core	DKP-T	36.9	36.9
	Glycerol distearate T-I	5.20	5.20
	Sodium starch glycolate type A	27.1	27.1
	Corn starch	49.6	54.2**
	Microcrystalline cellulose, 50µm	141	143**
	Purified water	-	200
	Core weight	260	260
Coated tablet	Propylene glycol	0.420	0.440***
	Dry lacquer	2.30	2.42***
	Purified water	-	17.8
	Coated weight	262.7	262.7

**Overages due to the loss of water. Purified water is a component that disappears during the manufacturing process.

***Overages due to the loss of process.

DKP-T coated tablets 25 mg manufacturing process includes five-unit operations: granulation, dying, mixing, compression, and coating. The detailed flowchart is summarized in Figure 2. 17.

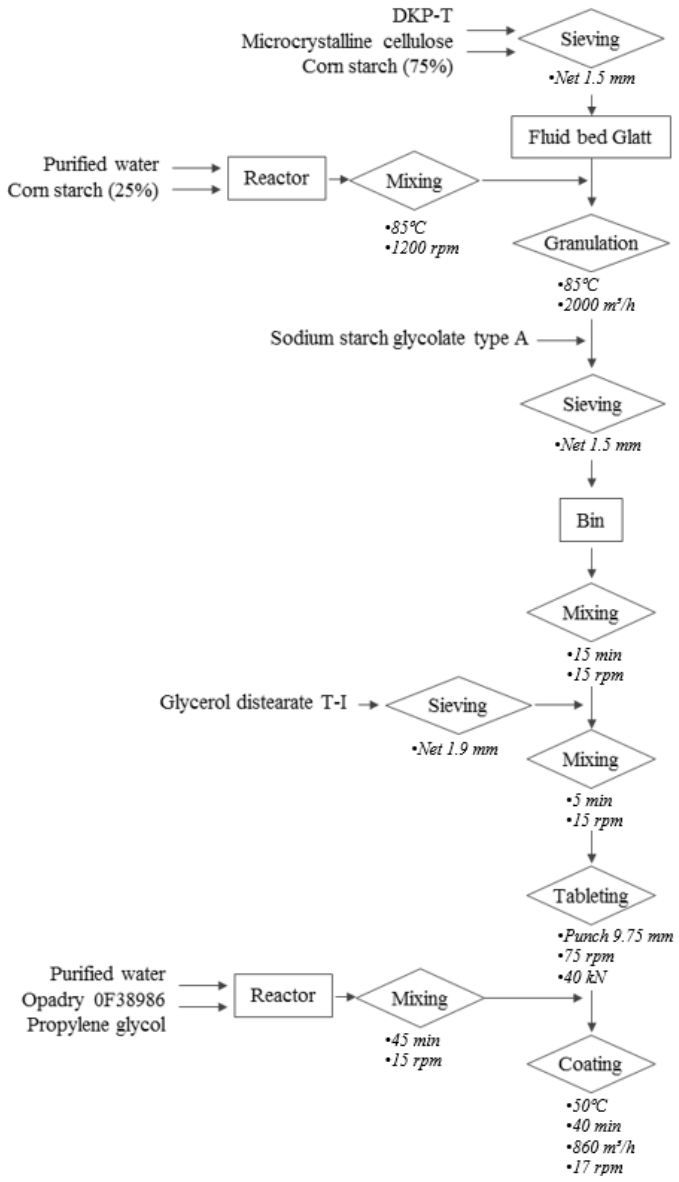


Figure 2. 17 Manufacturing LMSA process of DKP-T coated tablets 25mg.

2.2.2.2.Materials and Methods

This chapter describes the materials and methods involved during the project development.

2.2.2.2.1. Calibration and Validation Data Sets for Spectroscopic Qualitative Analysis

The samples involved in calibration and validation for the identification and qualification analysis for the DKP-T 25 mg tablet were manufacturing batches (from now on abbreviated as MB), including

cores and tablets, formulation excipients, powdered prepared placebos (see Table 2. 6), and a product that was formulated with the same active substance. The samples are detailed in Table 2. 5. For the Raman development procedure, only the excipients considered to be the majority were included, with a contribution to the formulation greater than 5%.

Table 2. 5 Calibration and validation sets for identification analysis.

Description	**Justification**
DKP-T 25 mg cores	Calibration and external validation set. Included for specificity evaluation due to the absence of a coating agent
DKP-T 25 mg tablets	Calibration and external validations set. Included for specificity and robustness evaluation. Target product
DKP-T	Calibration and external validation set. Included for specificity evaluation. Target product without the presence of the excipients
Corn starch	Calibration and external validation set. Included for specificity evaluation. Majority excipient
Glycerol distearate T-I	Calibration and external validation set. Included for specificity evaluation. Minority excipient
Microcrystalline cellulose, 50µm	Calibration and external validation set. Included for specificity evaluation. Majority excipient
Prepared placebo (powder)	Calibration and external validation set. Included for specificity evaluation. Target product without the presence of the active substance
Sodium starch glycolate type A	Calibration and external validation set. Included for specificity evaluation. Minority excipient

2.2.2.2.2. Calibration and Validation Data Sets for Spectroscopic Quantitative Analysis

Laboratory samples were prepared for the calibration and the validation data sets following the strategy of under and overdose production samples after the granulation and the final mixing step. Seven concentration levels were included between the expanded range: 70, 85, 100, 115, and 130, as well as 95 and 105%, assigned as the acceptance intervals for the active substance content. Using mixture design, five placebos were prepared to expand the excipient variability from the final DKP-T mixture to around 5%. The strategy followed was to modify around 5% of the composition for those considered majority (nominal value higher than 5%), from the nominal value of placebo composition; corn starch, microcrystalline cellulose, and a mixture from sodium starch glycolate type A, and glycerol distearate T-I reducing as much as possible correlations between formulation components (see Table 2. 6). A total amount of 3.5 kg for each placebo was prepared to be able to use the pilot plant compression and coating machines.

Table 2. 6 Content of each excipient for 3.5 kg of each placebo. Expanded excipient variability around 5% from nominal value from those considered majorities.

Corn starch (%, w/w)	Glycerol distearate T-I (%, w/w)	Microcrystalline cellulose, 50µm (%, w/w)	Sodium starch glycolate type A (%, w/w)	Placebo code
21.12	2.21	65.13	11.54	**1**
23.35	2.21	62.90	11.54	**2**
21.12	2.45	63.68	12.75	**3**
23.35	2.45	61.46	12.75	**4**
22.23	2.33	63.29	12.14	**5**

The calibration set included initially 12 prepared laboratory samples and two MB at the nominal value of the active substance. A total amount of 1.5 kg for each concentration level was prepared to be able to use the pilot plant compression and coating machines. At least four different batches were included, two batches for each homologated active substance manufacturer (assigned from B1 to B4). The prepared placebos were assigned randomly, expanding the variability from 95 to 107% for all concentration levels.

To capture as much variance as possible, batches from the manufacturing process were selected carefully (including, as far as possible, different SM batches, expected physical variability, granulated by different operators, and compressed with different tableting machines). The structure of the prepared samples is summarised in Table 2. 7.

Table 2. 7 Calibration set for the quantitative analysis of DKP-T tablets 25 mg.

Concentration level (%)	Number of tablets	Placebo code	Active substance supplier	Active substance batch number
70	10	4	-	-
85	10	3	-	-
95	10	2	-	-
100 MB	10	-	A1	B2
105	10	-	A1	B1
115	10	-	A1	B2
130	10	-	A1	B1
70	10	1	-	-
85	10	2	-	-
95	10	5	-	-
100 MB	10	-	A2	B3
105	10	-	A2	B4
115	10	-	A2	B3
130	10	-	A2	B4

Once the samples were mixed the final blend was compressed in the pilot plant and the core tablets were coated. 140 tablets were analyzed individually, 10 tablets per concentration level, considering each tablet's face.

The prepared placebos and the active substance batches involved in the validation set were completely independent of the calibration set. The placebos were prepared from different batches and on different days following[6] the composition shown in Table 2. 6. The validation set included initially 12 laboratory samples and 2 MB at the nominal value of the active substance, completely independent from the calibration set. At least four batches were included, ideally two different batches for each homologated active substance manufacturer (assigned from B5 to B8). This strategy expanded the placebo variability between 95 and 107% for all concentration levels. The structure of the prepared samples is summarized in Table 2. 8.

Table 2. 8 Validation set for the quantitative analysis for DKP-T 25 mg tablets.

Concentration level (%)	Number of tablets	Placebo code	Active substance supplier	Active substance batch number
70	10	2	-	-
85	10	1	-	-
95	30	3	-	-
100 MB	30	-	A1	B5
105	30	-	A1	B6
115	10	-	A1	B5
130	10	-	A1	B6
70	10	5	-	-
85	10	4	-	-
95	30	1	-	-
100 MB	30	-	A2	B8
105	30	-	A2	B7
115	10	-	A2	B8
130	10	-	A2	B7

Once the samples were mixed the final blend was compressed in the pilot plant and the core tablets were coated. 260 tablets were tested individually, 30 for the samples in the assigned acceptance range for the active substance content, and 10 for the rest, considering each tablet's face.

2.2.2.2.3. Instrumentation and Software

The FT-NIR MPA II (Bruker Optics, USA) instrument was coupled with an integration sphere adapted with a sample wheel measurement module of 30 positions for spectra acquisition. For laboratory-prepared samples, a specific spectral acquisition tablet holding device was used to measure a 2-mm diameter sample (see Figure 2. 18). Therefore, for MB was necessary to measure

the tablet spectra through a glass vial to reduce the travel distance of the light beam and increase the sample diameter from 2 mm to 12.5 mm. Each spectrum was recorded with 32 scans and 16 cm^{-1} of resolution.

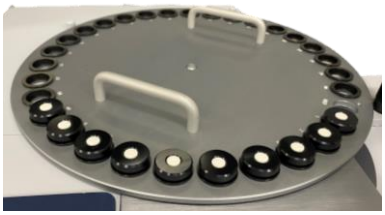


Figure 2. 18 Sample wheel accessory of 30 positions for laboratory-prepared tablets measurement using the FT-NIR MPA II instrument.

The TRS100 instrument was used for Raman data acquisition. The instrument and its related measure parameters are described in Table 2. 9.

Table 2. 9 RAMAN instrument configuration.

Parameter	Specification
Model	Analyzer TRS100 Raman
Software for data acquisition	Agilent ContentQC
Acquisition Method	External Transmission
Laser power (W)	0.65
Beam size (mm)	2
Accumulations	125 or 100
Exposure (secs)	0.4 or 0.5
Total scanning time (secs)	50
Software for data modelling	Solo Interface from Eigenvector Research

The UV-visible technique was used following the internal procedure as a reference method. The methodology and the related measures parameters were described in Table 2. 10.

Table 2. 10 UV-visible instrument configuration.

Parameter	Specification
Sampling accessory	1 cm cuvette
Solvent	1:3 H_2O/MeOH
Wavelength Interval	210-400 nm
Maximum Absorbance	258 nm

As a result, the active substance content for each analyzed sample was calculated as shown in Equation 2 19.

$$Active\ substance\ content\ \frac{mg}{tablet} = \frac{Abs\ x(100 - R)xF}{662}$$

Abs: absorbance of the problem solution at 258 nm.

$A^{1\%}_{1cm}$ (the absorbance of a 1% (w/v) solution in a 1 cm cell) from ketoprofen: 662.

F: quantification factor being 250.

R: percentage of the isomer R (-) ketoprofen.

Equation 2 19.

2.2.2.3.Risk Assessment

The Ishikawa diagram was used as a starting point for risk analysis to detect possible causes affecting the method development. See Figure 2. 19.

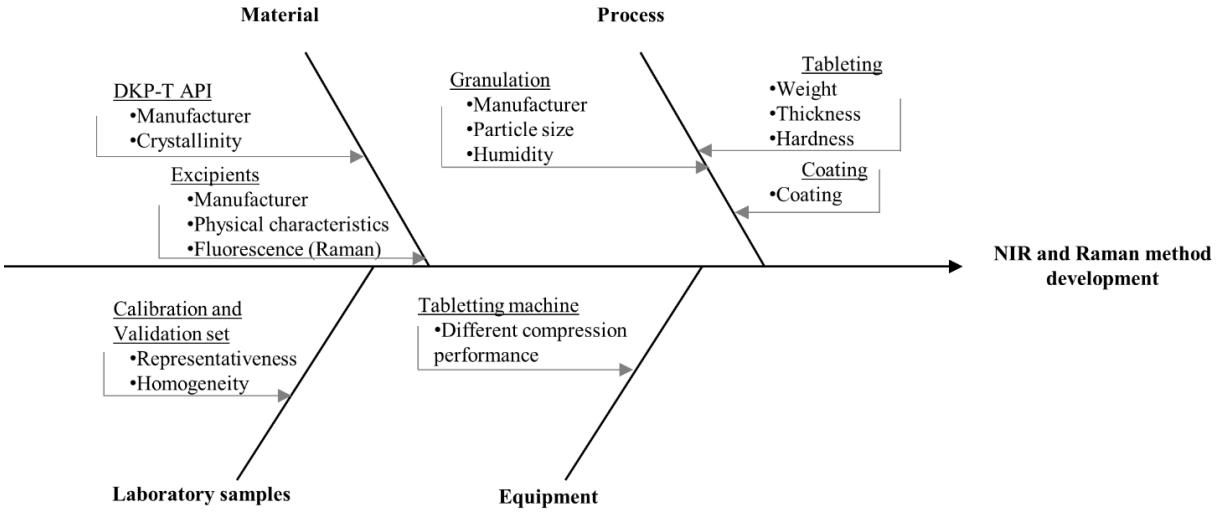


Figure 2. 19 Ishikawa diagram for DKP-T coated tablets 25mg.

Materials

DKP-T (active substance):

- Manufacturer: two manufacturers are homologated.
- Crystallinity: DKP becomes a mix of two crystalline forms during the granulation process.

Excipients:

- Manufacturers: for all excipients, except for sodium starch glycolate type A one manufactured is homologated. The two manufacturers of sodium starch glycolate type A are used indistinctly
- Physical characteristics: the effect of the initial physical characteristics such as particle size or water content is not important once the product is granulated.
- Fluorescence: microcrystalline cellulose shows a high fluorescence on the Raman technique.

Laboratory samples

Calibration and validation set development. Equivalent and homogeneous laboratory samples have to be achieved to guarantee process representativeness.

Equipment

Tableting machine: two tableting machines are suitable to compress DKP-T tablets. The compression performance between them should be equivalent. The performance directly impacts tablet weight, thickness, and hardness. If the compression between machines is not equivalent, commercial batches from each one should be included in the calibration and validation set.

Process

- Granulation: DKP-T is transformed into a mix of two crystalline forms. It is assumed that this variability will be included in the calibration and validation set with production samples.
- Tabletting: the pressure of compaction will affect tablet thickness and hardness. It is assumed that these differences will be included in the calibration and validation set with production samples.
- Coating: the coating will affect the NIR and Raman spectra; process differences will be assumed in the calibration and validation set with production samples.

As a result, the main areas of the DKP-T coated tablets manufacturing process as well as calibration and validation set development were investigated to find the probable risk factors that might cause the failure of the development of the active substance content quantification method based on NIR and Raman spectroscopy. The severity ranking used is shown in Table 2. 11.

Table 2. 11 Severity ranking.

High	6	Wrong model detected after regulatory application (or not detected at all)
	5	Wrong model detected before the regulatory application
Medium	4	Wrong model detected during the development phase
	3	Reworking of calibration/validation sets
Low	2	Issues on acquisition of spectra
	1	None effects

The cause and effect risk analysis for DKP-T 25 mg tablets is detailed in Table 2. 12.

Table 2. 12 Cause and effect risk analysis for the DKP-T 25 mg tablets manufacturing.

Cause	Rank	Effect	Mitigation
Material DKP-T active substance. Crystallinity	1	None expected. The active substance becomes a mix of polymorphic forms during the granulation process	For practical purposes, the equivalence between polymorphic forms is verified
Material DKP-T active substance. Manufacturer	3	Medium risk expected. Two manufacturers are homologated. Differences in NIR and Raman spectra	Two manufacturers were included in the calibration and validation set
Material Excipients. Manufacturer	1	None expected. For all excipients, except for sodium starch glycolate type A one manufactured is homologated	The two manufacturers of sodium starch glycolate type A were included with the MB. Both are used indistinctly
Material Excipients. Physical properties	1	None expected. The initial physical characteristics of the granular excipients don't affect the final particle size or water content	Does not apply
Material Excipients. Microcrystalline cellulose fluorescence	4	Medium risk expected for Raman. Interference in the spectrum	Select the appropriate spectral range during the procedure development as well as the correct pre-treatment.
	1	None is expected for NIR.	Does not apply.
Laboratory samples Calibration and validation set development	4	Heterogeneous laboratory samples or non-equivalent laboratory tablets from produced ones	Each laboratory step to obtain the galenic tablets was verified involving validated physicochemical assays
Equipment Tableting machine	2	Differences in NIR and Raman spectra due to differences in pressure performance of each piece of compression equipment	Expected variability introduced on the calibration and validation set with commercial batches manufactured with both machines
Process Wet granulation. Crystallinity	3	Medium risk expected. The active substance becomes a mix of crystalline forms. Differences in NIR and Raman spectra	Expected variability introduced on the calibration and validation set with production samples. A preliminary study was conducted to compare MB spectra
Process Wet granulation. Particle size	1	Low risk. Particle size could have a low effect on the Raman spectrum	Expected variability introduced with MB
	4	Medium risk. The particle size has a huge effect on the NIR spectrum (scattering effect).	Expected variability introduced on the calibration and validation set with MB
Process Wet granulation.	1	None expected. Water content doesn't affect the Raman spectrum	Does not apply

Table 2. 12 Cause and effect risk analysis for the DKP-T 25 mg tablets manufacturing.

Cause	Rank	Effect	Mitigation
Humidity	4	Medium risk. NIR is very sensitive to water content	Expected variability introduced on the calibration and validation set with MB
Process Tableting. Thickness and hardness	3	Medium risk expected. Differences in the NIR and Raman spectra due to differences in tablet thickness and hardness	The variability is introduced on the calibration and validation set with MB
Process Coating	3	Medium risk expected for Raman. Differences in the Raman spectrum due to differences in coating performance	Expected variability introduced on the calibration and validation set with MB. A preliminary study was conducted to compare spectra for cores and tablets
	2	Low risk expected for NIR. Differences in NIR spectrum due to differences in coating performance	Expected variability introduced on the calibration and validation set with MB. A preliminary study was conducted to compare spectra for cores and tablets

2.2.2.4.Investigation of the Coating Effect to the NIR and Raman Spectra

Physical and chemical factors must be controlled during the production process of a pharmaceutical drug due to its influence on its quality. The last step during the production of DKP-T 25 mg coated tablets involves coating with a thin lacquer containing polymers. The coating introduces two new chemical compounds in the formulation, increasing the weight of the 25 mg tablet by approximately 1%. The coating variable could have some effect on the NIR as well as the Raman spectrum for the sample of interest, therefore, it must be evaluated to be considered during the development of a multivariate calibration model to determine the active substance concentration (for content uniformity analysis).

The strategy followed to carry out this investigation is the next one: to capture as much variance as possible, three different batches from the manufacturing process were selected carefully. The sample set includes uncoated and coated tablets.

The spectra from sixty tablets were acquired individually with NIR and Raman instruments (see Table 2. 13. Then an exploratory analysis based on PCA was developed to extract the information from the obtained spectra.

Table 2. 13 Sample set for evaluation effect of the coating.

Batch sample	Active substance content (%)	Type of sample	Number of tablets
MB-A	100	Uncoated	10
		Coated	10
MB-B	100	Uncoated	10
		Coated	10
MB-C	100	Uncoated	10
		Coated	10

For NIR analysis a cluster trend was observed along the second PC in the scores plot after reducing the scattering effect by applying vector normalization as data pre-treatment. However, there was no specific absorbance band related to the presence coating agent (see Figure 2. 20).

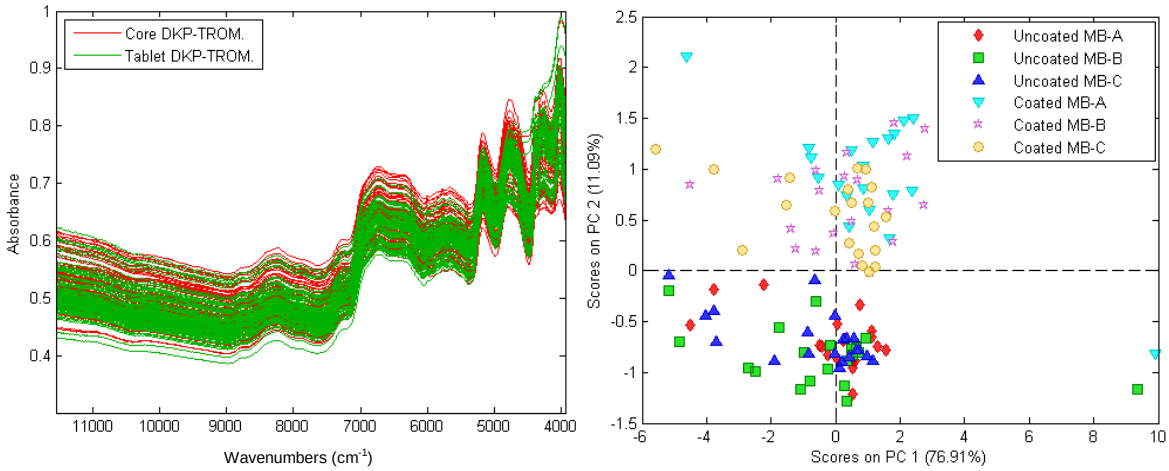


Figure 2. 20 Left) Spectral data in the reflection mode for core and tablets of DKP-T. Right) Scores plot for the PCA analysis for core and tablets of DKP-T.

The pre-treated spectra for the excipients included in the formulation for the coating show the same spectra profile as the second LV in the loading graph, confirming the cluster trend (see Figure 2. 21).

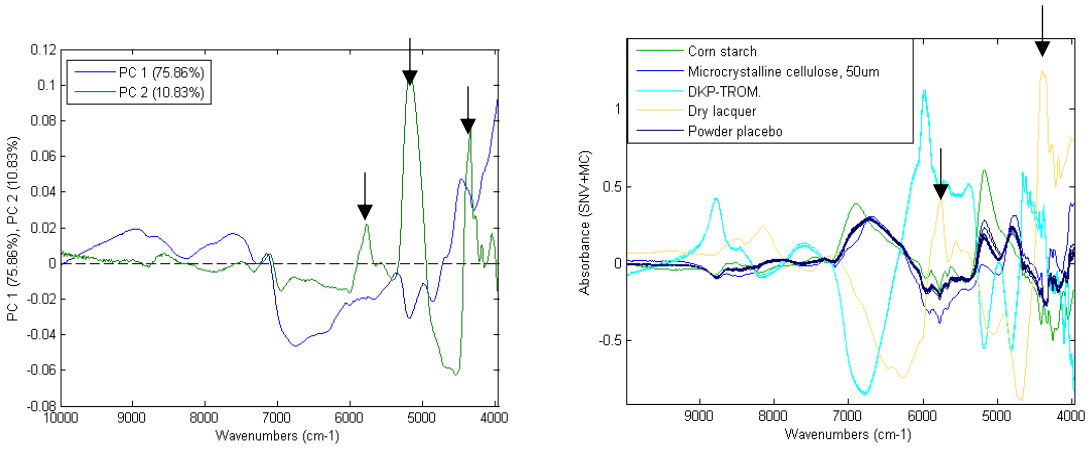


Figure 2. 21 Left) Loading graph for the significant PCs. Right) Pre-treated spectra for the majority formulation components of DKP-T tablets. Marked arrows for the coating spectra profile detection.

For Raman analysis blooming phenomena occurred due to the presence of Microcrystalline Cellulose as the main excipient in the pharmaceutical formulation (see Figure 2. 22).

The preliminary study shows a set apart within MB for cores and tablets depending on the coating below 200 cm^{-1}. A cluster trend through the PC3 was observed based on the presence of the TiO_2 included in the dry lacquer (see Figure 2. 23).

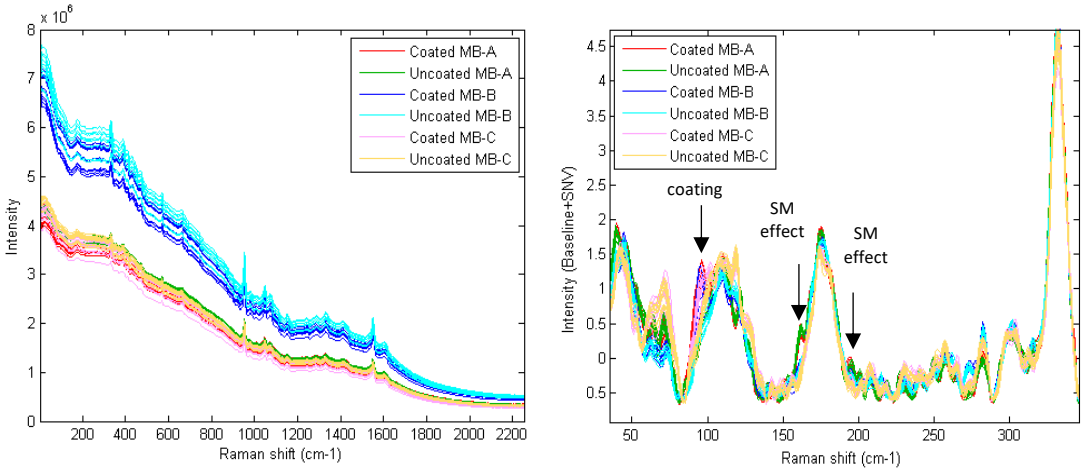


Figure 2. 22 Transmission Raman spectra for cores and tablets. Left) Raw data included. Right) Baseline correction and vector normalization were applied as spectra pre-treatment. Raman shift included from 45 to 350 cm^{-1}.

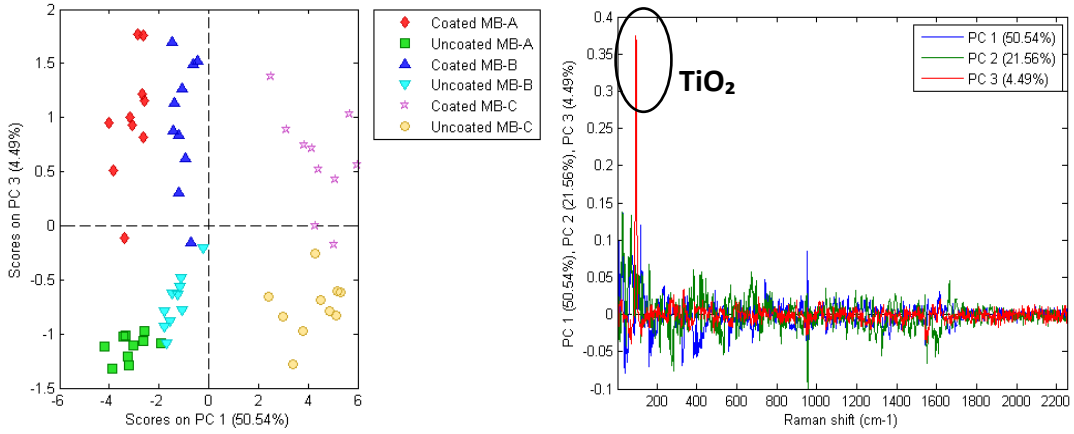


Figure 2. 23 Left) Scores plot for PCA analysis of cores and tablets. Right) Loading plot for the significant PCs for the exploratory analysis. Baseline correction and vector normalization were applied as spectra pre-treatment. Raman shift included from 45 to 350 cm^{-1}.

As a result, the calibration and validation data set were constructed with DKP-T laboratory and production tablets to include the variability related to the coating process.

2.2.2.5.Exploratory Analysis. Investigation of the Effect of the Active Substance Polymorphism on the Chemometric Raman Model

Raman spectroscopy is a useful technique for crystal structure characterization. Structural information for polymorphs can be achieved due to the different crystal packing forms and as a result,

different molecular vibrations. The crystal polymorphism contribution in Raman spectra was evaluated. A previous study demonstrates the presence of a mixture of polymorphic forms for DKP-T active substance after the manufacturing process (polymorph A and B)[11,12]. The effect of the polymorphic contribution in the Raman spectroscopy is evaluated in this section.

The main bands of the active substance in the transmission range appeared at 1000 cm^{-1} and between 1500 and 1600 cm^{-1} without interference with the placebo mostly. When the spectra were preprocessed with smoothing 21 windows width, baseline correction, and vector normalization, the spectra of the calibration set were arranged according to their concentration. Those bands were related to the active substance as can be checked in Figure 2. 24.

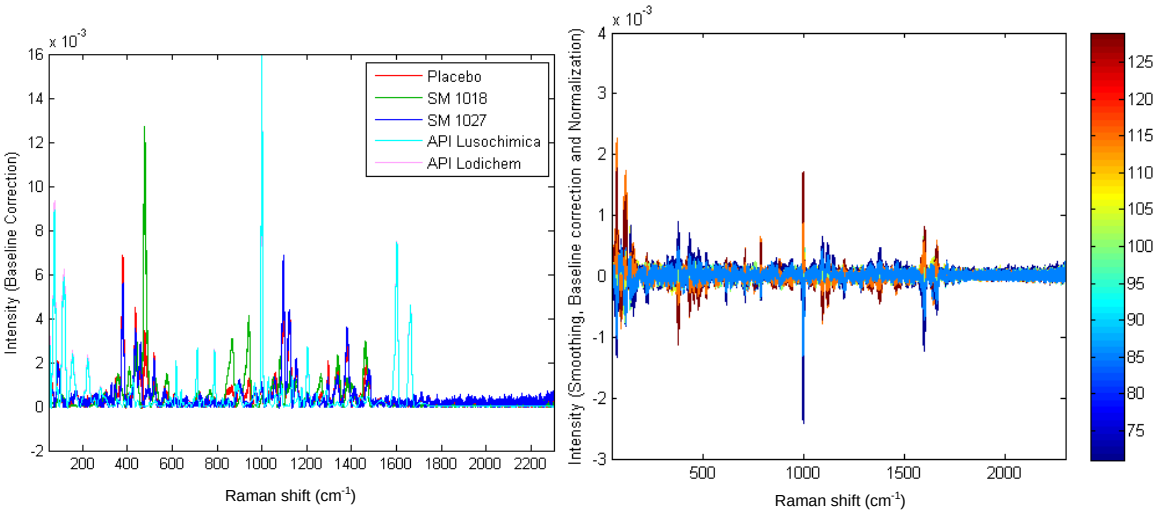


Figure 2. 24 Transmission Raman spectra. Left) Major excipients, placebo, and different active substance suppliers after baseline correction. Right) Laboratory sample set and MB included for calibration ordered by the active substance concentration after smoothing 21 windows width, baseline correction, and vector normalization.

The scores plot for the PCA of the calibration set, including wavenumber range from 979 to 1614 cm^{-1}, once it was applied smoothing with 15 points of windows width followed by baseline correction and normalization, showed throughout the PC1 the differences between samples according to the active substance concentration (see Figure 2. 25). PC2 explained the presence of different contributions between the polymorphic forms of the active substance at the included MB. This hypothesis was previously assumed during the MB analysis where the relative intensity between the characteristic active substance peak at 1000/1600 and 1000/1650 cm^{-1} was different (see Figure 2. 26). To be confirmed, a preliminary RX analysis for 4 different MBs randomly selected was evaluated[13]. Afterward, the hypothesis was confirmed.

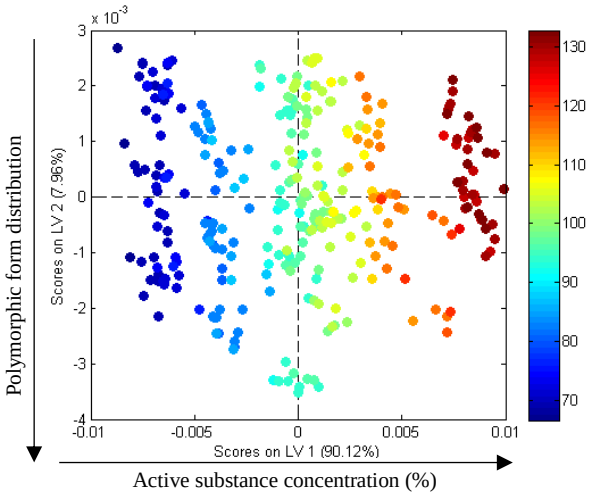


Figure 2. 25 Scores plot for the calibration set including PC1 and PC2. The 98.08 % of the variability is explained.

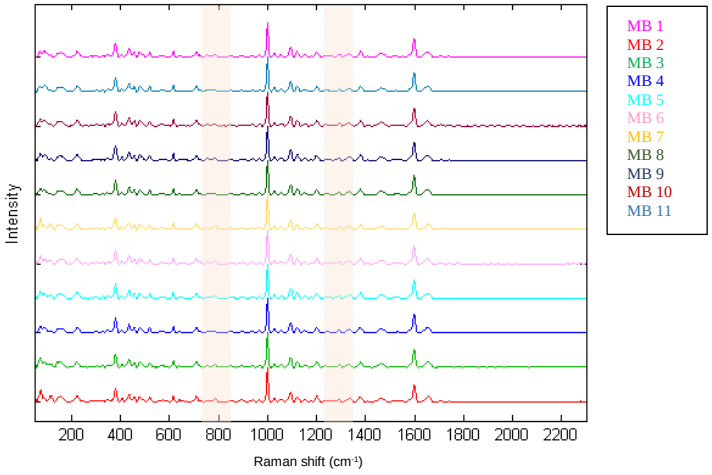


Figure 2. 26 Transmission Raman spectra changes depending on the polymorphic form contribution at the different MB for DKP-T 25 mg tablets.

Despite the polymorphic effect on the Raman spectra, the representativeness of the polymorphic distribution for the MB across the sample set along the different concentration ranges was demonstrated.

As a result, the calibration and the validation data sets were constructed with MB for DKP-T tablets to include the variability related to the manufacturing process.

2.2.2.6. Results and Discussion

This section presents the procedure development and validation for DKP-T 25mg tablets involving NIR and Raman spectroscopies.

2.2.2.6.1. Qualitative and Quantitative NIRS Procedures. Development and Validation

A PLS model was calculated for the active substance content of the DKP-T 25mg tablet allowing a linear relationship between the NIR spectra and the reference UV-visible values. Different PLS models were calculated. The strategy of leave-one-out cross-validation was used to optimize the model parameters. The figures of merit for the chemometric model development are shown in Table 2. 14.

Table 2. 14 Figures of merit for the chemometric model development of the active substance quantification by NIR spectroscopy for DKP-T 25mg tablet.

Parameter	PLS for active substance quantification
Number of samples included (laboratory samples and MB)	260 individual tablets (120 laboratory samples/ 140 MB)
Acquisition mode	Diffuse reflectance
Spectral pre-treatment	First derivative S-G with a window width of 5 points followed by vector normalization
Working range (cm^{-1})	9672-8168, 6488-5289, and 4920-4616
PLS factors	4
Explained variance (%)	94.5
Regression $Y_{UV\text{-}visible}$ vs. Y_{Raman}	
slope	0.91
Y-intercept	9.37
RMSEC/RMSECV (%)	3.7/3.96

A total of 260 individual tablets were included, of which 120 were prepared in the laboratory by the strategy of under and over-dosage of the manufacturing DKP-T granules and 140 belonged to MB.

The pre-treatment selected was the first derivative with a window width of 4 points followed by vector normalization. The working ranges included were from 9672-8168, 6488-5289, and 4920-4616 cm^{-1}.

Once the PLS was calculated the number of LV selected were 4, using as support information the graphic of RMSECV against the number of factors. During the model calculation, outlier spectra were found since the reference value was too low than the predicted value, so they were excluded.

The external validation was done with the validation set, including MB completely independent of the calibration set for model calculation[6,9]. The regression plot of predicted active substance content vs. the reference values for the external validation set is shown in Figure 2. 27. This figure includes the determination coefficient of the straight line as well the RMSEP and bias. Moreover, the scores plot for calibration and external validation set is shown in the same figure.

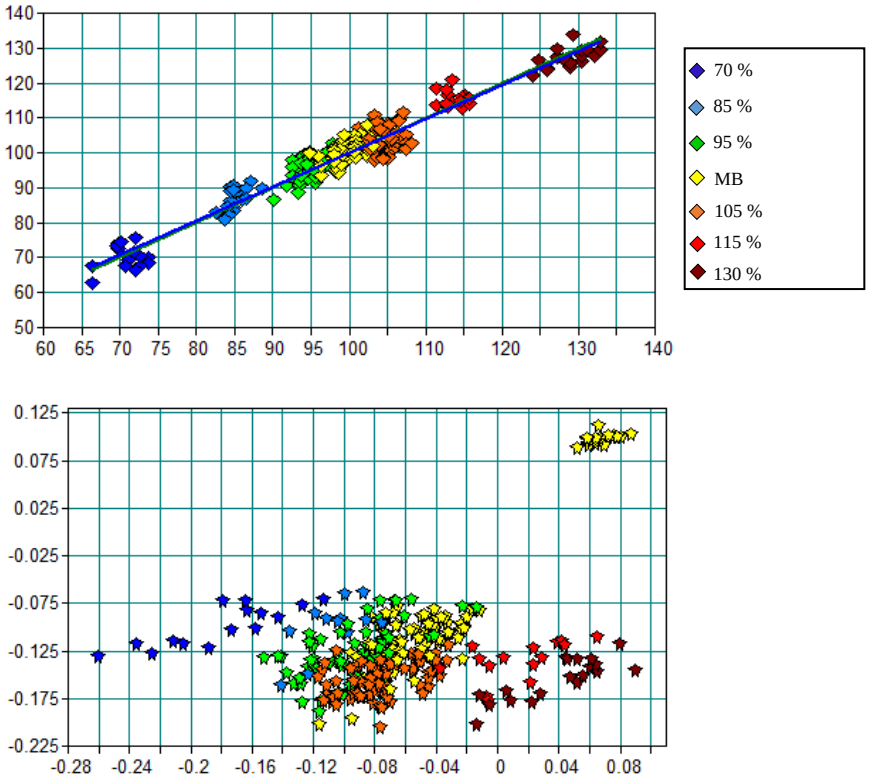


Figure 2. 27 Up) Predicted vs. reference values for the external validation set. Down) Scores plot for external validation set including first and second principal components.

The validation of specificity, linearity, accuracy, precision, and robustness was demonstrated following the specific guidelines[6,9,14]. The figures of merit are shown in Table 2. 15.

Table 2. 15 Figures of merit for the validation quantitative procedure for DKP-T 25mg tablet by NIRS.

Parameter	**Results**	
Specificity	-The main peaks of the active substance were presented in the PLS factors. -Final product library developed based on a sub-cascading spectral strategy for the calibration set, allowing the unequivocal identification of each group of samples included in the library (see Figure 2. 29) -MD was automatically evaluated for each included sample (see Figure 2. 30) as a threshold of similarity within samples	
Linearity	Num. of samples	200 individual tablets (120 laboratory samples/80 MB)
	Range (%)	65.87-133.0
	Explained variance (%)	95.0
	y-intercept ± IC (α=0.05) including 0	2.87 ± [-0.066:5.77]
	Slope ± IC (α=0.05) including 1	0.973 ± [0.94:1.00]
	R>0.95	0.968
	The residuals between the reference UV-visible and the NIR method did not show any trend regarding concentration levels (see Figure 2. 28)	
Accuracy	Number of samples	200 individual tablets (120 laboratory samples/80 MB)

Table 2. 15 Figures of merit for the validation quantitative procedure for DKP-T 25mg tablet by NIRS.

Parameter	Results	
	Paired t-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 0.99 *t*-tabulated: 1.97 *p*-value: 0.32
	RMSEP (%)	2.94
Precision	SEP<1.4SEL	2.94<1.4x3.04
	Validation bias value close to 0	-0.17
Repeatability	RSD (%) for 10 measures of the same sample less than 5%	2.9
Intermediate Precision	One-way ANOVA factor (day): *p*-value higher than 0.05 *F*-calculated lower than the *F*-tabulated	*F*-calculated: 2.46 *F*-tabulated: 3.35 *p*-value: 0.10
Robustness	Number of samples	1060 individual tablets (106 MB)
	Paired t-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 0.03 *t*-tabulated: 1.98 *p*-value: 0.97
Control Model Performance	Centre line (CL), $\overline{x}$ Upper and Lower action limits (UAL/LAL), ±3σ Upper and Lower warning limits (UWL, LWL), ±2σ *Based on residuals*	UAL : 7.05 UWL: 4.70 $\overline{x}$: 0.001 LWL : -4.70 LAL : -7.05

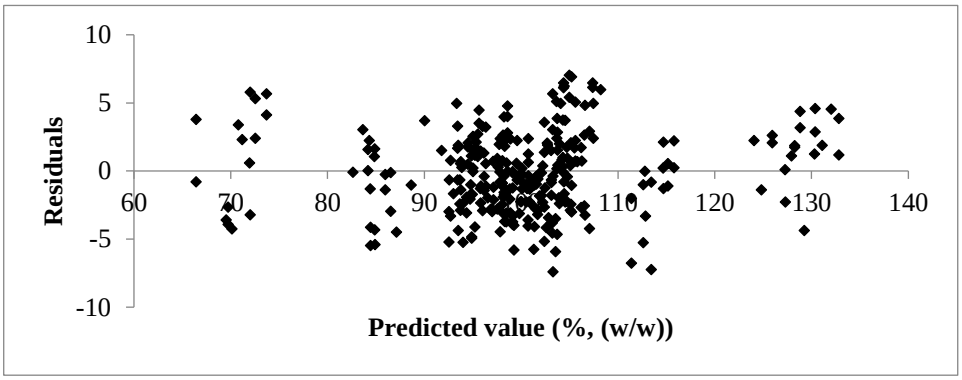


Figure 2. 28 Residuals plot for the external validation set ordered by concentration level for the NIR model

The specificity was tested considering the internal validation for the spectral library including the calibration spectra, based on the sub-cascading spectral library strategy. Each spectrum was used for calibration and internal validation. The internal validation showed that all calibration spectra belong to each assigned group of each SM and there were no ambiguities during the calibration identification. The reference libraries were validated by checking that the external spectra (referred to as the validation set) were correctly and unambiguously identified. The correct identification for the spectral library was guaranteed to assure the quality of the results when the amount of active substance will be calculated using the chemometric model developed. The library structure is shown in Figure 2. 29. At the first level all the SM were included. The ones with an ambiguous qualification

were included in a sub-library at the second level (the active substance and the cores and tablets for DKP-T 25 mg). In the end, a third-level sub-library was necessary to unambiguously qualify both cores and tablets.

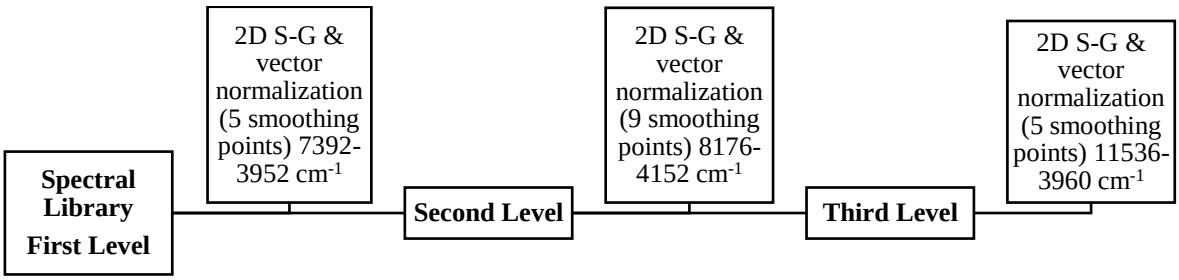


Figure 2. 29 Sub-cascading spectral library structure for DKP-T 25 mg tablets unequivocally identification.

Moreover, during the model validation, the MD was automatically calculated. This is a statistics useful spectral quality test to detect spectral outliers. The limit was fixed by the software at 0.055. Those new samples with lower MD than the critical value was considered spectrally similar to the calibration set, so the predicted values obtained were reliable. Whereas when the distance calculated was higher than the critical one, it meant that the spectrum of the new sample presented some differences regarding the calibration set, so the predicted value could be affected by this feature. As can be seen in Figure 2. 30 no outliers were detected.

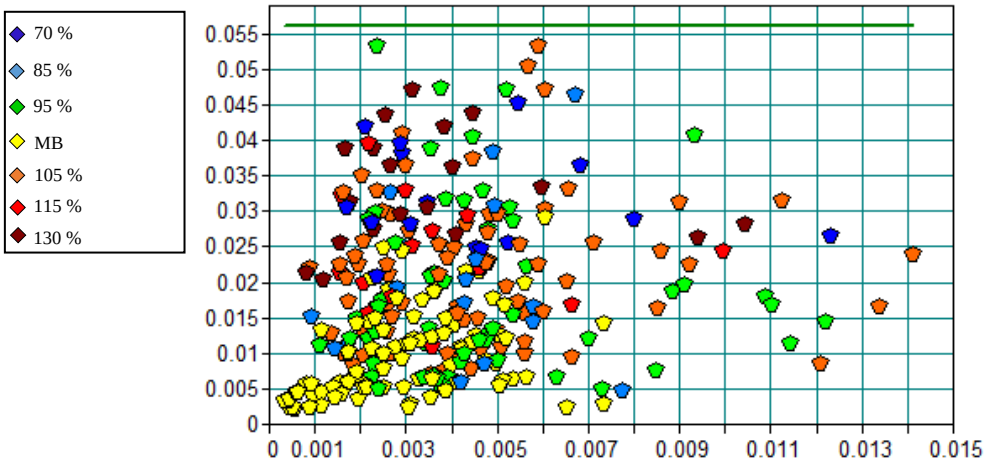


Figure 2. 30 Scatter plot for MD for each included sample into the validation set of DKP-T 25mg tablet.

2.2.2.6.2. Raman Spectroscopy Procedure. Development and Validation

A PLS model was calculated for the active substance content of the DKP-T 25mg tablet allowing a linear relationship between the Raman spectra and the reference UV-visible values. Different PLS models were calculated. The strategy of leave-one-out cross-validation was used to optimize the model parameters. The figures of merit for the model development are shown in Table 2. 16.

Table 2. 16 Figures of merit for the model development of the active substance quantification by Raman spectroscopy for DKP-T 25mg tablet.

Parameter	PLS for active substance quantification
Number of samples included (laboratory samples and MB)	140 individual tablets (120 laboratory samples/20 MB)
Acquisition method	Transmission
Spectral pre-treatment	Smoothing with a window width of 15 points followed by baseline correction by an automatic weighted least square second order, and normalization
Working range (cm^{-1})	979-1614
PLS factors	3
Explained variance (%)	98.63
Regression $Y_{UV\text{-}visible}$ vs. Y_{Raman}	
slope	2.05
Y-intercept	0.98
RMSEC/RMSECV (%)	2.65/2.72

A total of 140 individual tablets were included, of which 120 were prepared in the laboratory by the strategy of under and over-dosage of the manufacturing DKP-T granules and 20 belonged to MB.

The pre-treatment selected was smoothing with a window width of 15 points followed by baseline correction by an automatic weighted least square second order, and normalization. The working range included was 979-1614 cm^{-1}. The spectral range between 200 and 1700 cm^{-1} was not presented spectral noise. Can be observed as an over-saturation up to around 200 cm^{-1}. This spectral range was removed for future calculations and analysis. To optimize the model, the strategy of leave-one-out cross-validation was used.

Once the PLS was calculated the number of LV selected were 3, using as support information the graphic of RMSECV against the number of factors.

During the model calculation, outlier spectra were found since the reference value was too low than the predicted value, so they were excluded.

The external validation was done with the validation set and production samples completely independent of the calibration set for model calculation[6,9]. The regression plot of predicted active substance content vs. the reference values for the external validation set is shown in Figure 2. 31. This figure includes the determination coefficient of the straight line as well the RMESEC/RMSECV/RMSEP and the bias. Moreover, the scores plot including the calibration and the external validation sets is shown in the same figure showing a sample trend related with the active substance concentration.

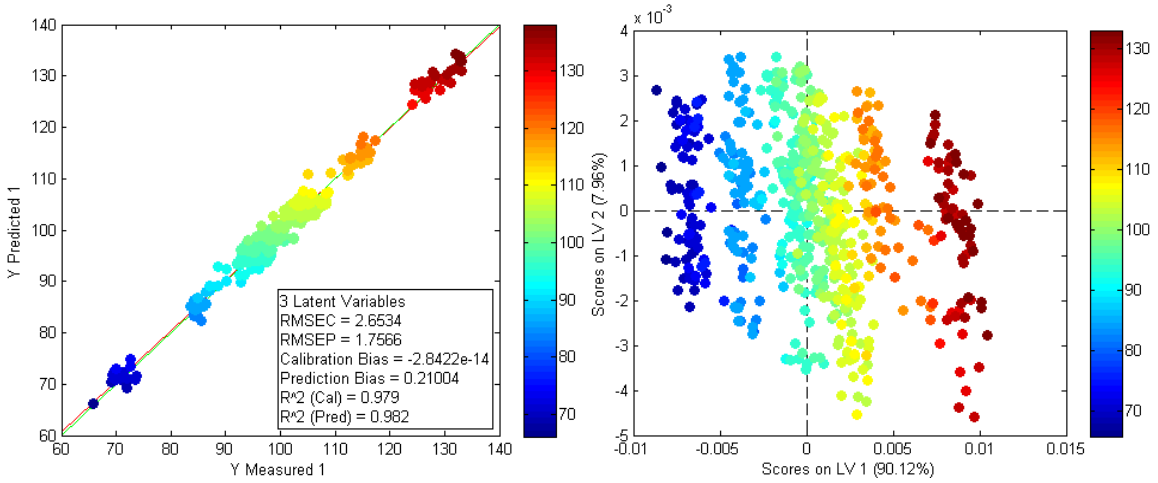


Figure 2. 31 Left) Predicted values vs. reference values for the external validation set. Right) Scores plot for calibration and external validation sets including the two first PCs. The 98.08 % of the total variability is explained in this plot.

The validation of specificity, linearity, accuracy, precision, and robustness were demonstrated following the specific guidelines[6,14,15]. The figures of merit are shown in Table 2. 17.

Table 2. 17 Figures of merit for the validation quantitative procedure for DKP-T 25mg tablet by Raman spectroscopy.

Parameter	**Results**	
Specificity	-The main peaks of the active substance were presented in the PLS loadings -Final product library developed based on a PLS-DA analysis for the calibration set, allowing the unequivocal identification of each group of samples included in the library evaluating the confusion matrix for the model without the presence of any false positive or negative results	
Linearity	Number of samples	140 individual tablets (120 laboratory samples/20 MB)
	Range (%)	65.87-133.0
	Explained variance (%)	97.93
	y-intercept ± IC (α=0.05) including 0	1.30 ± [-0.13:2.74]
	Slope ± IC (α=0.05) including 1	0.99 ± [0.98:1.00]
	R>0.95	0.98
	The residuals between the reference UV-visible and the Raman method did not show any trend regarding concentration levels (see Figure 2. 32)	
Accuracy	Number of samples	140 individual tablets (120 laboratory samples/20 MB)
	Paired *t*-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 1.92 *t*-tabulated: 1.97 *p*-value: 0.056
	RMSEP (%)	1.76
Precision	SEP<1.4SEL	1.76<1.4x3.04
	Calibration and validation bias value close to 0	-2.84^{-10}:1.19
Repeatability	RSD (%) for 10 measures of the same sample less than 5%	0.85

Table 2. 17 Figures of merit for the validation quantitative procedure for DKP-T 25mg tablet by Raman spectroscopy.

Parameter	Results	
Intermediate Precision	One-way ANOVA factor (day): *p*-value higher than 0.05 *F*-calculated lower than the *F*-tabulated	*F*-calculated: 1.19 *F*-tabulated: 3.35 *p*-value: 0.32
Robustness	Number of samples	300 individual tablets (26 MB)
	Paired t-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 1.46 *t*-tabulated: 1.96 *p*-value: 0.14
Control Model Performance	CL: $\bar{x}$ UAL/LAL, ±3σ UWL/LWL, ±2σ *Based on residuals*	UAL: 3.18 UWL: 2.10 $\bar{x}$: -0.07 LWL: -2.23 LAL: -3.32

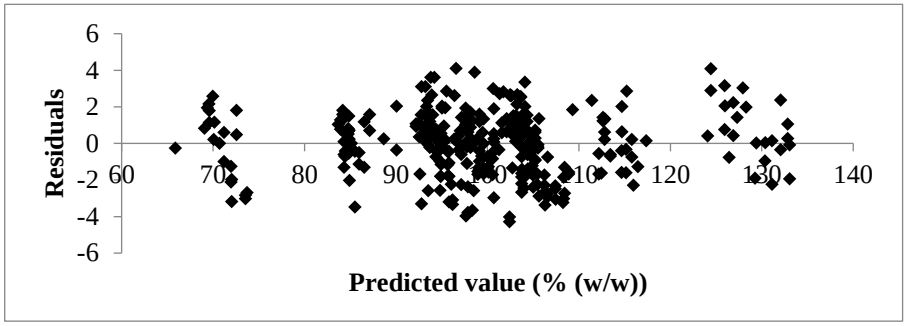


Figure 2. 32 Residuals plot for the external validation set ordered by the concentration level of the Raman method.

2.2.3. Atorvastatin Film-coated Tablets

Atorvastatin was selected as the third product of interest. This product is manufactured at AMMLS FL (Italy).

This section presents an evaluation of the composition and the manufacturing method for the product. Moreover, a risk analysis was developed to understand the factors that can affect the NIR and Raman. Finally, an exploratory analysis was developed to understand the behaviour of the samples and be able to establish a proper strategy for the calibration and the validation data sets preparation, including coating and manufacturing strategy effect on the NIR and Raman spectra.

2.2.3.1.Composition and Manufacturing Process

Atorvastatin product is a film-coated tablet available in four different strengths (10, 20, 40, and 80 mg/tablet). The information described in Table 2. 18 applies to the most frequent manufacture strengths being 10, 20, and 40 mg/tablet. The relative composition of the formulation is identically for all strengths. As a result, the desired strength is obtained modified the final tablet weight and size (see Table 2. 18).

Table 2. 18 Composition of the Atorvastatin film-coated tablet.

	Description	Content per tablet (mg)		
		10 mg	**20 mg**	**40 mg**
Core	Atorvastatin calcium	10.9	21.7	43.4
	Tween 80	0.400	0.800	1.60
	Magnesium stearate	0.500	1.00	2.00
	Hydroxypropyl cellulose	2.00	4.00	8.00
	Sodium croscarmellose	6.00	12.0	24.0
	Microcrystalline cellulose 50µm	20.0	40.0	80.0
	Lactose monohydrate	27.2	54.5	109
	Calcium carbonate	33.0	66.0	132
	Core weight	100	200	400
Film-coated tablet	Opadry lacquer	3.58	7.15	14.3
	Silicone emulsion	0.0240	0.0480	0.0960
Nominal target weight for the film-coated tablet		103.6	207.2	414.4

Through the manufacturing process, the powder is granulated by a wet granulation process. As a result, the dried granulate, after mixing with extra granular excipients, is tabletted to obtain the three strengths with the appropriate weights. Finally, the cores are film-coated (see Figure 2. 33).

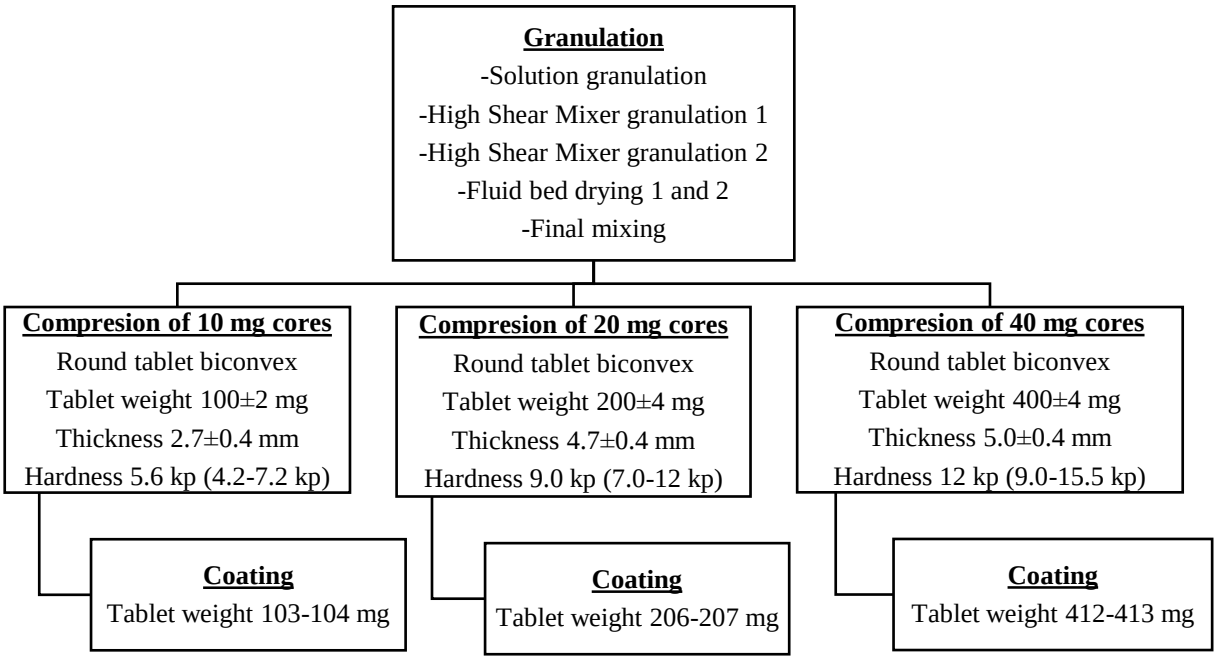


Figure 2. 33 General flow sheet for Atorvastatin film-coated tablet manufacturing process.

2.2.3.2.Samples for Preliminary Investigation

The main strategy for laboratory sample preparation was based on the under and overdose production samples after the granulation and the final mixing step. Then the samples were compressed at the

pilot plant. Three concentration levels were included between the expanded range ±30%: 70, 100, and 130. The laboratory sample (from now abbreviated as LB) prepared at the nominal active substance value includes 70% of granules from the production and 30% of the powder product mixture at the nominal active substance value. Moreover, cores from MB and cores prepared at the pilot plant from the granules of the production (referred to as MB-LB) were included. The sample description is detailed in Table 2. 19.

Table 2. 19 Samples included in the preliminary investigations.

Manufacturing steps included	Active substance concentration (%)	Sample name
Granulation and compression during the manufacturing process	100	MB
Manufacturing granulation, direct mixture by under and overdosage strategy, and pilot plant compression	70, 100, and 130	LB
Manufacturing granulation and pilot plant compression	100	MB-LB

2.2.3.3.Instrumentation

The FT-NIR MPA II equipment (Bruker Optics, USA) was used for NIR spectra acquisition. The spectra were acquired in diffuse reflectance mode involving a specific tablet accessory for independent measures. The parameters selected were 32 scans with 16 cm^{-1} of resolution.

The TRS100 instrument (Agilent, Oxford) was used for Raman data acquisition. The instrument and its related measure parameters are described in Table 2. 20.

Table 2. 20 TRS100 Raman parameters for Atorvastatin product spectra acquisition.

Parameter	Specification
Model	Analyzer TRS100 Raman
Sampling accessory	External Transmission
Laser power (W)	0.65
Beam size (mm)	2
Accumulations	100
Exposure (secs)	2.5
Total exposure (secs)	250

2.2.3.4.Risk Assessment

An Ishikawa diagram was developed for a cause-and-effect risk analysis (see Figure 2. 34).

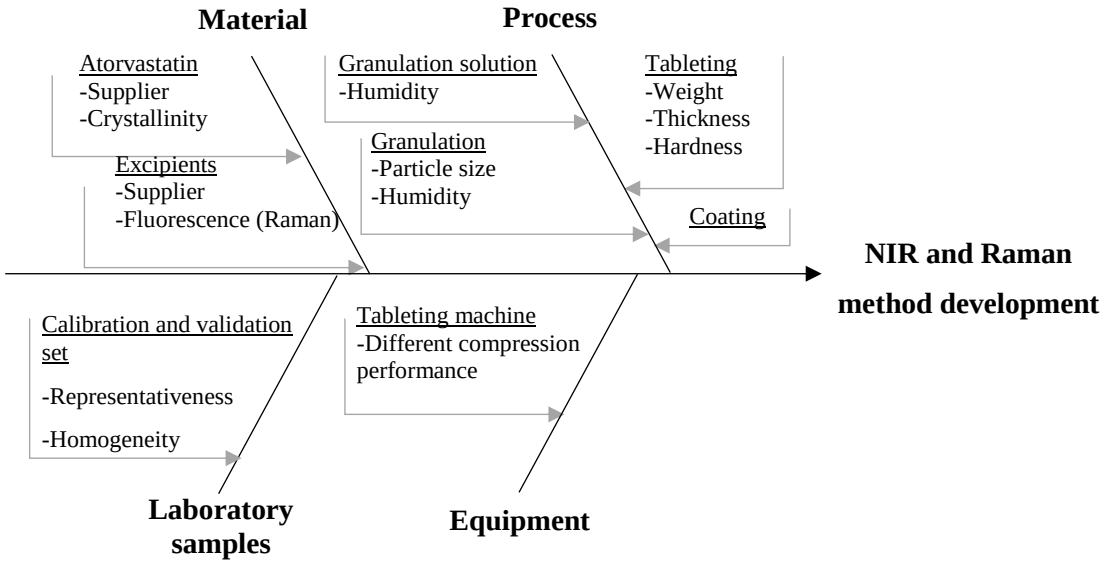


Figure 2. 34 Ishikawa diagram for Atorvastatin film-coated tablet NIR and Raman method development.

A risk mitigation strategy was adopted based on the cause-and-effect analysis which is described in Table 2. 21. The severity ranking was considered as detailed in Table 2. 11.

Table 2. 21 Cause and effect risk analysis for Atorvastatin film-coated tablet manufacturing.

Cause	Rank	Effect	Mitigation
Material Active substance	1	None expected	Comparison of possible active substance polymorphic transformation during manufacturing by X-ray analysis. Possible variability will be included with MB
Material Active substance. Supplier	1	None expected	One manufacturer will be included in the calibration and validation set.
Material Excipients. Supplier	1	None expected. For all excipients, only one manufacturer is homologated	Does not apply
Material Excipients. Physical properties	1	None expected. The initial physical characteristics of granular excipients don't affect the final particle size or water content	Does not apply

Table 2. 21 Cause and effect risk analysis for Atorvastatin film-coated tablet manufacturing.

Cause	Rank	Effect	Mitigation
Material Excipients. Microcrystalline cellulose. Fluorescence	4	Medium risk expected for Raman. Interference in the spectrum	Select the appropriate spectral range during the model development as well as the correct data pre-treatment
	1	None expected for NIR	Does not apply
Laboratory samples Calibration and validation set development	4	Heterogeneous laboratory samples or non-equivalent laboratory tablets from MB	Each laboratory step to obtain Atorvastatin galenic tablets will be verified involving validated physicochemical assays
Equipment Tableting machine	2	Differences in NIR and Raman spectra due to differences in pressure performance of each piece of equipment	The expected variability will be introduced on the calibration and validation set with MB with both machines
Process Granulation. Crystallinity	3	Medium risk expected. Active substance becomes a mix of crystalline forms. Differences in NIR and Raman spectra	Expected variability will be introduced on the calibration and validation set with production samples
Process Granulation. Particle size	1	Low risk. Particle size should have a low effect on the Raman spectrum	Expected variability will be introduced with MB
	4	Medium risk. The particle size has a huge effect on the NIR spectrum (scattering effect)	Expected variability will be introduced on the calibration and validation set with MB
Process Granulation. Humidity	1	None expected. Water content doesn't affect the Raman spectrum	Does not apply
	4	Medium risk. NIR is very sensitive to water content	Expected variability will be introduced on the calibration and validation set with MB
Process Tabletting. Thickness and hardness	3	Medium risk expected. Differences in the NIR and Raman spectra due to differences in tablet thickness and hardness	The variability will be introduced on the calibration and validation set with MB
Process Film-coating	3	Medium risk expected for Raman due to differences in coating performance	Expected variability will be introduced on the calibration and validation set with MB. A preliminary study was conducted to compare cores and tablets
	3	Medium risk expected for NIR due to differences in coating performance	Expected variability introduced on the calibration and validation set with MB. A preliminary study was conducted to compare spectra for cores and tablets

2.2.3.5. Investigation of the Coating Effect to the NIR and Raman Spectra

The strategy followed to carry out this investigation is the next one: to capture as much variance as possible, three different batches from the manufacturing process at the three strengths were selected carefully. The sample set includes uncoated and coated tablets. The spectrum from ten tablets was acquired individually with the NIR, and for five tablets with the Raman equipment (see Table 2. 22).

Table 2. 22 MB for Atorvastatin cores and film-coated tablets at 10, 20, and 40 mg strength.

Batch name	Strength (mg)	Active substance content (%)	Type of sample	Number of tablets measured by NIRS	Number of tablets measured by Raman
MB-A	10	100	Uncoated	10	5
			Coated	10	5
MB-B	20	100	Uncoated	10	5
			Coated	10	5
MB-C	40	100	Uncoated	10	5
			Coated	10	5

An exploratory analysis based on PCA analysis was developed to extract the information from the spectra obtained to evaluate the coating effect in the NIR and Raman spectra.

For the NIR spectra, significant differences between cores and film-coated tablets were observed due to the presence of the 2≡Si-OH bond between the wavelength range from 7300 to 4500 cm^{-1} included in the coated agent[16] (see Figure 2. 35). As a result, the scores of the PCA analysis showed a cluster trend along the PC1 referred to the coating effect. Moreover, the dosage trend was observed along the PC2 (see Figure 2. 36).

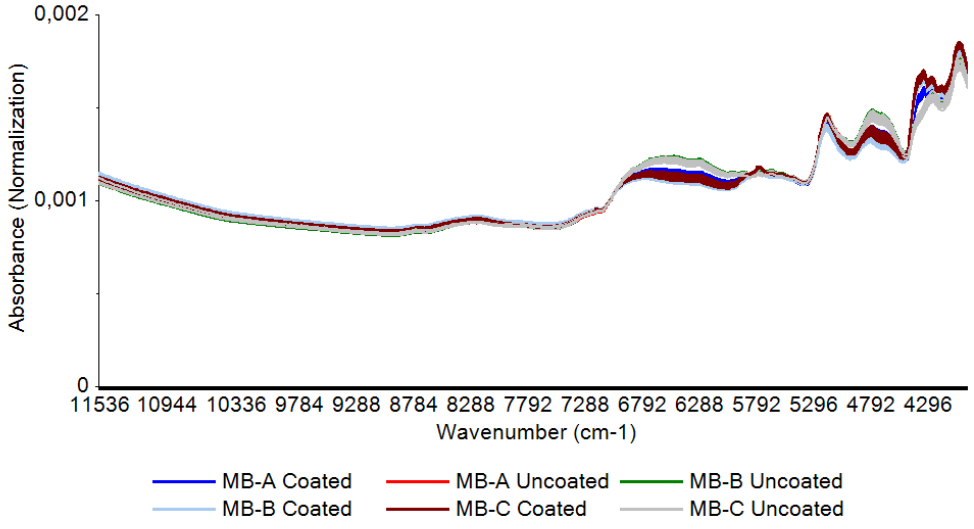


Figure 2. 35 Diffuse reflectance NIR spectra for Atorvastatin MB considering the three strengths after vector normalization as data pre-treatment including the range from 11500 to 4000 cm^{-1}.

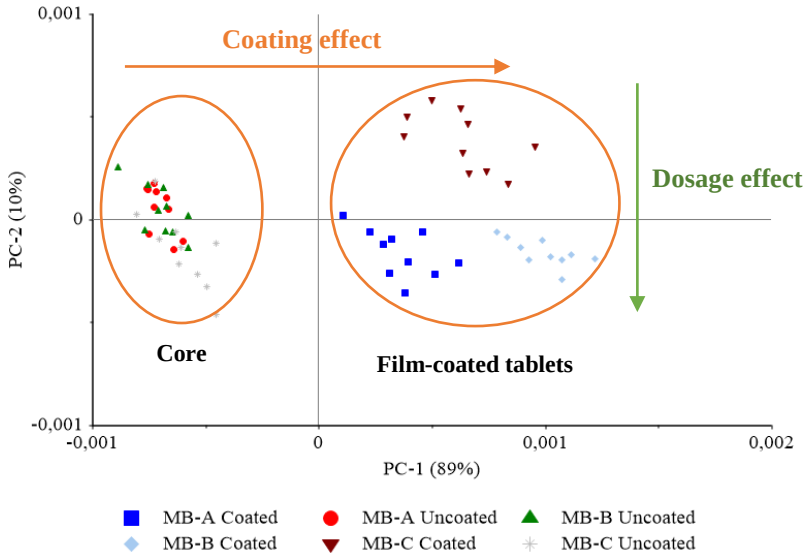


Figure 2. 36 PC1 vs PC2 for the scores plot for the cores and tablets of Atorvastatin product for the NIR spectra. The 99% of the variability is explained.

For the Raman instrument, the same spectra profile was observed for all the evaluated batches. Differences in the Raman intensity were observed in the interval range below 200 cm^{-1} (see Figure 2. 37). As a result, the PCA analysis showed a cluster trend along PC1 related to the coating effect and a cluster trend related with the different levels of the active substance along the PC2 (see Figure 2. 38).

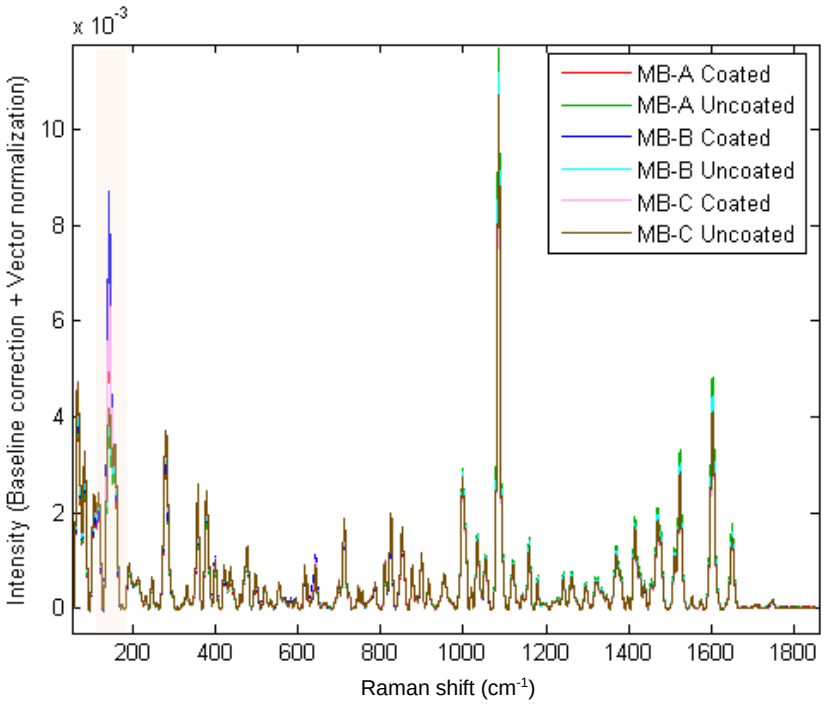


Figure 2. 37 Transmission Raman spectra for the Atorvastatin MB including cores and film-coated tablets, for the three strengths after baseline correction and vector normalization as data pre-treatment including the range from 50 to 1850 cm^{-1}.

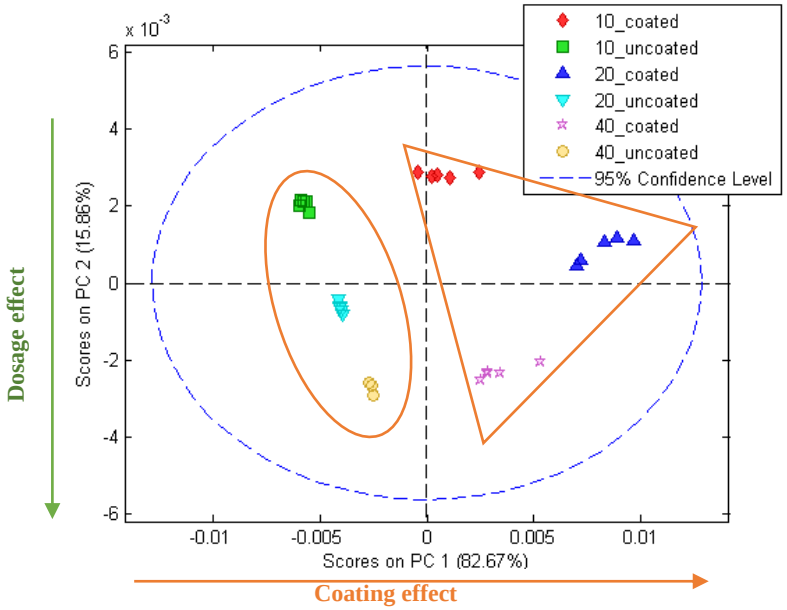


Figure 2. 38 Scores plot (PC1 vs PC2) for the cores and film-coated tablets of Atorvastatin product for the Raman spectra. The 98.53% of the total variability is explained.

As a result, tablets from MB and laboratory scale should be included in the calibration and validation sets.

2.2.3.6. Investigation of Laboratory Sample Preparation Strategy on NIR and Raman Spectra

The manufacturing process includes five steps for Atorvastatin film-coated product: granulation and drying, extra granular mixing, tableting, coating, and packaging (see Figure 2. 33).

The laboratory samples have to be as representative as possible from the MB[6]. To understand if there are significant spectral differences between the MB and the laboratory samples prepared by the under and over-dosage strategy, a preliminary study was developed.

Six different batches were analyzed. For each one ten tablets were analyzed by NIR spectroscopic technique. Three of these batches correspond to MB at the three studied strengths. The other three ones were laboratory batches prepared by under and over-dosage strategy including powder placebo and active substance at the nominal value, from the MB granules (named as MB-LB in Table 2. 19).

The spectra for each sample preparation (compression before fluid-bed granulation (MB) and compression before fluid-bed + under- and over-dosage strategy for laboratory samples (MB-LB)) exhibited an identical spectra profile (no difference in chemical composition between samples was apparent); however, the spectra were shifted by the effect of differences in physical properties between sample preparation processes. High absorbance due to high pressure due to the higher powder contribution in the tablet with a LB sample preparation[17] (see Figure 2. 39, as an example).

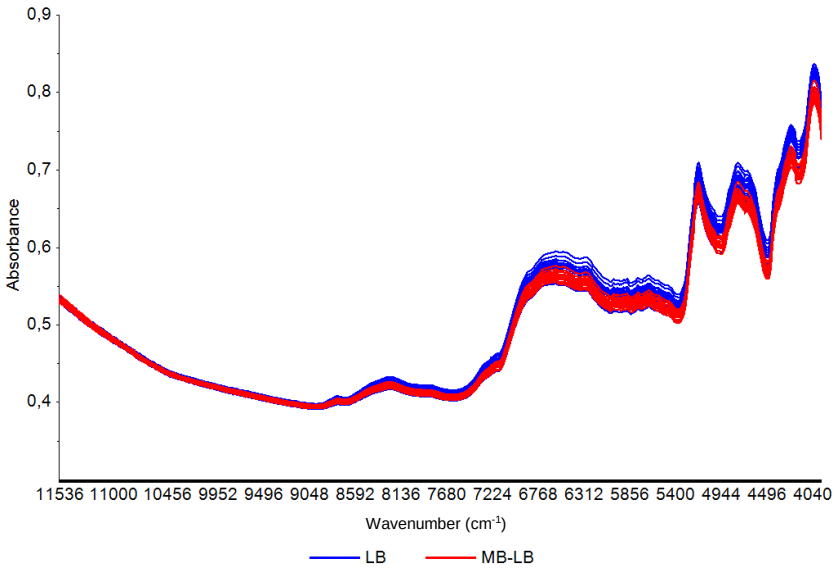


Figure 2. 39 Diffuse reflectance NIR spectra for the Atorvastatin cores at 40 mg.

There was not possible to evaluate the samples by Raman spectroscopy due to logistic inconvenience.

As a result, the under and overdosage strategies could be used for laboratory sample development.

2.2.3.7.Suitability Study for PLS Modelling by NIR and Raman Spectroscopy

The linearity between the NIR and Raman spectra and the active substance was evaluated. The strategy followed was the next one: the active substance concentration was expanded between ±30% from the nominal value. The lower and the upper concentration values for the studied range (70 and 130%, respectively) were prepared by the under and upper-dosage strategy, respectively, from the granulated MB, for each studied strength. Once the mixture was prepared the samples were compressed, with a process as representative as possible of the production one (named LB in Table 2. 19). As a result, three different concentrations were analyzed per strength. For each sample, ten cores were analyzed (see Table 2. 23).

Table 2. 23 Included samples for the suitability study for the PLS modelling by NIR and Raman spectroscopy.

Strength (mg)	Batch name	Active substance concentration (%)	Type of sample	Number of tablets measured by NIRS	Number of tablets measured by Raman
10	LB-A	70	Uncoated	20	10
	MB-A	100	Uncoated	10	5
	MB-LB-A	100	Uncoated	20	-
	LB-A	100	Uncoated	20	10
	LB-A	130	Uncoated	20	10
20	LB-B	70	Uncoated	20	10
	MB-B	100	Uncoated	10	5
	MB-LB-B	100	Uncoated	20	-
	LB-B	100	Uncoated	20	10
	LB-B	130	Uncoated	20	10
40	LB-C	70	Uncoated	20	10
	MB-C	100	Uncoated	10	5
	MB-LB-C	100	Uncoated	20	-
	LB-C	100	Uncoated	20	10
	LB-C	130	Uncoated	20	10

A trend based on the concentration active substance range was achieved for all the strengths before vector normalization and the second derivative S-G data pre-treatment for NIR spectra (see Figure 2. 40).

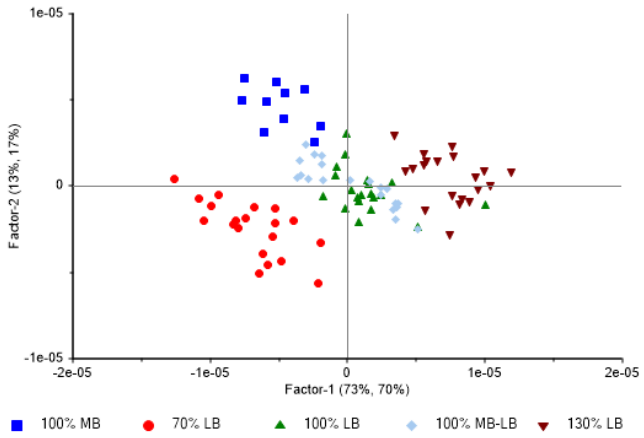


Figure 2. 40 PC1 vs PC2 for the scores plot for the cores and tablets of Atorvastatin product for the NIR spectra. The 83% of the total variability is explained.

A linear relationship based on the active substance concentration range was achieved for all the strengths (see Figure 2. 41).

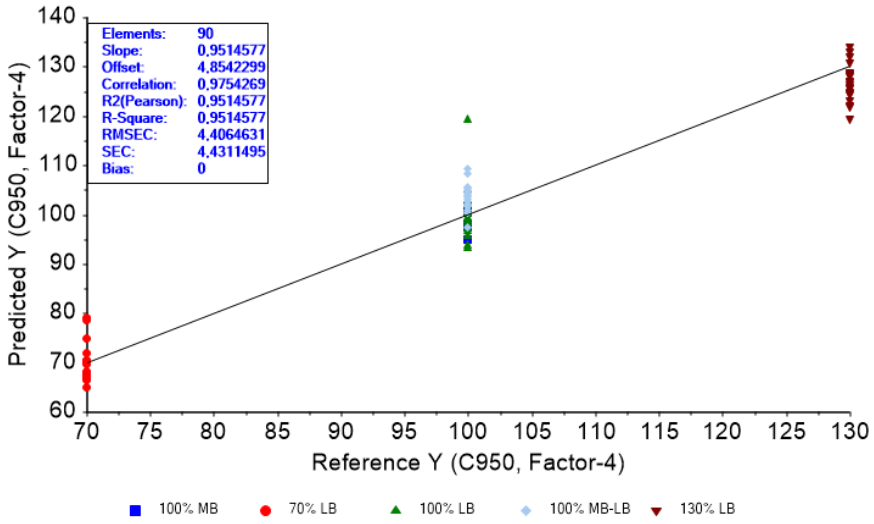


Figure 2. 41 Linear regresion based on PLS analysis for the cores and tablets of Atorvastatin product for the Raman spectra.

Moreover, for Raman spectra, the same results were achieved after baseline correction, and vector normalization as data pre-treatment

2.2.4. Conclusions

The application of qualitative analysis using PLS-DA and sub-cascading spectral library for the identification and qualification by Raman and NIR spectroscopies and the PLS for quantification of both technique spectra for the DKP-T 25 mg tablets was satisfactorily demonstrated.

Moreover, the preliminary study for the Atorvastatin tablet demonstrates the linear relationship between NIR and Raman spectra and the active substance concentration.

The design of experiments for the selection of the strategy for sample preparation, the optimization of the spectral acquisition parameters, and the construction of a representative calibration and validation set, were essential to guarantee the applicability of the models.

The prediction model ability and the development methodologies for the qualitative and quantitative analysis were validated considering the guidelines criteria. For both techniques, a correct ability was demonstrated. As a result, these newly developed methodologies will be replacing the current authorized method to determine assay and content uniformity routinely.

In conclusion CQA active substance content of a pharmaceutical preparation can be controlled during the manufacturing process.

2.3. Develop a NIRS Procedure for Identifying and Qualifying a Powder Bulk Product and Quantifying the Assay for the Content Uniformity Analysis

The homogeneity of a powder product is guaranteed involving a blending step as an essential unit operation during the pharmaceutical manufacturing process including all the formulation compounds (the active substance and the excipients). The final mixture homogeneity depends on the unit operation procedure, and the physical component properties (such as size, density, and shape). The main components in the studied formulation, for the Sucramal product, include two specific sorbitol types, presenting a 2:1 difference in particle size. This different nature can lead to segregation by a sifting mechanism and cause the separation of the compounds in the mixture into two layers, in the upper part those containing the larger particles and in the lower part the smaller ones[18,19]. Avoiding this phenomenon is key to ensuring the consistency of a high-quality product, guaranteed by achieving the CQA evaluated for each batch analyzed within the registered specifications. These determinations are necessary to release the product.

The determination of the CQAs in the mixture is currently done by the register HPLC method which is very time-consuming and expensive, requiring a specific mobile phase, column, and detector.

The scope of this chapter is to present a NIRS procedure as a non-invasive technique to determine in a representative manner[20] two CQAs (identification and the active substance concentration for uniformity content assay) of a Sucramal product in an offline mode, as an alternative to the reference HPLC method currently registered.

2.3.1. Composition and Manufacturing Process

The manufacturing process is summarized in Figure 2. 42. Nowadays, the determination of these CQAs, requires the use of time-consuming analysis in the QC laboratory.

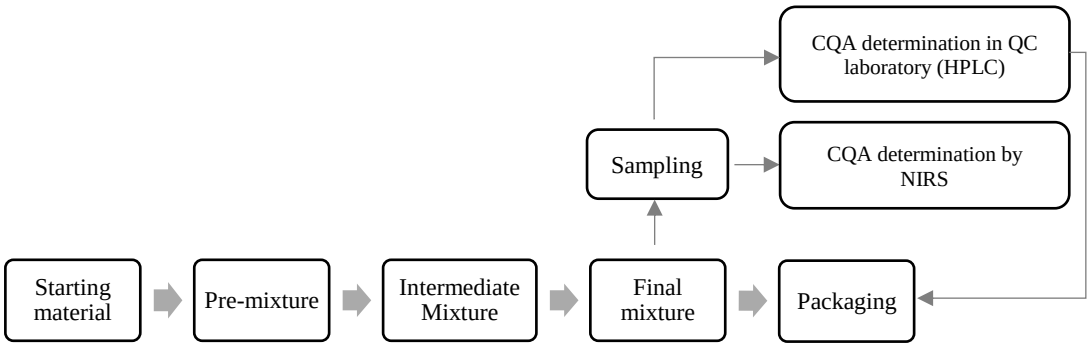


Figure 2. 42 Production process including the manufacturing and the CQAs determination for Sucramal product.

The composition of the studied product is presented in Table 2. 24. The product presents Sucralfate as the active substance and represents 16.7% (w/w) of the total amount. The sorbitol is the main excipient (82.8% (w/w)).

Table 2. 24 Composition of Sucramal product.

Description	Proportion (% (w/w))
Sucralfate	16.67
Sucrose	0.13
Vanilla scent	0.42
Sorbitol GDO P60G	20.70
Sorbitol GDO 30/60	62.08

Once the SM are sieved a pre-mixture is prepared including the vanilla scent and sucrose. Then an intermediate mixture is prepared in a mixer roller including a few percent of sorbitol and sucralfate. Then the final mixture is done including the rest of the sorbitol and the pre-mixture in a stainless-steel blender in V form.

The CQAs are listed in Table 2. 25. In this table are indicated the reference analytical methods used to determine each CQA as well as their internal specifications.

Table 2. 25 Reference method and specification of CQAs for Sucramal product.

CQA	Reference method	Specification
Identification	HPLC	Pass
Active substance content (mg/g)	HPLC	$52.4 \leq X \leq 60.9$
Variation coefficient related to the active substance (%)	HPLC	≤ 10.0 %

2.3.2. Materials and Methods

In this section, the materials and methods involved in the development and validation of qualitative and quantitative procedures involving offline NIR analysis for Sucramal product are detailed.

2.3.2.1.Calibration and Validation Data Sets for Spectroscopic Qualitative Analysis

The samples used for the development and validation of the qualification analysis by NIRS for the Sucramal product are summarized in Table 2. 26. All production batches were properly prior identified by the reference method. The active substance sucralfate, the excipients, and placebos, which were prepared with all SM except the active substance, were used to test the specificity of the NIRS procedure. The samples used in the calibration set should include all possible sources of

variability to ensure that the calibration covers all ranges of interest. A set of MB was carefully selected to include the variability source associated with the product and the manufacturing process.

Table 2. 26 Samples involved to calibrate, internally, and externally validate the qualification analysis for Sucramal product by NIRS.

Product	Sample type	Parameter	Justification
Sucramal product	Calibration, internal and external validation sets	Specificity and Robustness	Chemical and physical variability
Sucralfate	Calibration, internal and external validation sets	Specificity	Active substance
Placebo	Calibration, internal and external validation sets	Specificity	Without active substance
Sorbitol GDO P60G	Calibration, internal and external validation sets	Specificity	Main excipient
Sorbitol GDO 30/60	Calibration, internal and external validation sets	Specificity	Main excipient
Vanilla scent for sucramal	Calibration, internal and external validation sets	Specificity	Excipient
Sucrose	Calibration, internal and external validation sets	Specificity	Excipient

2.3.2.2.Calibration and Validation Data Sets for Spectroscopic Quantitative Analysis

The calibration set has to include the variability sources that potentially affect the NIR spectra. The samples used to develop the qualitative analysis come from different populations. The first one comes from production batches to consider the expected process variations, covering enough timeline to assure the introduction of the following factors: different product temperatures, different operators, different temperatures, and relative humidity from the environment, and NIR equipment wear away. The second one comes from laboratory samples to include the necessary concentration range[6,9,15]. The active substance content was expanded by ±30% regarding the nominal content. The strategy followed to expand the narrow range for active substance content determination was based on the under and over-dosage strategy in a pilot scale plant from a direct mixture of placebo or active substance, respectively, to expand the concentrations from 70 to 130%. Seven concentration levels were included: 70, 85, 100, 115, and 130%, as well as 92.5 and 107.5%, assigned as the acceptance intervals for the evaluated CQA. 100 g of galenic samples were prepared. Using a mixture design strategy, three placebos were prepared to expand the excipient variability for the final mixture around 7.5%. The strategy followed was to modify by around 7.5% the composition of those excipients considered majority (higher than 7.5%) from the nominal value of placebo composition; sorbitol GDO 30/60 and P60G, and a mixture from vanilla scent and sucrose, reducing as much as possible the correlation between formulation components (see Table 2. 27).

Table 2. 27 Placebo composition (% (w/w)) used for calibration and validation set manufacturing.

Sorbitol GDO 30/60	Sorbitol GDO P60G	Vanilla scent for Sucramal	Sucrose	Placebo code
76.37	22.98	0.50	0.16	1
74.51	24.83	0.50	0.16	2
72.65	26.69	0.50	0.16	3

Different batches were used for placebos included in calibration and validation sets. Only two replicates for each concentration were prepared due to the difficulties during the experimental procedure for the reference method already registered. The prepared laboratory samples are summarized in Table 2. 28.

Table 2. 28 Samples composition for calibration and validation sets prepared by direct mixture for Sucramal product.

Calibration set			Validation set		
Concentration level (%)	**Placebo code**	**Active Substance Code**	**Concentration level (%)**	**Placebo code**	**Active Substance Code**
70	2 1	B1 B2	70	3 1	B3
85	2 3	B1 B2	85	3 1	B3
92.5	3 1	B1 B2	92.5	1 2	B3
100	3 3	B1 B2	100	2 1	B3
107.5	3 1	B1 B2	107.5	2 3	B3
115	3 3	B1 B2	115	2 1	B3
130	1 2	B1 B2	130	3 1	B3

2.3.2.3.Instrumentation

The NIR spectra were recorded in the sampling room using the FT-NIR MPA II (Bruker Optics, USA) offline spectrometer. The instrument specifications and their related measure parameters are described in Table 2. 29. These were selected according to prior experience and supplier recommendations.

Table 2. 29 NIRS instrument configuration for Sucramal product.

Parameter	Specification
Model	FT-NIR MPA II
Measurement module	Integration sphere with sample holder device with sample rotator
Optical path length (mm)	2
Diameter of the measured area (mm)	12
Dispersion principle	Cubic corner mirrors on the ROCKSOLID™ Interferometer
Detector type	Thermos-electrically cooled InGaAs
Measurement mode	Diffuse reflection
Number of scans	32
Spectral range (cm^{-1})	11550-3950
Resolution (cm^{-1})	8
Number of spectra	5 for laboratory samples/10 for MB

The instrument was coupled with an integration sphere measurement module adapted with a sample rotator for sample cups of 51 mm in diameter. All spectra samples were recorded through a quartz vial allowing measurements of 12.5 mm sample diameter.

The NIR method during routine analysis batches can be summarized as follows: after the homogenization step, the batch is sampled in the bin storage. Before spectral acquisition, the background spectrum corresponding to a dark background is acquired. Therefore, 25 mg approximately of the product is added to the integration sphere rotation accessory, and 5 spectra are recorded for each laboratory sample and 10 for each MB, considering the description for uniformity dosage analysis in the Ph. Eur method of analysis chapter for solid dosage forms[5] (and if considered[21]). The sample is homogenized manually with a stainless-steel spatula between each spectra acquisition to guarantee representativeness.

The HPLC technique is used as a reference. A specific Refractive Index Detector (from now on RID) is necessary. A specific column with a silica-based aminopropyl bonded sorbent as a stationary phase is used. The mobile phase consisted of an ammonium sulfate solution at pH 3.5.

The active substance content is calculated as shown in Equation 2 20 as an equivalence to sucrose octasulfate.

$$Active\ substance\ content\ \left(\frac{mg}{g}\right) = Ap\frac{PstdxRstd}{AstdxPp}x0.01$$

$$Assay\ (\%) = \frac{Active\ substance\ contentxPx100}{340}$$

Ap: chromatographic peak for sucrose octasulfate in the evaluated sample.

Pp: weight of the evaluated sample (g).

Pstd: weigh for the sucralfate standard (mg).

Rstd: assay content for the sucralfate standard (%).

Astd: chromatographic peak for the sucralfate standard.

P: average weight of the unit-dose sachets for the sample evaluated.

340: theoretical content (mg) of sucrose octasulfate in each sachet.

Equation 2 20.

2.3.2.4.Software

The OPUS (Buker Optics, USA) software was implicated during the development procedure and the routine analysis. The software was used for instrument cGMP verification, spectra acquisition, chemometric model development, and procedure validation.

2.3.3. Risk Assessment

The NIR spectra can be affected by many variables from sample, instrument, operation conditions, and manufacturing process. These can be minimized but not removed by mathematical pre-treatments. The models often should be recalibrated and validated to incorporate variables not expected during the initial model development. Table 2. 11 shows the severity ranking for associated risks to a NIRS model. Table 2. 30 summarize the risks detected and their control measure for the Sucramal product.

Table 2. 30 Risk to NIRS procedure for Sucramal product.

Variable	Risk	Range	Model Effect	Mitigation action
Starting material	Physic-chemical properties of the active substance	1	Non-expected. Only components from homologated manufacturers may be used	Do not apply
	Physic-chemical properties for excipients	1	Non-expected. Only components from homologated manufacturers may be used	Do not apply
		4	The initial physical characteristics of the excipients have an effect on the behavior of the final product "tapping", due to the difference in particle size	Spectra acquisition with a higher representative area. Integration sphere with sample rotator accessory

Table 2. 30 Risk to NIRS procedure for Sucramal product.

Variable	Risk	Range	Model Effect	Mitigation action
Laboratory samples	Not representative samples for laboratory strategy preparation	4	No representative model	Considering the guide of manufacturing to construct calibration and validation samples set
	Ambient conditions	2	Incorrect results	Including the variability with the production samples
Instrument	Lack of cleanliness of the sample interface	4	Cross-contamination involving incorrect results	Assuring the interface cleaning
	The optical quality of the sapphire window	4	Incorrect results	Good storage for sapphire window
	Replacement equipment consumables	3	Variations in the results	Comparability test. If it is necessary re-calibration and re-validation of the model
Production	Manufacturing process changes	3	No representative model	The manufacturing variability is going to be included in production samples

2.3.4. Development and Validation of NIR Spectroscopic Qualitative Analysis

Sub-cascading spectral library was used as an identification and qualification strategy.

The identification spectral library was structured in one general library, which includes all SM tested, and two sub-libraries at the second level. This cascading strategy was applied to allow for solving ambiguities. The detailed structure of the library is shown in Figure 2. 43. The IP1 code was referred to the first level, and was assigned to the general library. The IP2 code was referred to the second level sub-library and was assigned to the samples not able to identify without ambiguities in the first level.

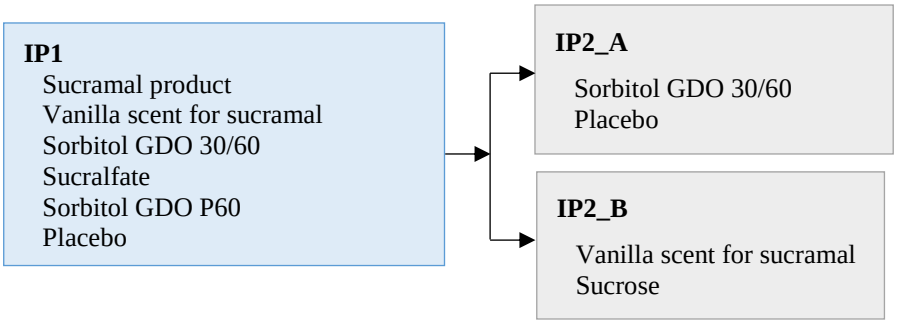


Figure 2. 43 Structure and content of the Sucramal product library.

The identification library contains all the materials intended to be used for the procedure application. This classification procedure aimed to allow the unequivocal identification of each sample included in the library, and to distinguish them from those which are not included.

During the calibration of the general library, distance as factorization in a PC space was used as a discriminant criterion. A wide spectral range was selected and derivative followed by vector normalization was used as data processing. An ambiguous identification for some of the materials included was obtained and automatic detected by the chemometric software. The sub-library cascading strategy was applied. There is no limit number, composition and levels of sub-libraries. The discriminant method applied for the sub-library was distance as factoring decomposition in a PC space. The spectral range and data pre-treatment were different. Narrow spectral ranges and a complex data processing were used. The global threshold was adjusted after internal validation allowing correcting unambiguous identification.

To understand the difference between included materials, Figure 2. 44 shows vector normalization spectra of a batch of Sucramal product and its formulation compounds. Sorbitol powder, as a major excipient, and placebo presented similar spectra profile. There was a significant difference in the wavelength where the active substance presented the main peak, between 7300 and 6800 cm^{-1} and between 5200 and 4800 cm^{-1}. It should be enough to differentiate between target product and its placebo.

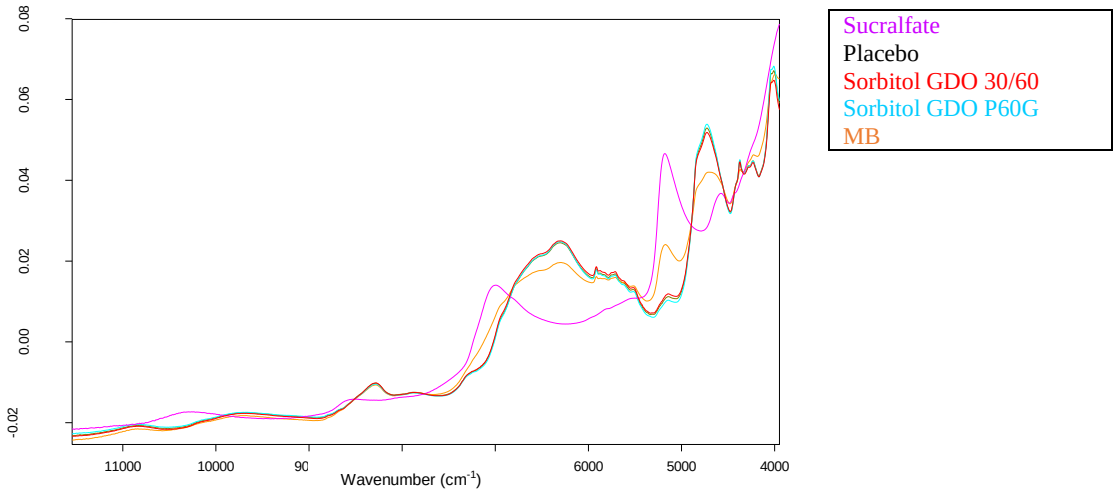


Figure 2. 44 Spectral data in diffuse reflectance for Sucramal formulation including placebo, major excipients, the active substance, and an MB. Vector normalization as data pre-treatment was applied in the complete spectral range.

The identification procedure for Sucramal product by NIR was validated following the specific guidelines[6,9,14] including specificity and robustness parameters.

The specificity was tested by internal validation for each library and sub-library considered the spectra including in the calibration set. Each spectrum was used for calibration and internal validation. As a result, the internal validation showed that all calibration spectra belong to each of assigned classes created (referred to each SM included), and there were no ambiguities during the calibration identification. All calibration spectra were used as positive and negative samples to confirm the ability to identify all classes and to discriminate within them.

The robustness was assessed by including a period timeline of one year being enough to include all possible relevant variables such as SM variation, operating and environmental manufacturing conditions modifications, and instrument changes.

2.3.5. Development and Validation of NIR Spectroscopic Quantitative Analysis

PLS has been the mathematical algorithm used for the assay determination of Sucramal product. The calibration set was used to calculate the model and the external validation set to optimize it.

The wavelength range used was from 9000 to 4928 cm^{-1}, where the correlation between NIR signal and active substance content changes was demonstrated (see Figure 2. 45). The excluded ranges were not provided relevant information to improve the predictive ability of the calibration model.

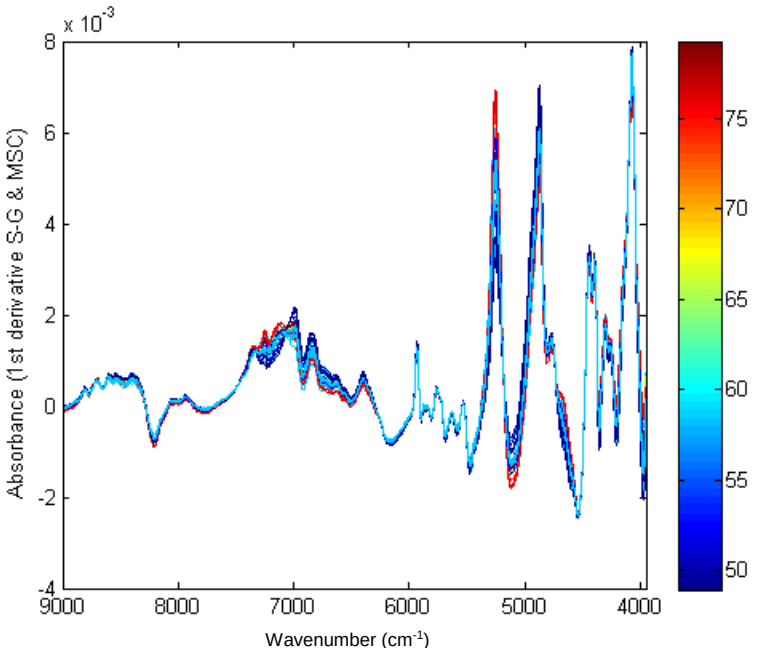


Figure 2. 45 Spectral data acquired in diffuse reflectance for the calibration data set. 1st derivative Savitzky-Golay with 13 points of window width followed by multiple scatter correction as data pre-treatment. Spectral range from 9000 to 4928 cm^{-1}.

The main bands of the active substance in the reflection range appeared at 7500-6500 cm^{-1} and 5500-4500 cm^{-1} (see Figure 2. 44).

The number of factors was optimized with the validation set to achieve the lowest RMSEP value. Five factors were finally selected. The factors included in the model were related to the principle active substance bands at 7500-6500 and 5500-5000 cm^{-1}. See Figure 2. 46.

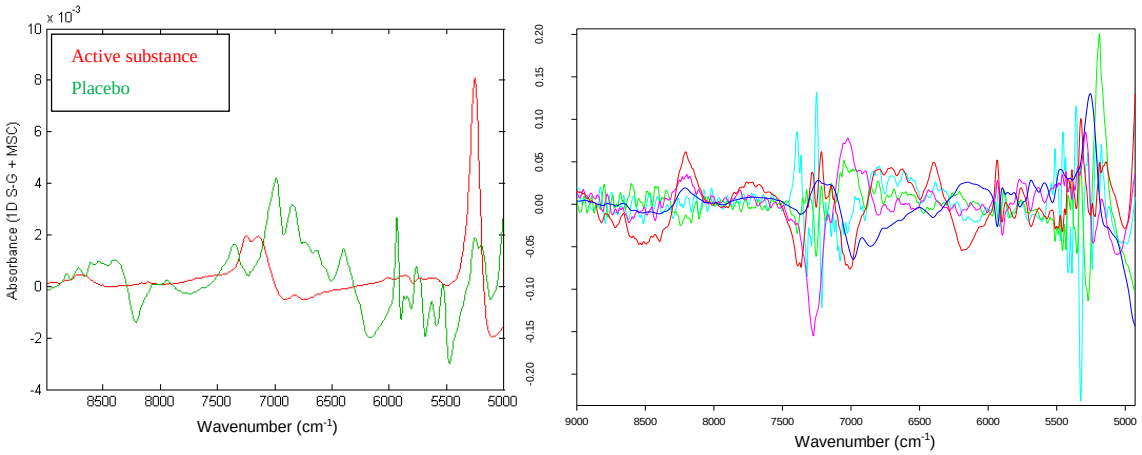


Figure 2. 46 Left) Pretreated reflectance diffuse spectra for the active substance and placebo. 1st derivative S-G with 13 window width followed by MSC. Right) Loading plot for the LV included in the PLS model. Spectral range included from 9000 to 4928 cm-1. LV1, LV2, LV3, LV4, LV5.

The scatter plot of the active substance content against the reference HPLC values, showed a linear relationship for calibration and validation sets (see Figure 2. 47).

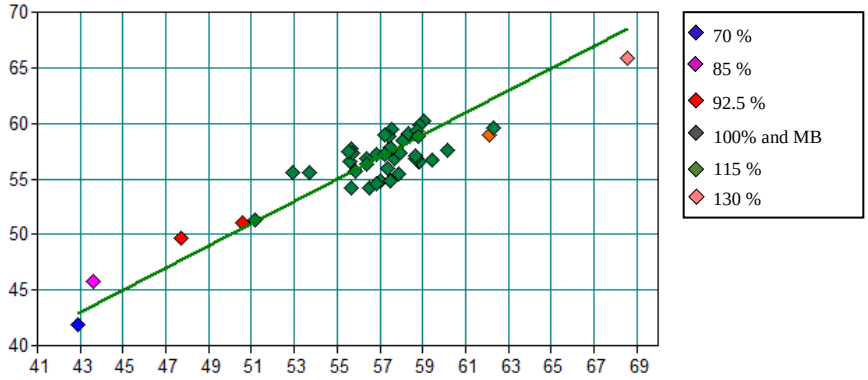


Figure 2. 47 Predicted vs. true values represented in a scatter plot for the validation set of Sucramal product after PLS modelling.

The quantitative procedure validation was designed according to the specific guidelines[6,9,14]. The figures of merit of the quantitative analysis validation for Sucramal product are presented in Table 2. 31.

Table 2. 31 Figures of merit for the validation quantitative analysis for Sucramal product by NIR spectroscopy.

Parameter	Results
Specificity	-The main peaks of active substance were presented in the loadings for the PLS modelling (see Figure 2. 46). -Final product library developed sub-cascading spectral library strategy analysis for the calibration set, allowing the unequivocally identification of each group of samples included in the library (see chapter 2.3.4). -MD evaluation. The limit was fixed by software in 0.40. All the spectra included for the validation set presented a lower MD than the critical one and were considered spectrally similar to the calibration set, so the predicted values obtained were reliable out samples up to the limit value (see Figure 2. 48).

Table 2. 31 Figures of merit for the validation quantitative analysis for Sucramal product by NIR spectroscopy.

Parameter	Results		
Linearity	Number of samples	17 (15 laboratory samples/2 MB)	
	Range (mg/g)	42.86-68.54	
	Explained variance (%)	94.35	
	y-intercept ± IC (α=0.05) including 0	7.07 ± [-0.033:14.18]	
	Slope ± IC (α=0.05) including 1	0.87 ± [0.75:1.00]	
	R>0.95	0.97	
	The residuals between the reference HPLC and NIRS methods did not show any trend regarding concentration levels (see Figure 2. 49 and Figure 2. 32).		
Accuracy	Number of samples	17 (15 laboratory samples/2 MB)	
	Paired *t*-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 0.19 *t*-tabulated: 2.11 *p*-value: 0.85	
	RMSEP (mg/g)	1.72	
Precision	SEP<1.4SEL	1.72<1.4x3.00	
	Validation bias value close to 0	0.3	
Repeatability	RSD (%) for 10 measures of the same batch less than 10%	4.44	
Intermediate Precision	Two-way ANOVA factor (day and analyst) *p*-value higher than 0.05 *F*-calculated lower than the *F*-tabulated	Analyst *F*-calculated: 1.64 *F*-tabulated: 18.51 *p*-value: 0.32	Day *F*-calculated: 0.11 *F*-tabulated: 19.00 *p*-value: 0.91
Robustness	Number of samples	76 MB	
	Paired *t*-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 1.94 *t*-tabulated: 1.99 *p*-value: 0.056	
Control Model Performance	Number of samples	65 MB	
	CL: $\bar{x}$ UAL/LAL, ±3σ UWL/LWL, ±2σ *Based on residuals* (see Figure 2. 50)	UAL: 5.04 UWL: 3.48 CL: 0.35 LWL: -2.78 LAL: -4.35	

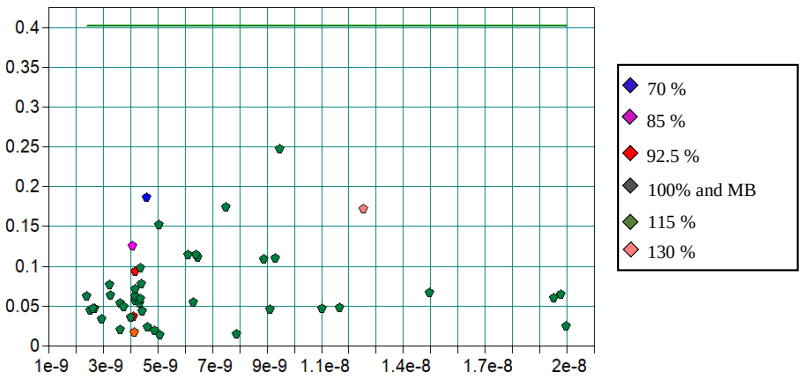


Figure 2. 48 Scatter plot for MD for each included sample into the validation set of Sucramal product.

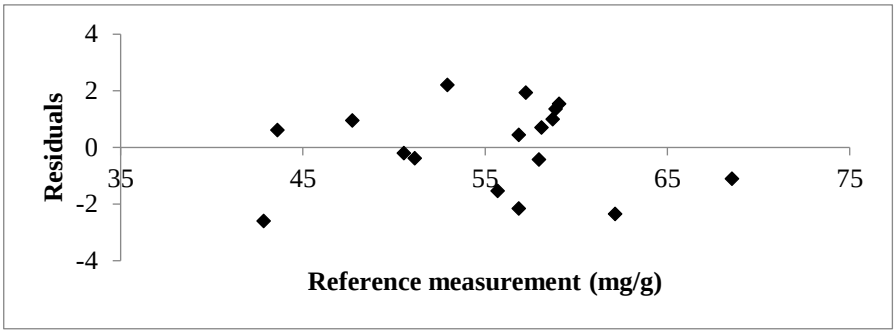
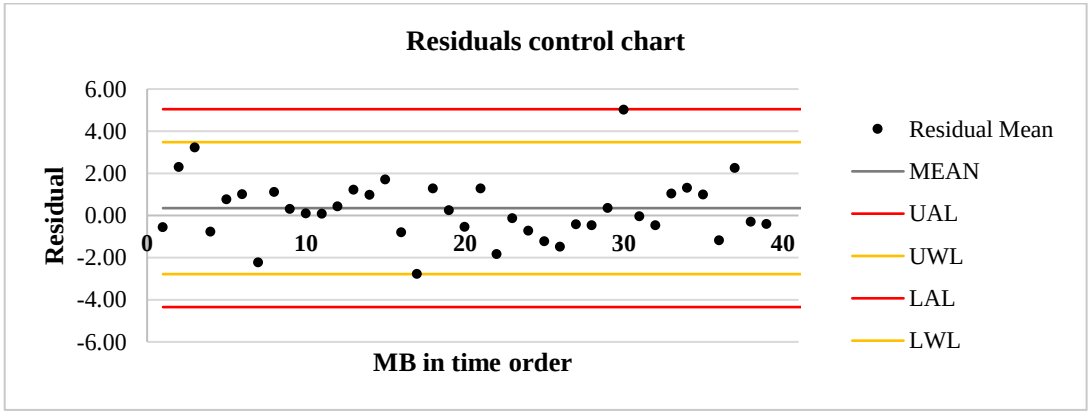


Figure 2. 49 Residuals plot for linearity validation of Sucramal NIR model.



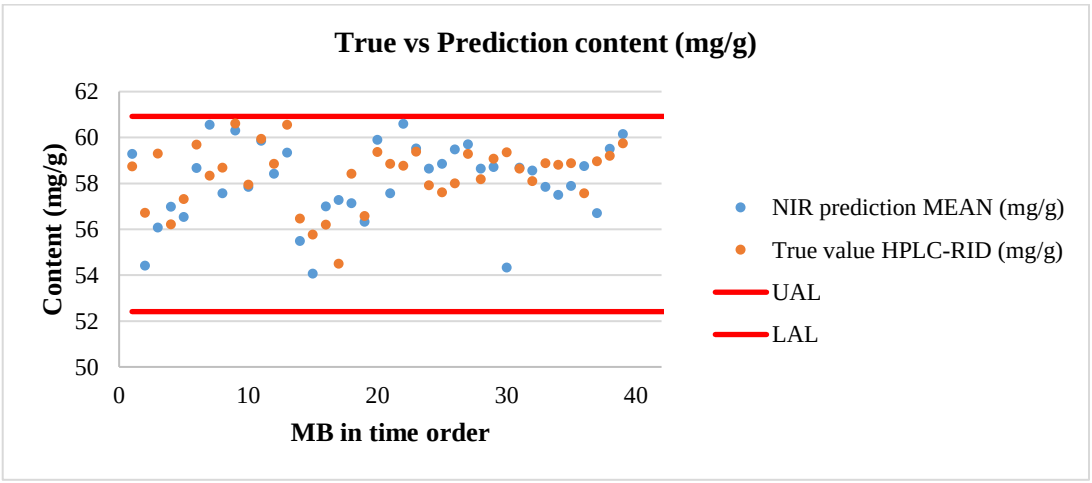


Figure 2. 50 Control chart for MB monitoring NIR model.

2.3.6. Conclusions

The results obtained show that the use of acquisition spectra in diffuse reflectance mode with the integrating sphere measurement module adapted with a sample rotator with the measurement channel macrosample, allows to achieve a robust NIRS procedure for qualifying and quantifying Sucramal product. The procedure was satisfactory validated in terms of specificity, linearity, accuracy, precision and robustness.

For some of the analyzed samples, the analysis could be inhomogeneous even though the entire sample is homogeneous since the penetration depth of NIR radiation was less than 1 mm and only a part of the surface sample was analyzed. This effect can cause large within-sample variability for this commodity resulting in a higher RSD in percentage value.

An effective and easier analysis for powder Sucramal product was presented as an alternative for the reference HPLC method. As a result, the specifications for the release of the manufactured batches were summited for variations in the commercialization product regions to involve the NIRS qualification and quantification analysis routinely instead of HPLC. To evaluate the applicability of the NIR analysis during the product life cycle, it was proposed to carry out the CQAs determinations with both techniques with a periodicity of 1 for 10 MB.

2.4. Bibliography

(1) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 30, 2015, pp 1–16.

(2) European Medicines Agency. Chapter 1: Pharmaceutical Quality System. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, December 14, 2012, pp 1–7.

(3) Blanco, M.; Bautista, M.; Alcalà, M. API Determination by NIR Spectroscopy Across Pharmaceutical Production Process. *AAPS PharmSciTech* **2008**, *9* (4), 1130–1135. https://doi.org/10.1208/s12249-008-9156-3.

(4) Blanco, M.; Villarroya, I. *NIR Spectroscopy: A Rapid-Response Analytical Tool*; 2002; Vol. 21, pp 240–250.

(5) European Pharmacopoeia. 2.9.40. Uniformity of Dosage Units. In *European Pharmacopoeia 11th edition*; 2023; Vol. 0, pp 421–423.

(6) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. London, UK, January 20, 2012.

(7) Blanco, M.; Peguero, A. Influence of Physical Factors on the Accuracy of Calibration Models for NIR Spectroscopy. *J Pharm Biomed Anal* **2010**, *52* (1), 59–65. https://doi.org/10.1016/j.jpba.2009.12.009.

(8) Blanco, M.; Romero, M. A.; Alcalà, M. Strategies for Constructing the Calibration Set for a Near Infrared Spectroscopic Quantitation Method. *Talanta* **2004**, *64* (3), 597–602. https://doi.org/10.1016/j.talanta.2004.03.027.

(9) European Medicines Agency. ICH Guideline Q2(R2) on Validation of Analytical Procedures. *EMA/CHMP/ICH/82072/2006*. London, UK, March 31, 2022.

(10) Biagi, D.; Nencioni, P.; Valleri, M.; Calamassi, N.; Mura, P. Development of a Near Infrared Spectroscopy Method for the In-Line Quantitative Bilastine Drug Determination During Pharmaceutical Powders Blending. *J Pharm Biomed Anal* **2021**, *204*. https://doi.org/10.1016/j.jpba.2021.114277.

(11) Bosch, M.; Mannucci, S.; Torras, E.; Falorni, R.; Gonzalez, J. Ma. Polymorphic Forms of Dexketoprofen Trometamol, Preparation and Pharmaceutical Compositions Thereof. EP1739072A1, January 3, 2007.

(12) Blanco, M.; Alcalá, M.; González, J. M.; Torras, E. Near Infrared Spectroscopy in the Study of Polymorphic Transformations. *Anal Chim Acta* **2006**, *567* (2), 262–268. https://doi.org/10.1016/j.aca.2006.03.036.

(13) Pandiscia, L. *Analysis of Crystal Polymorphism by Raman Spectroscopy for Medicine Development*; USA, 2022. http://www.jascoinc.com/applications.

(14) Food and Drug Administration. Analytical Procedures and Methods Validation for Drugs and Biologics Guidance for Industry. Silver Spring, MD, July 2015, pp 1–18.

(15) European Medicines Agency. ICH Q2 (R1). Validation of Analytical Procedures: Text and Methodology. *CPMP/ICH/381/95*. London, UK, June 1995.

(16) Elliot, C. R.; Newns, G. R. Near Infrared Absorption Spectra of Silica: OH Overtones. *Appl Spectrosc* **1971**, *25* (3), 378–379. https://doi.org/10.1366/000370271779949294.

(17) Chalus, P.; Roggo, Y.; Walter, S.; Ulmschneider, M. Near-Infrared Determination of Active Substance Content in Intact Low-Dosage Tablets. *Talanta* **2005**, *66* (5), 1294–1302. https://doi.org/10.1016/j.talanta.2005.01.051.

(18) Cárdenas Espitia, V.; Blanco, M.; Manel, R.; Bernàrdez, A. Development of New Analytical Methods Using Near Infrared and Chemometrics for Pharmaceutical Quality Control: Enhancement of Modelling Strategies Towards a Better Product Understanding, Universitat Autònoma de Barcelona, Bellaterra, 2015.

(19) Jian, F. A Review of Distribution and Segregation Mechanisms of Dockage and Foreign Materials in On-Farm Grain Silos for Central Spout Loading. *KONA Powder and Particle Journal* **2022**, *39*, 100–109. https://doi.org/10.14356/kona.2022014.

(20) He, X.; Han, X.; Ladyzhynsky, N.; Deanne, R. Assessing Powder Segregation Potential by Near Infrared (NIR) Spectroscopy and Correlating Segregation Tendency to Tabletting Performance. *Powder Technol* **2013**, *236*, 85–99. https://doi.org/10.1016/j.powtec.2012.05.021.

(21) European Pharmacopoeia. 2.9.47. Uniformity of Dosage Units Using Large Sample Size. *Ph. Eur. 11th*. Council of Europe: Strasbourg, France, January 2023, pp 432–434.

Experimental Chapter 3. In-Process Control Involving PAT Approach

3. IMPLEMENTATION OF A PAT METHODOLOGY FOR IN-PROCESS CONTROL

As an innovative approach to improving efficiency and quality in drug manufacturing, PAT allows quality-based adjustments in pharmaceutical processes involving proactive work. This strategy makes it possible to transfer the functions performed in the QC laboratory to the manufacturing plant[1–3].

This project aims to detail the development, maintenance, and implementation of the NIR methodology for in-process control, at the end of the final blending step, and as a quality control tool before batch release for CQA determination in real-time[4]. The proposed PAT system would allow the release of the granulated bulk product to the next unitary operation, increasing productivity. With this strategy, timely measurements of the intermediate generated during the process could be acquired to become a representative measurement of the product studied. As a result, the set of determinations, including the assay for the content uniformity dosage, should be carried out on the finished product by the reference method to comply with the currently registered methododology[4]. To optimize the resources and the release time of the finished product, this project also proposes the development, validation, and implementation of an offline NIR procedure at the end of the manufacturing process, to replace the currently registered methodology on a routine basis. It is proposed as a control strategy, carrying out the analysis in parallel with the reference method with a periodicity of 1 in 10 batches, considering also the determination of the microbiological quality required of the batch Figure 3. 7 shows the Nimesulide product's manufacturing process with the NIR analysis's incorporation.

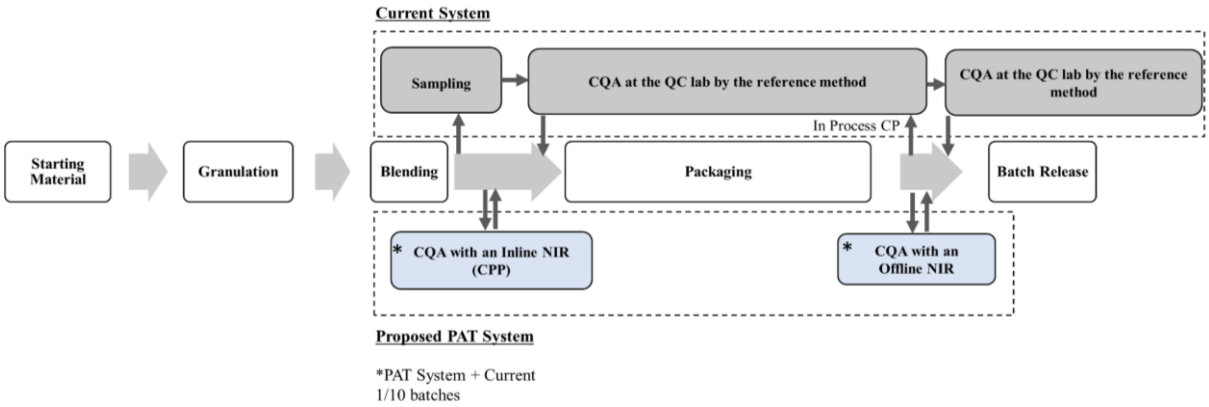


Figure 3. 7 Conventional and proposed systems for CQA determination of Nimesulide product.

In 2016 the structure for inline NIR analysis was finalized. The implementation of the PAT system was finished in 2018 establishing an information monitoring and distribution system as a key tool for recording the set of operations to be carried out during the mixing process and the subsequent analysis by NIR which is now tested[5].

3.1.Product Composition

The Nimesulide product is a solid granular product in unit-dose sachet form with an active substance concentration of approximately 5% (w/w). Contains sucrose as main excipient. The composition is summarized in Table 3. 2.

Table 3. 2 Nimesulide product composition.

Description	Content per unit-dose sachet % (w/w)
Nimesulide	5.00
Sucrose powder	81.05
Sucrose crystal	9.20
Orange essence	2.10
Citric acid	1.50
Maltodextrin	0.75
Cetostearyl macrogol	0.40

Two different formulas are manufactured depending on the active substance supplier.

A ligand solution is prepared with the cetostearyl macrogol, the sucrose crystal, and the maltodextrin. Then the granulation is assessed with the rest of the SM at a fluid bed. When the product is granulated, it is discharged into a stainless-steel tumbling bin, and the orange essence is added to be homogenized.

The specifications for the CQAs to be determined are shown in Table 3. 3.

Table 3. 3 CQAs specifications for Nimesulide product.

CQA	Reference Method	Specification
Identification	NIR and HPLC	Pass
Active substance content (mg/g)	HPLC	47.5-52.5
CV related to active substance content (%)	-	≤5.0
Loss on drying (%)	Moisture analyzer	≤1.0

3.2.Preliminary Studies

3.2.1. Materials and Methods

This section details the materials and methods involved in the development and validation of a qualitative and quantitative procedures for the inline NIR analysis for the Nimesulide granules.

3.2.2. Instrumentation

The NIR LabSpec Pro 2500 (ASD inc., USA) was used for the spectral acquisition. The instrument configuration is summarized in Table 3. 4.

Table 3. 4 Instrument configuration for the NIR LabSpec Pro 2500.

Parameter	Specification
Model	LabSpec Pro 2500 (ASD inc., USA)
Measurement accessory	A solid probe of 3 m, through a sapphire window (3mm thickness, 2 cm diameter)
Dispersion principle	-UV-vis-NIR region: diode array with filters (350-1000 nm) -NIR region: two monochromators with concave holographic grating (1000-1800 and 1800-2500 nm)
Detector type	-UV-vis-NIR: photodiode array -NIR region: high Sensitivity Thermo-Electric Cooler InGaAs
Acquisition mode	Diffuse reflectance
Number of scans	32
Spectral range (nm)	350-2500
Resolution (nm)	UV-vis-NIR region: 4; NIR region: 10
Number of spectra per batch	3

For NIR measurements it was necessary to modify the stainless-steel bin cup involving a sapphire window. The strategy for NIR measurements, during routine batch analysis, can be summarized as follows: before the blending step, the granulated product is discharged in a stainless-steel bin, afterward, the final blending step is performed. Once the blending is finished, the bin is positioned upside down allowing two key points: that the product completely covers the sapphire window and being able to couple the NIR probe to the bin. The NIR spectrum reference is acquired with a *spectralon* plate. Once the probe is coupled in a static mode to the bin, one spectrum is acquired. Once the process is finished, the probe is uncoupled and the bin is rotated 360º allowing the sample to be homogenized. This procedure is carried out in triplicate for each batch to be analyzed[5].

The instrument is daily qualified and maintained under a cGMP environment[6].

The reference values are acquired with the validated and authorized methods by LMSA for product analysis.

The HPLC technique is used as a reference for active substance content. The instrument configuration is summarized in Table 3. 5.

Table 3. 5 Instrument configuration for the HPLC as a reference technique for the active substance content analysis of Nimesulide granulate.

Parameter	Specification
Model	HPLC-DAD (Agilent Technology, USA)
Column	C18 4.6x50 mm
Detector type	DAD
Mobile phase	Acidified water and methanol (40:60 v/v)
Number of samples per batch	3
Software	Empower (Waters, USA)

The reference values for the active substance content involving HPLC method are calculated as shown in Equation 3. 3.

$$Active\ substance\ content\ \left(\frac{mg}{sachet}\right) = Ap\frac{PstdxRstd}{Astd}x10^{-2}$$

$$Assay\ (\%) = \frac{\frac{mg}{sachet}x100}{100}$$

Ap: Chromatographic peak for Nimesulide in the problem solution.

Pstd: standard Nimesulide weight (mg).

Rstd: assay for the standard of Nimesulide (%).

Astd: Chromatographic peak for Nimesulide in the standard solution.

100: Theoric content for Nimesulide in the sachets (mg).

Equation 3. 3.

The loss-on-drying technique is used as a reference for the analysis of moisture content with moisture equipment from Mettler Toledo (USA) analyzing a representative sample of the batch at 90ºC for 5 min. The result is given automatically by the equipment.

3.2.3. Software

The MB spectra were acquired involving the real-time process monitoring software Unscrambler X Process Pulse 1.1M (CAMO Analytics AS, Norway). Moreover, this software was involved in CQAs prediction. For model maintenance the Unscrambler X 10.3. (CAMO Analytics AS, Norway) was used. This software is a modified version of LMSA that allows communication between the Unscrambler and the internal data management program in the PLC. The installation and operation qualification (IQ and OQ)[6–8] of this system were carried out.

3.3. Life Cycle Management

Once the NIR procedure is validated, it must be maintained during the product's life cycle, considering its sensitivity to possible changes in parameters related to the SM, the manufacturing process, and/or the instrument[9]. If a change in trend or a deficiency in the prediction of the NIR model is detected, by the obtention of an OOS value confirmed by the reference method, it must be recalculated.

The trend was evaluated by control charts based on Pearson's residuals. When any of the patterns defined was detected, according to Nelson's rules, an evaluation is carried out to determine the need to update the said models: 1 value outside the upper or lower rejection limit, 9 consecutive points above or below the centre line, or 6 consecutive points increasing or decreasing. This possible trend obtained was corroborated by an internal diagnosis based on the model used, which includes the influence of Hotelling T^2 and the Q residuals established with the included calibration set. Higher unusual Hotelling T^2 and/or higher Q residuals limited at 95% confidence result in a decrease in accuracy for the values predicted by NIR[10,11].

In this section, the maintenance of quantitative procedures for online NIR analysis for CQAs of Nimesulide granules is detailed.

3.3.1. Maintenance of the Quantification Procedure for the Determination of the Moisture Content

After the procedure validation, continuous parallel testing, including the reference and the NIR method, was carried out to evaluate the life cycle of the chemometric model. The residuals obtained, defined as the difference between the reference moisture and the predicted water content, were plotted in a control chart, including the control limits calculated during the initial PLS model validation (LAL, UAL, LWL, UWL, CL, respectively)[5]. As can be seen in Figure 3. 8, from point 27 to the end there is shown a negative trend, more than 9 points in the same part of the control plot, considering Nelson rules, moreover, OOS values were acquired with the NIR model (red colour). The variation in the behaviour of the sample during the manufacturing process is observed as a root cause.

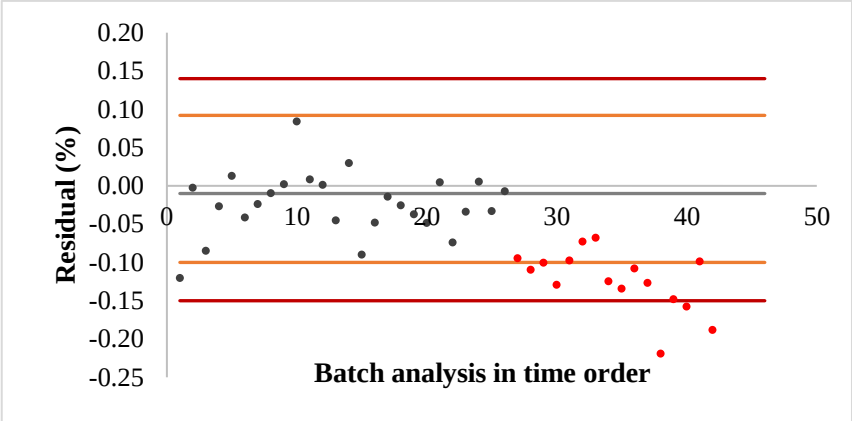


Figure 3. 8 Control plot of residuals (%) obtained for Nimesulide moisture content. Red line: UWL, LWL; orange line: UAL, LAL; grey line: mean for residuals.

The NIR model was updated to determine as representative as possible the moisture content of the Nimesulide granules.

Production samples were included in the calibration set besides laboratory samples, to introduce the new expected process variability.

The figures of merit of the calibrated PLS model for the moisture content are presented in Table 3. 6. The optimization of the selected parameters was based on de cross-validation results.

Table 3. 6 Figures of merit for the calibration model for the maintenance of the NIR model for moisture content of Nimesulide granules.

Set	Characteristics		Value
Calibration set	**Number of samples**		98
	Spectral pre-treatment		1st Derivative Savitzky-Golay (polynomial order 2, window 17 points)
	Wavelength range (nm)		1120-1968
	Calibration range (%)		0.12-1.67
	Internal/External specification		≤0.5/≤1.0
	Number of PLS factors		3
	Explained variance (Y) (%)		92.8
	Regression Y_{ref} vs. Y_{NIR}	**Slope**	0.928
		Offset	0.03
	RMSEC/RMSECV (mg/g)		0.08/0.08

After S-G's first derivative in the second polynomial order with 17 window widths the samples were ordered through the PC1 based on the active substance content into the scores plot (see Figure 3. 9).

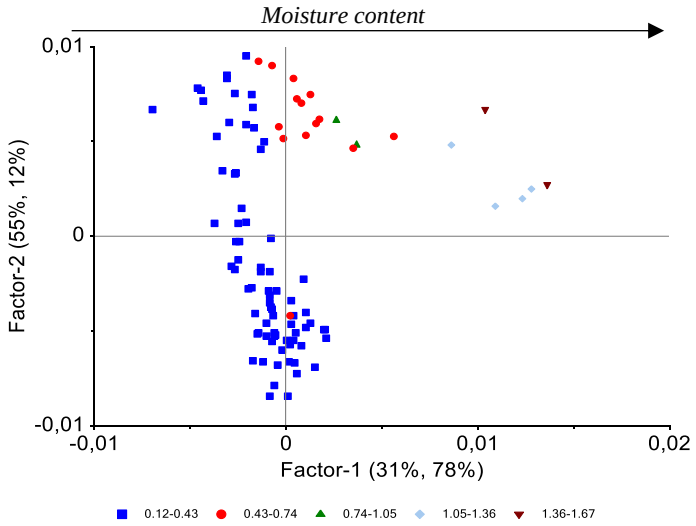


Figure 3. 9 Scores plot PC1 vs. PC2 for the calibration set. 86% of the total variability was explained. first derivative S-G in the second polynomial order with 17 window widths was applied as data pre-treatment.

The figures of merit for the validation of the analytical NIR procedure for the determination of the moisture content in the Nimesulide granules are detailed in Table 3. 7.

Table 3. 7 Summary of the results obtained in the validation of the analytical procedure for the determination of moisture content in Nimesulide granules.

Parameter	**Acceptance criteria**	**Result**	**Conclusion**
System suitability	Pass the verification test	Complies	The system is suitable
Specificity	Correlation coefficient >0.998	Complies	The NIR procedure is specific
Linearity	Number of samples	13	The NIR procedure is linear
	Validation range (mg/g)	0.16-1.23	
	Slope ± CI(α=0,05) include 1	1.01 ± [0.94;1.09]	
	Intercept ± CI(α=0.05) include 0	0.01 ± [-0.031;0.052]	
	r for calibration test set and validation test >0.95	0.99	
	RMSEP (%)	0.05	
	Residual values are random. The mean is close to 0	Complies	
Accuracy	Number of samples	13	The NIR procedure is accurate
	Paired *t*-test: *t*-calculated<*t*-tabulated	*t*-calculated: 1.68 *t*-tabulated: 2.18	
Precision. Repeatability	RSD of 6 measures ≤2%	0.27	The NIR procedure is precise

Table 3. 7 Summary of the results obtained in the validation of the analytical procedure for the determination of moisture content in Nimesulide granules.

Parameter	Acceptance criteria	Result		Conclusion
Precision. Intermediate Precision	No differences between days and analysts for 6 measures. *F*-calculated<*F*-tabulated *p*-value>0.05	Day $F_{calulated.}$: 0.55 $F_{tabulated}$: 19.00 *p*-value:0.64	Analyst $F_{calculated}$: 4.49 $F_{tabulated}$: 18.51 *p*-value:0.17	The NIR procedure is precise
Robustness	Number of samples	317		The NIR procedure is robust
	Paired *t*-test: *t*-calculated<*t*-tabulated	*t*-calculated: 0.73 *t*-tabulated: 1.97		
Control Model Performance	UAL: $\bar{x}$+3σ UWL: $\bar{x}$+2σ CL: $\bar{x}$ LWL: $\bar{x}$-2σ LAL: $\bar{x}$-3σ	0.16 0.10 -0.002 -0.11 -0.16		The NIR procedure is under control

3.3.2. Maintenance of the Quantification Procedure for the Determination of the Active Substance Content

After the procedure validation, continuous parallel testing for the active substance content, including the reference and the NIR method were achieved to evaluate the life cycle of the chemometric model. The residuals obtained were plotted in a control chart, including the control limits calculated during the PLS model validation[5]. As can be seen, at the end of the graph a positive trend is shown, more than 9 points in the same part of the control plot, considering Nelson's rules. The variation in the behaviour of the sample during the manufacturing process is observed as a root cause. The NIR model resulted in lower content values than the reference HPLC model. See Figure 3. 10.

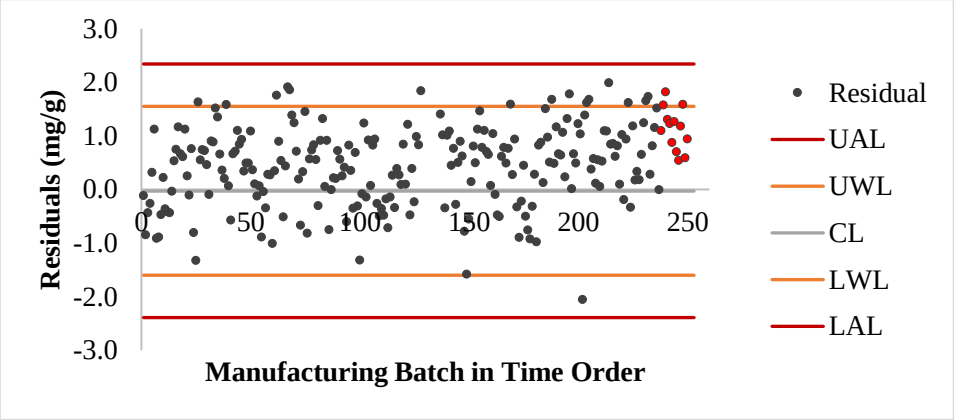


Figure 3. 10 Residuals plot for the MB of Nimesulide granules in time order. Positive trend marked with red samples.

The NIR model was updated to determine as representative as possible the CQA assay of the Nimesulide granules.

After the first derivative S-G in the second polynomial order with 19 window widths followed by area normalization, the samples were ordered through the PC1 based on the active substance content in the scores plot (see Figure 3. 11).

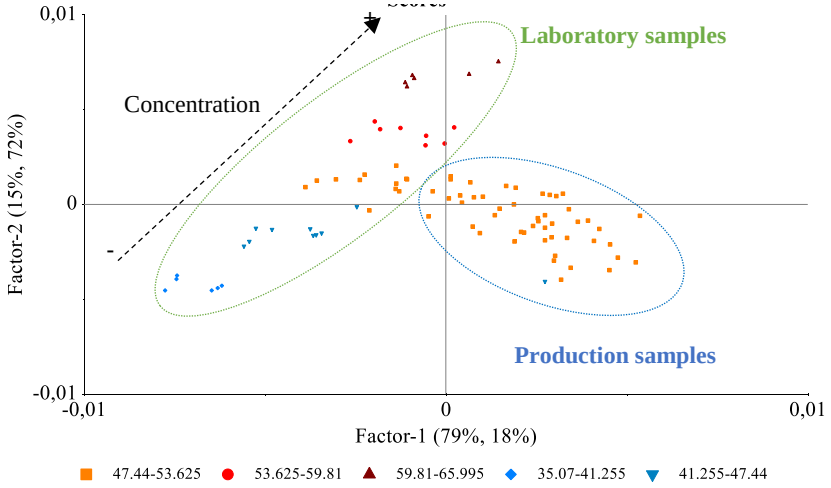


Figure 3. 11 Scores plot PC1 vs. PC2 for the validation set. 94% of the total variability was explained. First derivative S-G in the second polynomial order with 19 window widths followed by area normalization was applied as data pre-treatment.

The figures of merit of the calibrated PLS model for the active substance content are presented in Table 3. 8. The optimization of the selected parameters was based on de cross-validation results.

Table 3. 8 Figures of merit for the calibration model for the maintenance of the NIR model for active substance content of Nimesulide granules.

Set	Characteristics		Value
Calibration set	**Number of included samples**		95
	Spectral pre-treatment		1st Derivative S-G (polynomial order 2, window 19 points) followed by area normalization
	Wavelength range (nm)		1007-2313
	Calibration range (mg/g)		35.07-66.00
	Nominal value (mg/g)		50.0±2.5
	Number of PLS factors		5
	Explained variance (Y) (%)		97.60
	Regression Y_{ref} vs. Y_{NIR}	**Slope**	0.976
		Offset	1.19
	RMSEC/RMSECV (mg/g)		0.94/1.10

The figures of merit for the validation of the analytical NIR procedure are detailed in Table 3. 9.

Table 3. 9 Summary of the results obtained in the validation of the analytical NIR procedure for the determination of the active substance content in Nimesulide granules.

Parameter	**Acceptance criteria**	**Result**		**Conclusion**
System suitability	Pass the verification test	Complies		The system is suitable
Specificity	Correlation coefficient >0.998	Complies		The NIR procedure is specific
Linearity	Number of samples	16		The NIR procedure is linear
	Validation range (mg/g)	35.32-63.96		
	Slope ± CI(α=0,05) include 1	0.955 ± [0.906;1.003]		
	Intercept ± CI(α=0.05) include 0	1.92 ± [-0.55;4.42]		
	r for calibration test set and validation test >0.95	0.99		
	RMSEP (mg/g)	0.85		
	Residual values are random. The mean is close to 0.	Complies		
Accuracy	Number of samples	16		The NIR procedure is accurate
	Paired *t*-test: *t*-calculated<*t*-tabulated	*t*-calculated: 1.74 *t*-tabulated: 2.13		
Precision. Repeatability	RSD of 6 measures ≤2%	0.65		The NIR procedure is precise
Precision. Intermediate Precision	No differences between days and analysts for 6 measures. *F*-calculated<*F*-tabulated *p*-value>0.05	Day $F_{calculated.}$: 0.059 $F_{tabulated}$: 19.00 *p*-value:0.95	Analyst $F_{calculated}$: 2.80 $F_{tabulated}$: 18.51 *p*-value:0.23	
Range	Range≥10SEP	30≥10*0.85		The NIR procedure presents a valid range
Robustness	Number of samples	371		The NIR procedure is robust
	Paired *t*-test: *t*-calculated<*t*-tabulated *p*-value>0.05	*t*-calculated: 1.14 *t*-tabulated: 1.97 *p*-value: 0.26		
Control Model Performance	UAL: $\bar{x}$+3σ UWL: $\bar{x}$+2σ CL $\bar{x}$ LWL: $\bar{x}$-2σ LAL: $\bar{x}$-3σ *Based on the calculated residuals*	2.41 1.62 0.04 -1.53 -2.32		The NIR procedure is under control

3.3.3. Historical Data Evaluation

The set of historical data acquired, up to the date of writing this chapter, for the determinations of the CQAs of the Nimesulide granules, including identification, active substance content, and moisture content by the inline NIRS, during the mixing unit orperation, as well as by the currently authorized method, are reported.

The objective was to evaluate the set of results acquired by NIRS based on control charts to implement this technique during the routine control of the quality of the bulk product.

The comparison of the predicted values with the current NIRS methods and the reference values was carried out using control charts of Pearson's residuals.

The data presented covers the period from December 2018 until March 2023.

Identification

The set of data for the identification parameter is shown in Figure 3. 12.

The internally established lower rejection limit for product identification, based on the correlation coefficient was <0.998.

During the included data period, only one update to the identification chemometric model needed to be developed in April 2020 to include new product-associated variability, since a MB was identified with a correlation coefficient lower than the rejection limit. Up to this date, all the identification results have been satisfactory.

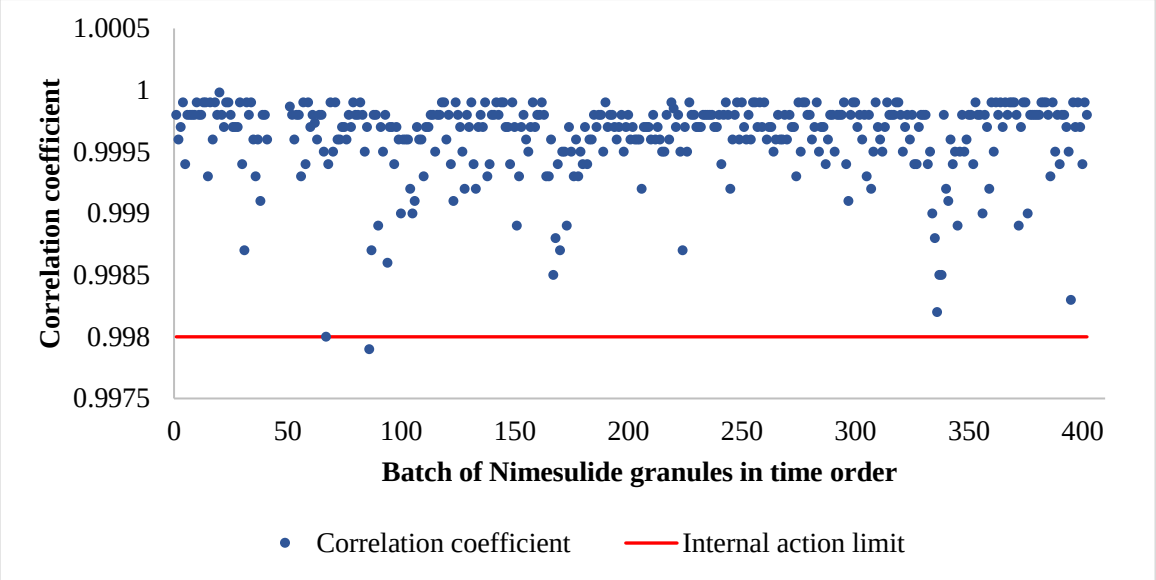


Figure 3. 12 Set of determinations for the identification of the Nimesulide granules prior to dosing unit operation by NIRS.

Moisture content

The control chart based on Pearson's residuals for the set of batches determined by the NIRS and the loss on drying reference methods, evaluating the water content (%), is shown in Figure 3. 13. The different NIR model updates can be seen as a result of the recalculation centre line and action and warning limits. The last update of the model was made in June 2022 due to the obtaining of OOT, according to the Nelson rules, in the residual control table. The addition of product variability was required.

The comparison for both models including the internal water content specification (≤0.5 %) is shown in Figure 3. 14. The authorized specification is ≤1.0%.

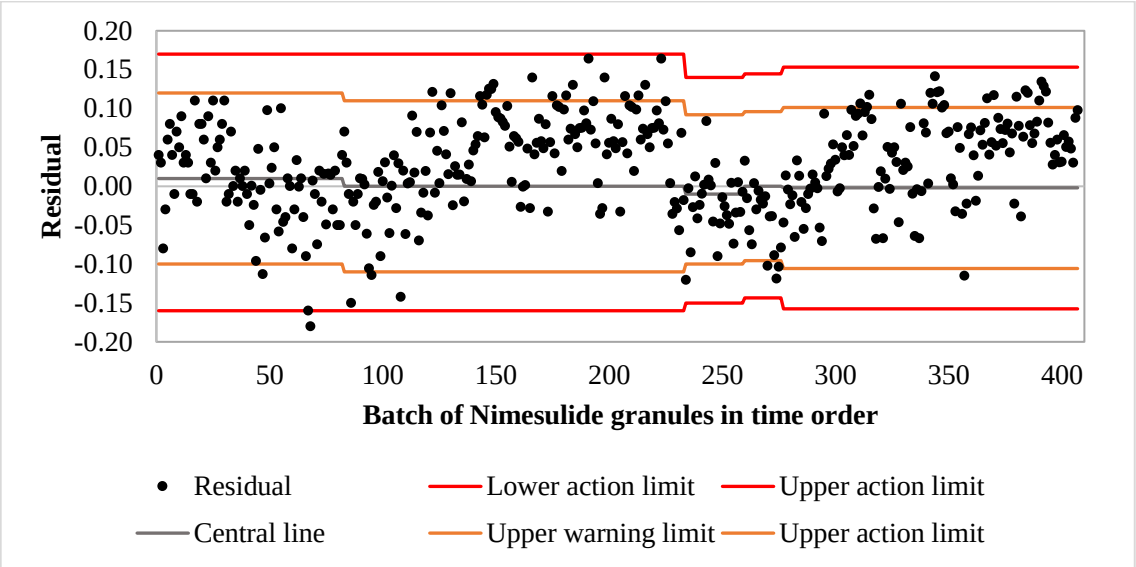


Figure 3. 13 Residual plot for the moisture determination of Nimesulide granules.

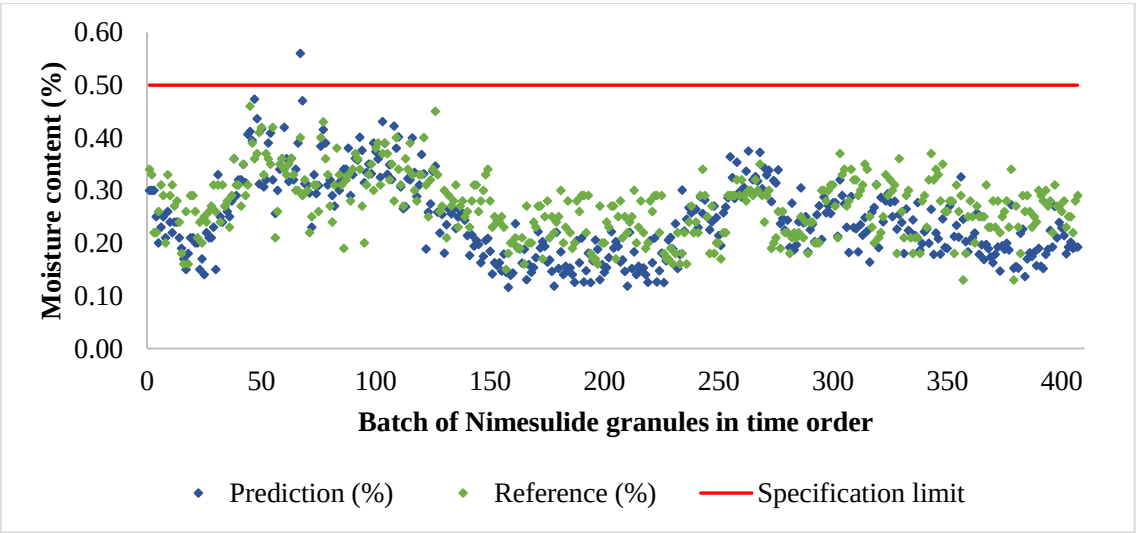


Figure 3. 14 Set of determinations for the moisture content (%) by NIRS and the loss on drying reference methods.

Active substance content

The set of batches analyzed by NIRS and the HPLC reference methods, evaluating the active substance content (mg/g) are included. A control chart based on Pearson's residuals is shown in Figure 3. 15 (calculated by the difference between the value obtained by the reference method and the predicted NIRS value). The different NIR model updates can be identified as a result of the different calculated centre line, warning, and action limits (based on $\bar{X}\pm2\sigma$ and $\bar{X}\pm3\sigma$, respectively). The last update of the active substance content model was made on January 2023 due to obtaining an OOT, according to Nelson's rules, in the residual control chart. Up to this date, all the determinations have been satisfactory.

The comparison of the active substance content acquired with both methods is shown on Figure 3. 16, using the currently registered rejection limits (50.0±5.0 mg/g), moreover, an internal warning limit is shown (50.0±2.5 mg/g).

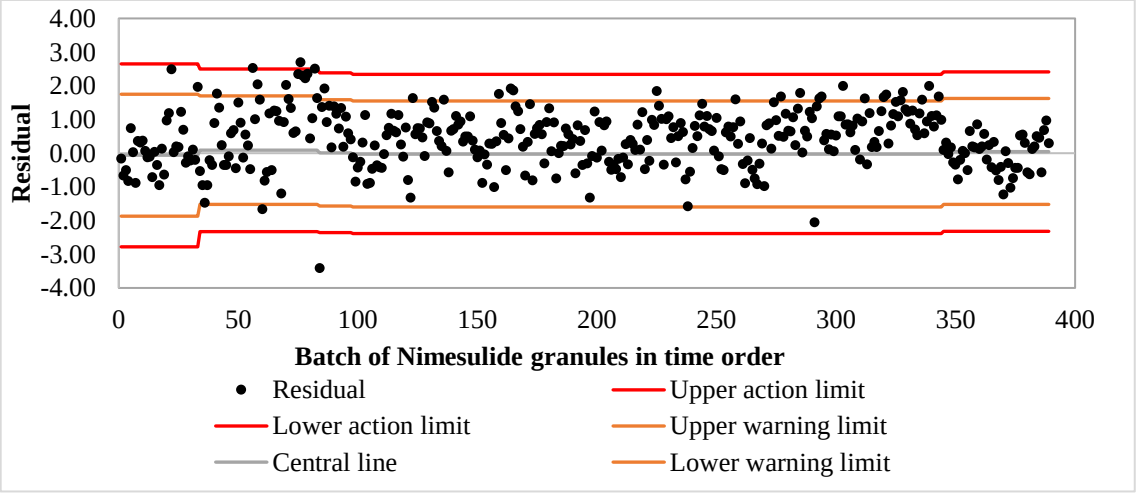


Figure 3. 15 Residual plot for the active substance content determination of Nimesulide granules.

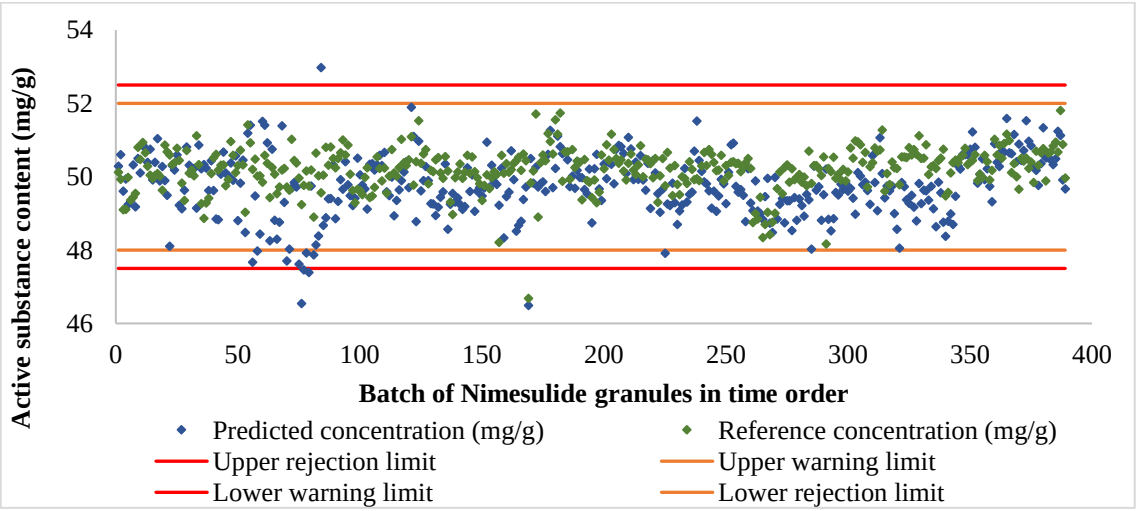


Figure 3. 16 Set of determinations for the active substance content (mg/g) by NIRS and the reference methods.

It was demonstrated the feasibility of the NIRS models developed and validated for the determination of the CQAs under study for the Nimesulide granulate for unit-dose sachets before the dosage operation. Comparable results were obtained considering the reference methods currently authorized for the release of each of the manufactured batches based on the control chart shown above.

3.4.CQAs Determination Involving an Offline NIR Method for Batch Release

3.4.1. Materials and Methods

In this section, the materials and methods involved in the development and validation of the qualitative and quantitative procedures involving an offline NIR analysis for the Nimesulide product are detailed.

3.4.1.1.Calibration and Validation Sets for Qualitative Analysis

The samples used for the development and validation of the qualification analysis by NIRS for the Nimesulide product are summarized in Table 3. 10. All production batches were properly prior identified by the reference method. The active substance, the excipients, and the placebos, which were prepared with all SM except the active substance, were used to test the specificity of the procedure. Moreover, a pharmaceutical product that involved the same mixture of equipment was included as a negative sample. The reasons for the choice of each product are described in the presented table (see Table 3. 10). The samples used in the calibration set should include all possible sources of variability to ensure that the calibration covers all ranges of interest. A set of MB was carefully selected to include the variability source associated with the product and the manufacturing process.

Table 3. 10 Samples involved to calibrate, internally, and externally validating the qualification method for the Nimesulide product.

Product	Sample type	Parameter	Justification
Nimesulide product	Calibration, internal and external validation sets	Specificity and Robustness	Chemical and physical variability
Nimesulide	Calibration, internal and external validation sets	Specificity	Active substance
Placebo	Calibration, internal and external validation sets	Specificity	Without active substance content
Sucrose	Calibration, internal and external validation sets	Specificity	Main excipient
Orange essence	Calibration, internal and external validation sets	Specificity	Excipient
Citric acid	Calibration, internal and external validation sets	Specificity	Excipient
Maltodextrin	Calibration, internal and external validation sets	Specificity	Excipient
Cetostearyl macrogol	Calibration, internal and external validation sets	Specificity	Excipient
Dexketoprofen Trometamol 25 mg product	Negative sample	Specificity	Materials handled close to the process

3.4.1.2.Calibration and Validation Sets for Quantitative Analysis

The calibration set has to include the variability sources that potentially affect the NIR spectra. The controllable factors were the active substance content, expanded ±30% regarding the nominal content, as well as the active substance manufacturer. Other uncontrollable factors could affect the spectra, such as the water content and particle size. To take them into account, as an expected variability, production samples used for the calibration set were accurately selected.

The strategy followed to expand the narrow range for the active substance content determination was the next one: production samples after the granulation and final blending step were under and overdosed with powder placebo and active substance respectively, to expand the concentration from 70 to 130% relative to the nominal value. Seven concentrations were included: 70, 85, 100, 115, 130, as well as 95 and 105%, assigned as the acceptance intervals for the active content. 100 g was considered as the prepared amount of samples. Using mixture design, three placebos were prepared to expand the excipient variability from the final Nimesulide granulated mixture to around 5%. The strategy followed was to modify around 5% of the composition for those considered majority (nominal value higher than 5%), from the nominal value of placebo composition; sucrose and a mixture from citric acid, macrogol, maltodextrin, and orange essence, reducing as much as possible correlations between formulation components (information on placebos preparation is summarised in Table 3. 11.

Table 3. 11 Content in (%, w/w) for the three prepared placebos, expanded the excipient variability around 5 percent from the nominal value from those considered majority.

Sucrose powder	Sucrose crystal	Citric Acid	Cetostearyl macrogol	Maltodextrin	Orange essence	Placebo Code
81.1	14.2	1.50	0.40	0.75	2.10	**1**
89.6	5.17	1.66	0.44	0.83	2.32	**2**
85.3	9.69	1.58	0.42	0.79	2.21	**3**

The calibration set included thirty-six laboratory samples and three runs at the nominal value for each MB. Three different MB were included from one homologated active substance manufacturer due to the arrival frequency. The prepared placebos were assigned randomly, expanding the placebo variability to around 15% for sucrose crystal and 5% for the rest of the excipients for all the concentration levels.

MB were carefully selected, including, where possible, different batches of SM, different expected physical variability, and granulation by different operators. The structure of the prepared calibration samples for each included MB is summarized in Table 3. 12.

Table 3. 12 Composition of the calibration and internal validation set for each MB of Nimesulide product. The active substance was expanded from 70 to 130 % (%, w/w).

Use	Batch number	Concentration level (%)	MB (g)	API (g)	Placebo (g)
Calibration and Internal validation	MB	70	70	0	30
			70	0	30
		85	85	0	15
			85	0	15
		95	95	0	5
			95	0	5
		MB-100	100	0	0
		105	99.7368	0.2632	0
			99.7368	0.2632	0
		115	99.2100	0.7890	0
			99.2100	0.7890	0
		130	98.4200	1.5790	0
			98.4200	1.5790	0

The same strategy was followed for the external validation set preparation. The batches used for the placebo preparation, and the active substance included were completely independent of the ones involved in the calibration set preparation. Moreover, the three placebos in the validation set were prepared on different days[12]. The validation set included twelve laboratory-prepared samples and one

run at the nominal value for the MB selected, completely independent from the ones involved in the calibration set.

The amounts for placebos were identical to the ones described in Table 3. 11. Also, the amounts involved at each concentration level were identical to the ones described in Table 3. 12.

3.4.1.3.Instrumentation

The NIR instrumentation parameters for the FT-NIR MPA II instrument are summarized in Table 3. 13. The conditions were optimized on a suitability study where different spectral parameters were tested to achieve the best spectral response.

Table 3. 13 Instrument configuration for the FT-NIR MPA II for Nimesulide product.

Parameter	Specification
Model	FT-NIR MPA II (Bruker)
Measurement module	Individual sample holder with vial using an integrating sphere (22 mm diameter) and the measuring sample wheel of 30 positions
Dispersion principle	Mirrors with cube corners in the interferometer ROCKSOLIDTM
Detector type	High Sensitivity Thermo-Electric Cooler InGaAs
Acquisition mode	Diffuse reflectance
Number of scans	32
Diameter of the area measured with the measuring accessory Integrating sphere	Integrating sphere Macro sample: 12 mm diameter measured area
Spectral range (cm^{-1})	11550-3950
Resolution (cm^{-1})	16
Number of spectra per batch	10

The NIR method during routine analysis batches can be summarized as follows: after the packaging step, the product was sampled[4]. Before the spectral acquisition, the background spectrum corresponding to the dark background was acquired. Therefore, the content of each sachet was added independently at different certified glass vials of measurement (22 mm diameter), and 10 spectra, one for each sachet, considering the description for uniformity dosage analysis in the Ph. Eur chapter for solid dosage forms[13], were recorded automatically with the integrating sphere macrosample measurement module adapted with a sample wheel of 30 positions. The samples were homogenized manually before the acquisition to guarantee spectra representativeness.

The instrument was daily qualified and maintained under a cGMP environment[6].

The HPLC technique was used as a reference (see chapter 3.2.2).

3.4.1.4.Software

Different software were involved during the offline NIR procedure development and validation:

- OPUS from Bruker Optics: instrument daily verification. Spectra acquisition for laboratory and manufacturing samples. OPUS IDENT and QUANT package for qualitative and quantitative chemometric model development and validation.
- Solo by Pls_toolbox: exploratory analysis and preliminary model development.

3.4.2. Risk Assessment

Ishikawa diagram (Figure 3. 17) was done as a starting point for the risk analysis to detect possible causes that can affect the NIR method development for the Nimesulide product[14].

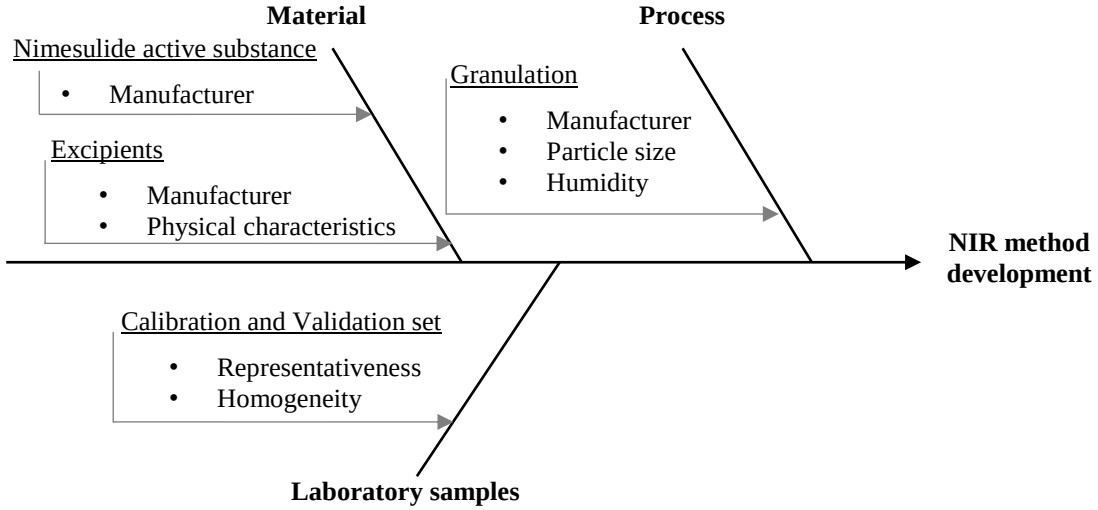


Figure 3. 17 Ishikawa diagram for the NIR method development for the Nimesulide product.

The main areas of the Nimesulide manufacturing process, as well as the calibration and the validation set development, were investigated to find the probable risk factors that might cause the failure of the development of the active substance content quantification method based on NIR spectroscopy. The risk analysis is shown in Table 3. 14. The severity ranking for associated risks to a NIRS method was described as shown in Table 2. 11.

Table 3. 14 Risk analysis and mitigation actions for the NIR procedure for the Nimesulide product.

Cause	Rank	Effect	Mitigation
Material Active substance. Manufacturer	3	Medium risk expected. Two manufacturers are homologated. Differences in NIR spectra	One manufacturer was included in the calibration and validation set with the MB due to the periodicity of arrival
Material Excipients. Manufacturer	1	None expected. For all excipients, except for sucrose powder, one manufacturer is homologated	The two homologated manufacturers were included with the MB. Both are used indistinctly
Material Excipients. Physical properties	1	None expected. The initial physical characteristics of granular excipients don't affect the final particle size or water content	Do not apply
Laboratory samples Calibration and validation set development	4	Medium risk. Heterogeneous or non-equivalent laboratory samples from MB	Each laboratory step to obtain Nimesulide product at the laboratory scale was verified involving validated physicochemical assays
Process Wet granulation. Particle size	4	Medium risk. The particle size has a huge effect on the NIR spectrum (scattering effect)	The process was validated and the particle size didn't have much variability. Expected variability introduced on the calibration and validation sets with MB
Process Wet granulation. Humidity	4	Medium risk. NIR is very sensitive to water content	The process was validated and the water content had to be less or equal to 5%. Expected variability introduced on the calibration and validation set with production samples

3.4.3. Development and Validation of a NIR Spectroscopic Qualitative Analysis

Sub-cascading spectral library was used as an identification and qualification strategy for the Nimesulide product. The identification spectral library was structured in one general library, which includes all SM tested (placebo, excipients, active substance, and the studied pharmaceutical product), and two sub-libraries at the second level. This cascading strategy was necessary to allow resolving qualification ambiguities. Moreover, a pharmaceutical product manufactured with the same equipment was included as a negative sample.

The spectral differences between the included materials (active substance, placebo, major excipients, and MB) are shown in Figure 3. 18. The sucrose, as the major excipient, the placebo, and the MB presented similar spectra. The active substance presented a specific spectral profile. A slight variation in the wavelength was observed where the active substance presents one of its main peaks; between 6200 and 5300 cm^{-1} (carbonyl group (R_2CO)), as well as characteristics peaks in the area between

9000 and 8000 cm^{-1} (C-H bond), from 7000 to 6500 cm^{-1} (R_2NH bond) as well as from 4700 to 4000 cm^{-1} (C-H bond)[15]. These regions could be enough to observe differences between the Nimesulide product and the placebo. The interval from 6200 to 5200 cm^{-1} was considered as the interval range with the main peak for the active substance to evaluate the possible differentiation between the concentration levels of manufacturing products under-dosed and over-dosed.

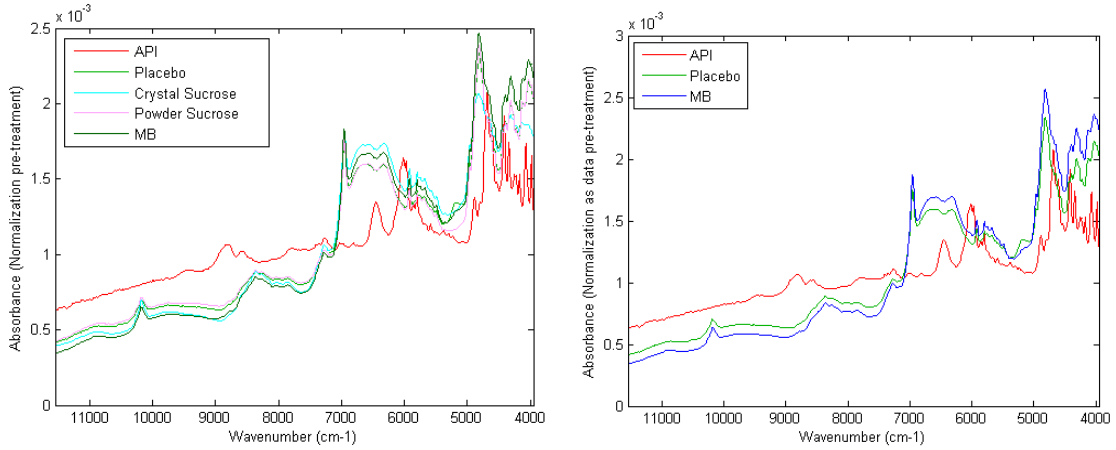


Figure 3. 18 Diffuse reflectance spectra for the active substance, placebo, major excipients (powder and crystal sucrose) and MB for Nimesulide product. Normalization was applied as data pre-treatment. All wavelength ranges included.

Distance as factorization in a PC space was used as a discriminant criterion. For the first level, a wide spectral range was selected (11000-4016 cm^{-1}) and first derivative S-G with 9 points of window width followed by vector normalization were used as data pre-treatments. An ambiguous identification for some of the materials included was obtained and automatically detected by the chemometric software. The sub-library cascading strategy was applied including as a discriminant method the factoring decomposition in a PC space.

The global threshold was adjusted after internal validation allowing correcting unambiguous identification.

The detailed structure and parameters of the library are shown in Figure 3. 19.

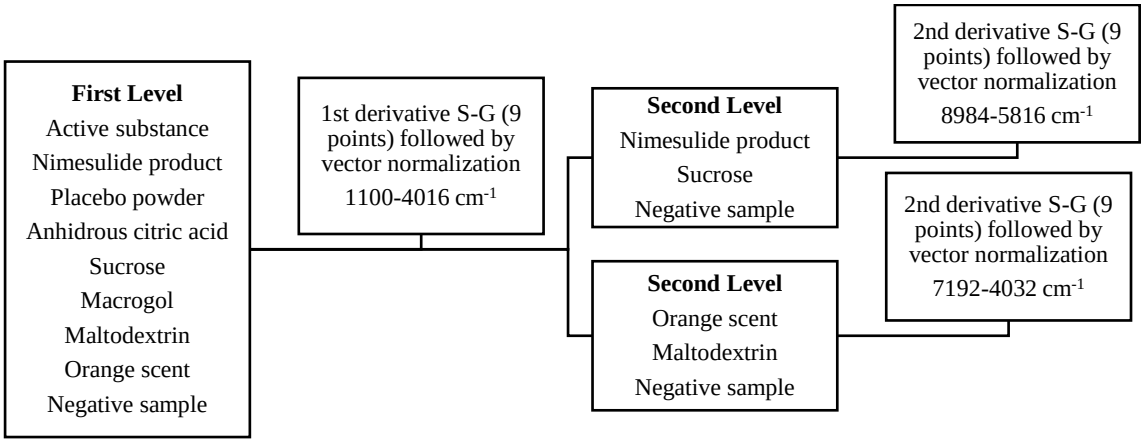


Figure 3. 19 Spectral library for the identification and qualification of Nimesulide product.

The qualitative procedure of the Nimesulide product was validated following the specific guidelines, including specificity and robustness parameters[2,12].

The specificity was tested by internal validation for each library and the sub-library considered the spectra included in the calibration set. Each spectrum was used for calibration and internal validation. As a result, the internal validation showed that all calibration spectra belong to each of the assigned classes created (referred to each SM included). There were no ambiguities during the calibration identification. All calibration spectra were used as positive and negative samples to confirm the ability to identify all classes and to discriminate within them. The negative samples included were correctly identified without ambiguities with the included classes. Moreover, the selectivity parameter, automatically calculated by the OPUS software, demonstrated the unequivocal identification of all the materials included in the library. The selectivity parameter showed a value higher than 2 for 97 % of the included data thus showing that the classes are completely independent. The remaining 3% includes those classes with grouping in contact but not with overlap (see Figure 3. 20).

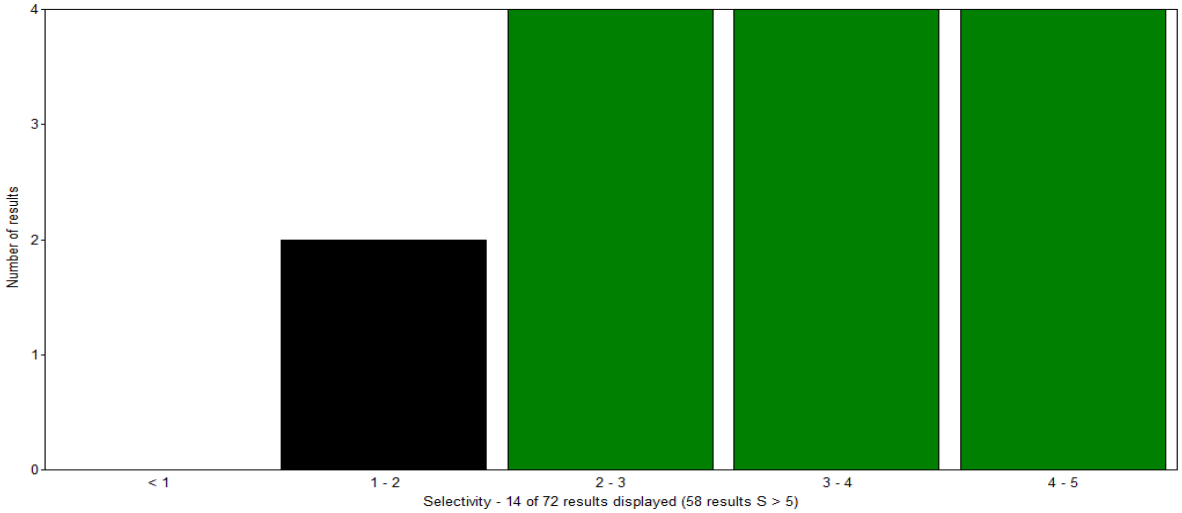


Figure 3. 20 Selectivity histogram for the qualitative model of Nimesulide product.

The robustness was assessed by including a period timeline of one year, considering all the materials included for the external validation set, being enough to include all possible relevant variables such as SM variation, operating and environmental manufacturing conditions modifications, and instrument-changes related to the Nimesulide manufacturing process.

If satisfactory results with the identification were obtained the quantitative procedure could be applied.

3.4.4. Development and Validation of a NIR Spectroscopic Quantitative Analysis

PLS has been the mathematical algorithm used for the active substance for uniformity of dosage unit determination of Nimesulide product. The calibration set was used to calculate the model and the external validation to optimize it.

The specific wavelength ranges included were from 9296 to 8648, 6280 to 5880, and 5048 to 4368 cm^{-1}, where the correlation between the NIR signal and the active substance content changes was demonstrated. First derivative S-G with 17 of window width calculated with 2^{nd} polynomial order followed by MSC was selected to maximize differences between concentrations (see Figure 3. 21). The excluded ranges were not provided relevant information to improve the predictive ability of the calibration model.

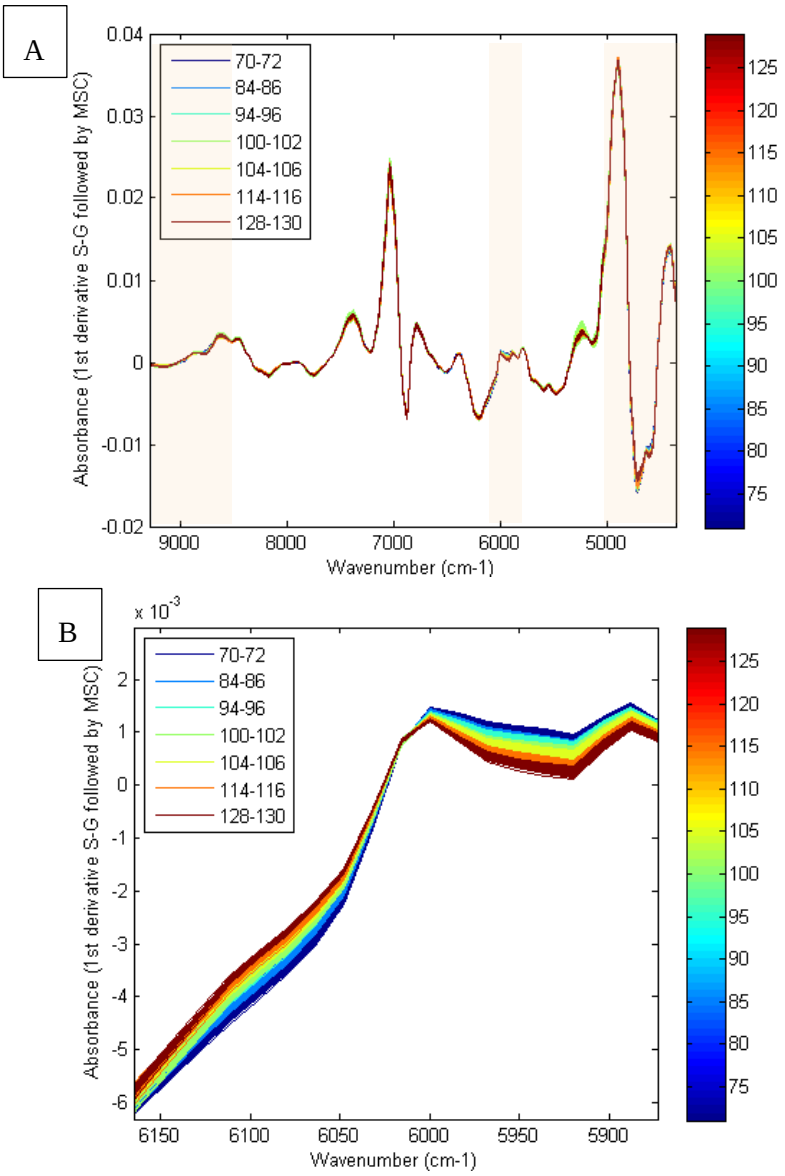


Figure 3. 21 Diffuse spectral data for the calibration set. First derivative S-G with 17 points of widow width followed by MSC as data pre-treatment. A) Spectral range from 9296 to 8648, 6280 to 5880, and 5048 to 4368 cm^{-1}. B) Spectral range from 6152 to 5880 cm^{-1}.

The number of factors was optimized with the external validation set to achieve the lowest RMSEP value. Four factors were finally selected. Some spectra were excluded for presenting a predicted content higher than the calculated by the reference method.

The scores plot for the factors included shows a sample trend related to the active substance content through the second factor mainly (see Figure 3. 22).

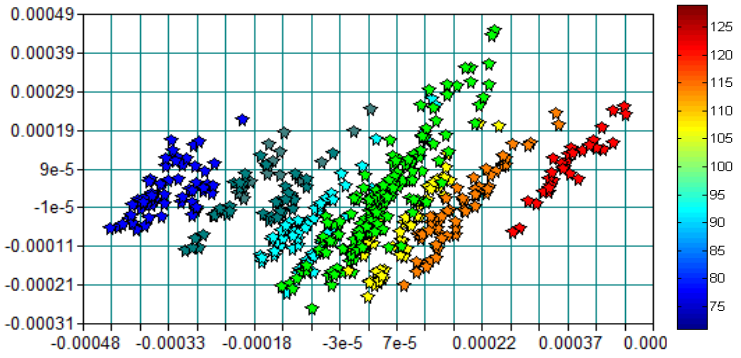


Figure 3. 22 Scores plot for the calibration set of the Nimesulide product NIRS PLS model.

Moreover, the loading plot for the factors included in the model was related to the principle active substance bands (see Figure 3. 23).

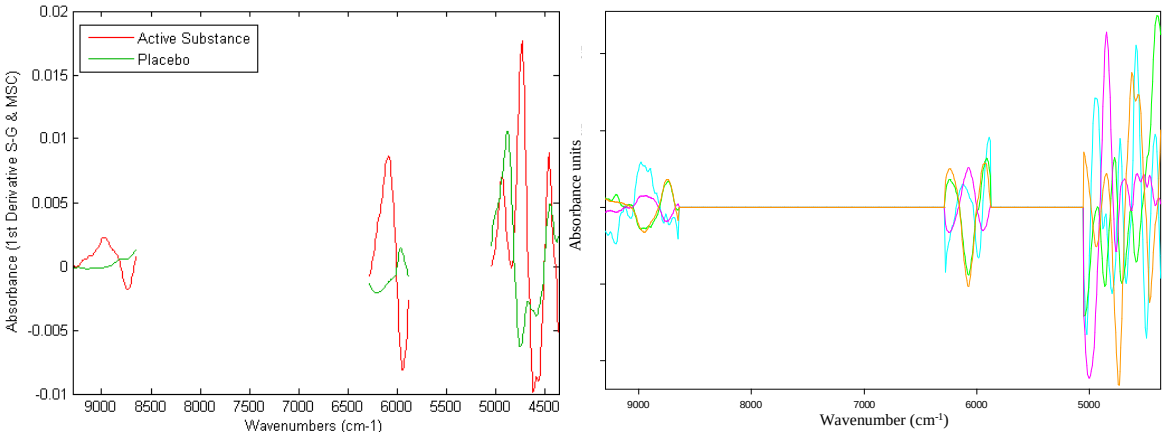


Figure 3. 23 Left) Comparison plot between active substance and placebo spectra. Right) Loadings plot for the model factors (LV1 orange, LV2 pink, LV3 green, and LV4 light blue).

The figures of merit for the model development are shown in Table 3. 15.

Table 3. 15 Figures of merit for the model development of the active substance quantification by NIRS for Nimesulide product.

Parameter	**PLS for active substance quantification**
Number of samples included	61 (42 laboratory samples/ 19 MB)
Acquisition mode	Diffuse reflectance
Spectral pre-treatment	First derivative S-G with 17 points of window width followed by MSC
Range (%)	69.86-133.02
Working range (cm^{-1})	9296 to 8648, 6280 to 5880, and 5048 to 4368 cm^{-1}
PLS factors	4
Explained variance (%)	96.6
Regression Y_{HPLC} vs. Y_{NIRS}	
slope	0.98
Y-intercept	3.37
RMSEC/RMSECV (%)	2.74/2.81

The scatter plot of the active substance content against the reference HPLC values showed a linear relationship between the calibration and the validation sets (see Figure 3. 24).

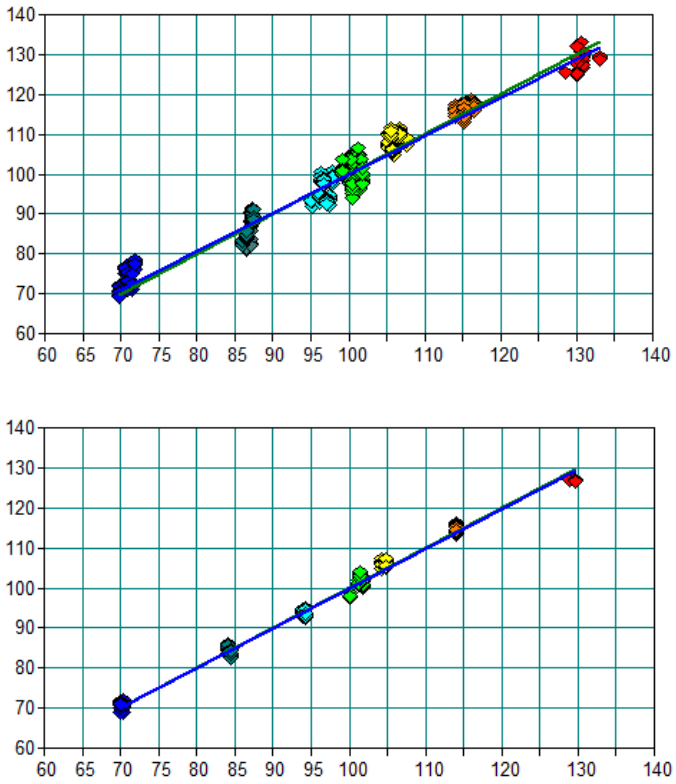


Figure 3. 24 Predicted vs. true values represented in a scatter plot for the calibration (up) and the validation set (down) for Nimesulide product after PLS modelling.

The procedure validation including specificity, linearity, accuracy, precision, and robustness, was designed according to the specific guidelines[12,16,17]. The figures of merit are presented in Table 3. 16.

Table 3. 16 Figures of merit for the validation of the quantitative NIR procedure for Nimesulide product.

Parameter	Results		Conclusion
Specificity	-The main peaks of the active substance were presented in the loadings for the PLS modelling (see Figure 3. 23). -Final product library developed by sub-cascading spectral library strategy analysis for the calibration set, allows the unequivocal identification of each group of samples included in the library (see chapter 3.4.3). -MD evaluation. The limit was fixed by the software at 0.035. All the spectra included for the validation set presented a lower MD than the critical one and were considered spectrally similar to the calibration set, so the predicted values obtained were reliable out samples up to the limit value (see Figure 3. 25).		The NIR procedure is specific
Linearity	Number of samples	17 (12 laboratory samples/5 MB)	The NIR procedure is linear
	Range (%)	70.0-129.6	
	Explained variance (%)		
	y-intercept ± IC (α=0.05) including 0	0.940 ± [-0.989:2.10]	
	Slope ± IC (α=0.05) including 1	0.994± [0.979:1.01]	
	R>0.95	0.996	
	The residuals between the reference HPLC and NIRS methods did not show any trend regarding concentration levels (see Figure 3. 26).		
Accuracy	Number of samples	17 (12 laboratory samples/5 MB)	The NIR procedure is accurate
	A paired t-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 0.03 *t*-tabulated: 1.98 *p*-value: 0.970	
	RMSEP (mg/g)	1.32	
Precision	SEP<1.4SEL	1.39<1.4x0.99	The NIR procedure is precise
	SEP/SEL close to 1	1.32/0.99	
	Validation bias value close to 0	0.00544	
Repeatability	RSD (%) for 10 measures of the same batch less than 5%	1.54	
Intermediate Precision	One-way ANOVA factor (day) *p*-value higher than 0.05 *F*-calculated lower than the *F*-tabulated	*F*-calculated: 1.61 *F*-tabulated: 3.35 *p*-value: 0.220	
Range	Range≥10SEP	30≥10*1.32	The NIR procedure presents a valid range
Robustness	Number of samples	162 MB	The NIR procedure is robust
	A paired *t*-test (two tiles): *t*-calculated lower than *t*-tabulated *p*-value higher than 0.05	*t*-calculated: 0.930 *t*-tabulated: 1.98 *p*-value: 0.350	
Control Model Performance	UAL: $\bar{x}$+3σ UWL: $\bar{x}$+2σ CL: $\bar{x}$ LWL: $\bar{x}$-2σ LAL: $\bar{x}$-3σ *Based on the calculated residuals*	5.42 3.67 0.16 -3.35 -5.10	The NIR procedure is under control

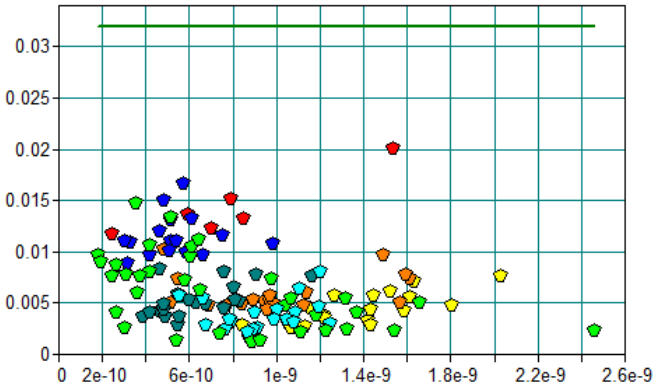


Figure 3. 25 Scatter plot for MD for each included sample into the validation set of Nimesulide product.

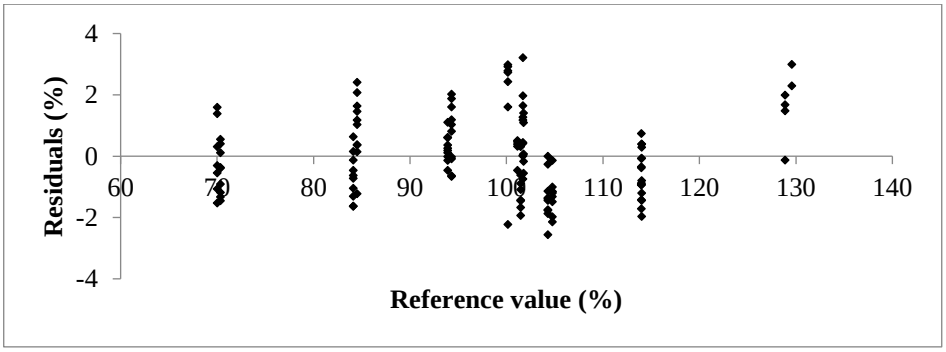


Figure 3. 26 Residuals plot vs assay values.

3.5.Conclusions

The correct maintenance of the CQAs for the Nimesulide granules before the dosage step is demonstrated. The implementation of the PAT by NIRS demonstrated the applicability of the technique through the product and process life cycle.

Moreover, the proposed offline NIR procedures allows the assay and content uniformity determination of the pharmaceutical product only if the analyzed batch was previously satisfactorily identified as a Nimesulide product using the validated NIR reference spectral library.

It was necessary to achieve the spectral data in diffuse reflectance with 32 scans and 16 cm^{-1} of a resolution, involving the integration sphere measurement module adapted with a sample wheel of 30 positions. As a result, the two-level spectral library structure demonstrated the unequivocal ability to identify and qualify the Nimesulide product. Moreover, the PLS model including the first 4 factors and working range from 9296 to 8648 cm^{-1}, 6280 to 5880 cm^{-1}, and 5048 to 4368 cm^{-1}, after the first derivative S-G with 17 points of window width followed by MSC as data pre-treatment, was satisfactorily validated in terms of specificity, linearity, accuracy, precision, and robustness.

Different parameters for spectral acquisition were tested without finding optimal results in terms of linearity and accuracy.

The proposed NIR methodologies allow the correct CQAs determination during the manufacturing process and at the final Nimesulide product to guarantee the product quality before the batch release.

3.6.Bibliography

(1) Swarbrick, B.; Gadsby, M.; Wempen, F. *QbD & PAT*, Optimal Special.; Mengle, R., Ed.; John Wiley & Sons: Chichester, West Sussex, 2018.

(2) Food and Drug Administration. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Rockville, MD, September 2004, pp 1–19.

(3) Gouveia, F. F.; Felizardo, P. M.; Menezes, J. C. Lifecycle Management of PAT Procedures. In *Multivariate Analysis in the Pharmaceutical Industry*; Elsevier, 2018; pp 323–345. https://doi.org/10.1016/b978-0-12-811065-2.00016-3.

(4) European Medicines Agency. Chapter 6: Quality Control. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 28, 2014, pp 1–8.

(5) Brassier Font, J. Implementación de Técnicas Analíticas Rápidas Para La Fabricación Farmacéutica Avanzada, Universitat Autònoma de Barcelona, Bellaterra, 2019.

(6) European Medicines Agency. Annex 11: Computerised Systems. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, June 30, 2011, pp 1–5.

(7) European Medicines Agency. Chapter 4: Documentation. *Good Manufacturing Practice Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, May 17, 2011, pp 1–9.

(8) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 30, 2015, pp 1–16.

(9) Blanco Romía, M.; Alcalá Bernárdez, M. NIR Spectroscopy in Pharmaceutical Analysis: Off-Line and At-Line PAT Applications. In *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*; Bakeev, A. K., Ed.; John Wiley & Sons, 2010; pp 463–491.

(10) Cogdill, R. P.; Anderson, C. A.; Drennen Iii, J. K. Process Analytical Technology Case Study, Part III: Calibration Monitoring and Transfer. *AAPS PharmSciTech* **2005**, *6* (2), 284–297. https://doi.org/10.1208/pt060239.

(11) Wise, B. M.; Roginski, R. T. A Calibration Model Maintenance Roadmap. *IFAC-PapersOnLine* **2015**, *28* (8), 260–265. https://doi.org/10.1016/j.ifacol.2015.08.191.

(12) Medicines Agency, E. Guideline on the Use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. *EMEA/CHMP/CVMP/QWP/17760/2009*. London, UK, January 20, 2012.

(13) European Pharmacopoeia. 2.9.40. Uniformity of Dosage Units. In *European Pharmacopoeia 11th edition*; 2023; Vol. 0, pp 421–423.

(14) European Medicines Agency. ICH Guideline Q9 on Quality Risk Management. *EMA/CHMP/ICH/24235/2006*. London, UK, September 2015.

(15) Workman, Jerry.; Weyer, Lois. *Practical Guide to Interpretive Near-Infrared Spectroscopy*, 1st Edition.; CRC Press, 2008. https://doi.org/https://doi.org/10.1201/9781420018318.

(16) European Medicines Agency. ICH Guideline Q2(R2) on Validation of Analytical Procedures. *EMA/CHMP/ICH/82072/2006*. London, UK, March 31, 2022.

(17) Food and Drug Administration. Analytical Procedures and Methods Validation for Drugs and Biologics Guidance for Industry. Silver Spring, MD, July 2015, pp 1–18.

4. QUALITY BY DESIGN AND CLEANING VALIDATION, IMPLEMENTING A TOTAL ORGANIC CARBON METHODOLOGY

CV is a regulatory requirement based on the cGMP to prevent cross-contamination or adulteration from the manufacturing equipment/systems, which can be safely used to manufacture consecutive products (the same or different products are included). The word "product" could be referred to as an active substance, a drug product, or other kinds of formulations[1,2].

The implementation of CV implies process understanding due to its criticality in pharmaceutical manufacturing.

The lifecycle for the CV approach includes different steps[1]:

- Cleaning process design and development. The Standard Operation Procedure (from now on SOP) has to be developed to reflect sufficient details to guarantee process consistency. The SOP issues are detailed in Figure 4. 3.

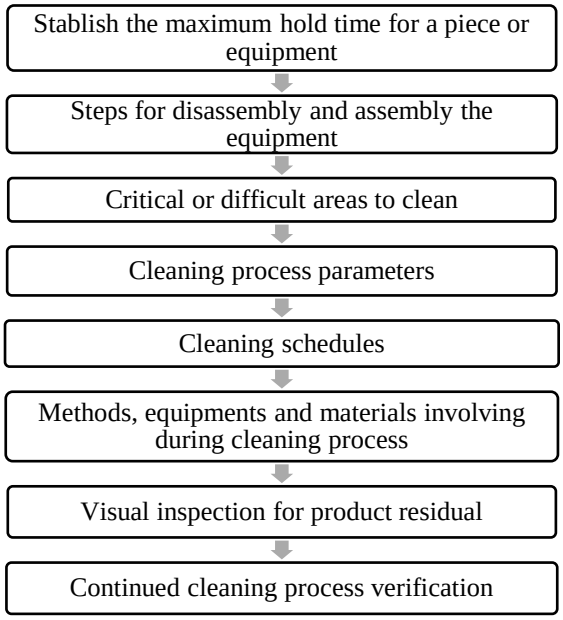


Figure 4. 3 SOP issues for cleaning process development.

- Process qualification activities have to be performed to demonstrate the consistency of the cleaning process presented, including specific protocol runs.
- Ongoing validation maintenance.
- Cleaning verification for a specific cleaning event, including the visual inspection or/and analytical verification.

The cleaning verification, which implies the evaluation of the effectiveness of the cleaning process, involves the analysis of the samples carried out on the surfaces of the equipment based on the

previously established limits. The analytical methods used to verify the absence of the contaminant must be properly validated to demonstrate their ability to detect any remaining agent which could compromise the safety, quality, and efficacy of the following batches of a drug within the studied equipment.

The application of AQbD to the validation and, if considered, to the verification of the cleanliness of the manufacturing processes of pharmaceutical products is presented in this case study. The workflow is detailed in Figure 4. 4. The first step consists of the identification of the ATP and selecting the analytical appropriate method. Then, the implementation of risk assessment to obtain a rational understanding of CMP that affects the CAA of the analytical method. Then the strategy for method development takes place including the implementation of the DoE, establishment of the MODR, and the NOC. Finally, the ACS is defined, and the monitorization for continuous improvement[3,4].

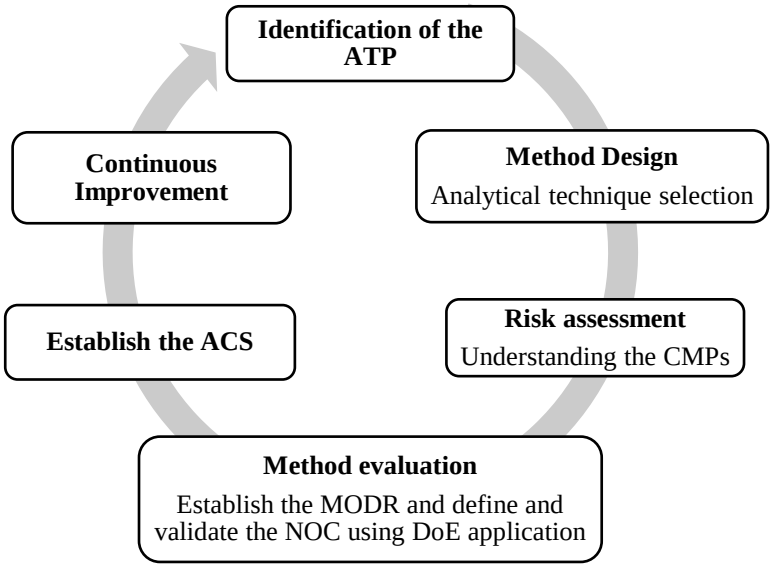


Figure 4. 4 Workflow for AQbD implementation

Specific documentation has to be prepared to assure a reliable, reproducible, and effective cleaning procedure allowing the reduction of the interested contaminants below the specified acceptable limit. The guidelines published related to CV are detailed in Table 4. 1.

Table 4. 1 List of guides published by organizations and regulatory agencies related to CV.

Organism	Title	Status
FDA	*Guide to Inspections of Validation of Cleaning Processes*[5]	Approved (1993)
WHO	*Supplementary Guidelines on Good Manufacturing Practices: Validation*[6]	Approved (2006)
PDA	*Technical Report Nº 29. Points to Consider for Cleaning Validation*[1]	Approved (2012)
APIC	*Guidance on aspects of Cleaning Validation in Active Pharmaceutical Ingredients Plants*[7]	Approved (2014)
EMA	*Guidelines on Setting Health-Based Exposure Limits for use in Risk Identification in the Manufacture of Different Medicinal Products in Shared Facilities*[8]	Approved (2014)
EU-GMP	*Annex 15: Qualification and Validation*[2]	Approved (2015)

Active Pharmaceutical Ingredients Committee (APIC).
European Medical Agency (EMA).
Parenteral Drug Association (PDA).

Different analytical methods were used for the CV. Historically, specific analytical methods such as HPLC, have been the most commonly used for the quantitation of the residual agent for CV. Over time, the indirect TOC technique is involved due to its better process understanding, faster sample analysis, the opportunity for atline or online efficiency gains, and lower operating costs.

TOC could be applied to any residue containing significant amounts of organic carbon in its chemical formula. The common strategy to use the TOC technique is to assume that all residues detected come from the worst-case target residue (for example the API). Selecting a compound with adequate aqueous solubility is a key characteristic for the good use of the technique[1]. To consider the feasibility of the TOC analysis three main factors have to be considered[9]:

- Cleaning process and sample compatibility. Understand the possibilities of the sampling process (see chapter 4.1.2)
- Limit acceptable criteria for the cleaning process (see chapter 4.1.2).
- Agent solubility and recovery (from now on abbreviated as R) during cleaning (see chapter 4.2.2).

TOC methodology is also described in the Pharmacopoeia (Ph. Eur., USP, and the Japanese Pharmacopoeia)[10–12]. Pharmacopoeia specifies a maximum limit value of TOC according to the type of water evaluated being the specific limit for the purified water involved in the pharmaceutical industry ≤500 ppb TOC.

This chapter aims to describe the strategy followed to approach the CV after the pharmaceutical manufacturing of any product that contains DKP-T as an active substance for the TOC technique installed in an online mode as an alternative to the actual reference analytical method based on HPLC.

4.1. Materials and Methods

This section describes the instrumentation involved, the samples analyzed, and the software used through the development of this experimental chapter.

4.1.1. Total Organic Carbon Instrument

The instrument involved was an online M9 TOC Sievers adapted with an online sampling system (iOS) for offline measurements. Key parameters for TOC experiments include the acid and oxidant flow rate (1.0 and 1.8 µL/min), optimized during the analytical method development, and the default values by the instrument vendor: 4 injections per sample analyzed (first cleaning injection of 7 min that are rejected, and 3 more injections around 2 min each).

When performing TOC testing, it is important to minimize contamination risks. Certified low TOC vials (with TOC value <10 ppb) were used to ensure the representativeness of TOC measurements.

4.1.2. Sampling Methods

To evaluate the cleaning effectiveness, the residue amount in the equipment after the cleaning can be determined involving two different sampling methods:

- **Rinse sampling.** The sampling strategy is internally standardized at LMSA and involves a flowing solvent to remove the residue (which is commonly water, if not a water/organic solvent mixture, an aqueous solution, or an organic solvent). The strategy followed for the rinse sampling at the production equipment is based on taking an aliquot from the final rinse solvent of the cleaning process. In the case of laboratory-prepared samples, rinse sampling was simulated by suspending a spike coupon over a clean collected surface, from the prepared stock solution, and the solvent solution cascaded across the surface into the TOC vial (prepared in 25 mL of solvent) [1]. See Figure 4. 5.

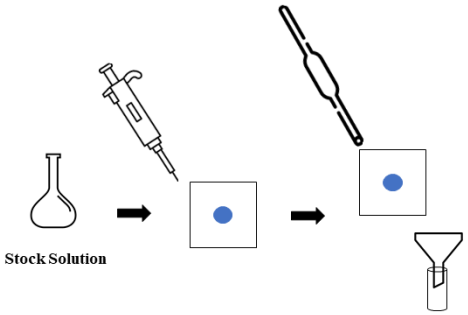


Figure 4. 5 Rinse sampling for laboratory-prepared samples.

- **Swabbing.** The strategy followed for the swabbing is internally standardized at LMSA involving wiping the equipment surface with fibrous material (polyester-tipper swab) and transferring the residue from the surface to the fibrous material. Then, the material is placed in a solvent solution which is analyzed. In the case of laboratory-prepared samples, the swabbing was simulated to suspend a spike coupon, from the stock solution, above a clean collected surface. Once the amount spiked was dried, the residue is removed with swabbing. A polyester-tipped swab was moistened with the solvent and passed through the surface previously spiked with an aliquot of the target residual, sampling 25 cm^2, and poured into the TOC vial [1]. See Figure 4. 6.

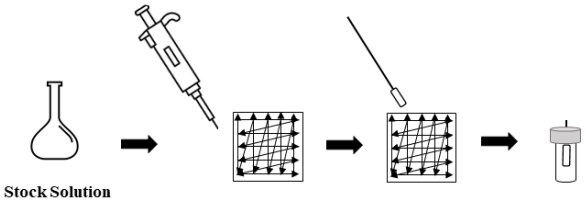


Figure 4. 6 Swabbing for laboratory-prepared samples.

The choice for rise sampling or swabbing depends on the type of manufacturing equipment. Rinse sampling is used as the starting point for evaluating equipment cleanliness. During rinsing, the entire product-containing the surface is wetted[1]. However, swabbing is commonly involved in hard-to-clean areas, allowing the analysis of the target residue on specific surfaces. If swabbing of the selected areas is not easy enough, rinse sampling is presented as an alternative[7].

Different types of surfaces, including stainless steel, epoxy resin, and plastic, could be involved. All of them should have different recoveries during sampling. For this reason, recovery studies during method validation are necessary when applying sampling methods to determine residues (see chapter 4.2.2)[9].

Moreover, three specific samples in the solvent solution (a solvent in which the target residual is completely soluble) have to be prepared to achieve representative results:

- Stock solution is prepared in a volumetric flask as a concentrated solution from which laboratory rinse samples and swabbing, as well as reference samples, are prepared by dilution to multiple concentrations.
- Reference sample is prepared in a volumetric flask at the studied concentration. The same amount of the agent that the once-to-be spiked up to the studied surface is added directly to a specific volume of solvent solution (25 mL).
- Blank sample for each sampling methodology, prepared in the laboratory but without the contaminant agent studied, following the same strategy as for the samples.

Finally, the sealed TOC vials are sonicated for 10s to be homogenized before TOC analysis (see chapter 4.2.2.).

4.1.3. Acceptance Limit Calculation

The calculation of the acceptance criteria defined as the maximum allowable carryover (from now on MACO) is based on the acceptable daily exposure (from now on ADE) following Equation 4. 1[7].

$$\frac{\text{ADExMBS}}{\text{TDD}} = MACO \text{ (mg)}$$

ADE: mg/day.

MBS: the minimum batch size for the next product (mg).

TDD: the standard therapeutic daily dose for the next product (mg/day).

Equation 4. 1.

If a homogeneous distribution is assumed on all studied surfaces, the organic carbon concentrations for the stock solutions are calculated as shown in Equation 4. 2 and Equation 4. 3 for the rinse and swab test limit, respectively[7,9].

$$\frac{\frac{\text{MACO}}{\text{Total sampled surface}}}{\text{V}} \text{ x \%C} = \text{ppm C}$$

$$\%C = \frac{\text{Milligrams C}}{\text{Molecular Weight}}$$

C: carbon atom.

Milligrams C: nº C*molecular weight C.

Total sampled surface: surface sampled for the evaluated equipment (dm^2). Commonly 0.25 dm^2.

V: solvent volume in the TOC vial (L).

Equation 4. 2.

$$\frac{\text{MACO}}{\text{V}} \text{ x \%C} = \text{ppm C}$$

Equation 4. 3.

In the case of the swab test, during CV and verification, hard-to-clean parts should be determined and selected as the worst-case limit for the whole equipment train.

In this study, specific acceptance limits were calculated for both sampling methods described in Table 4. 2.

Table 4. 2 Acceptance limits for rinse sampling and swabbing.

Sampling Method	UAL TOC (ppm)	LAL TOC (ppm)
Rinse sampling	6.21	0.58
Swabbing	2.56	0.13

4.1.4. Sofware

The M9 laboratory analyzer interface (Suez, USA) was used for data acquisition, and the M9 online analyzer interface (Suez, USA) for instrument control.

Excel 2019 (Microsoft Office) was used for general statistical calculations.

4.2. Method Development

This section describes the method development involving the risk assessment for method variable selection and the method optimization.

4.2.1. Risk Assessment Implementation

A risk assessment based on FMEA methodology[13] was developed to evaluate the impact of each method variable (CQAs), including selectivity, linearity, accuracy, the LoD, the LoQ, precision, and robustness (see Table 4. 3)[14]. A factor W for each CQA was considered (see Table 4. 4), to be the resulting quality of the S, see Table 4. 5), the P of the error occurrence (see Table 4. 6), and the D of the error (see Table 4. 7). As a result, the RPN or the CN [3,4] were calculated (see Equation 1. 12. from Part III).

Table 4. 3 CQA description[15].

CQA	Interpretation
Selectivity	Ability to unequivocally ensure that the presence of organic carbon is due to the tested substance and not to any other expected external substance
Linearity	Ability to confirm that the concentration of the target residue is directly proportional to the TOC measurement
Accuracy	Referring to the closeness between the experimental results and the theoretical results. Recovery (rinse sampling/swabbing) experimental concentration vs. theoretical concentration
LoD	The lowest detectable concentration. Evaluation of whether a signal is the result of instrument noise or the response of the compound
LoQ	The lowest quantifiable concentration
Precision	Measured as the standard deviation or the RSD %. Related to the closeness with which multiple analyzes of a given sample agree with each other
Robustness	Ability to remain unaffected by small but determined variations

Table 4. 4 W criteria.

W	Description
1	Low impact on the total quality of the result to be obtained
3	Intermediate impact on the total quality of the result to be obtained
5	High impact on the total quality of the result to be obtained

Table 4. 5 S criteria.

S	Description
1	Low impact: CQA is affected by a severe failure in the method variable
3	Intermediate impact: CQA is significantly affected by a severe failure in the method variable
9	High impact: CQA is significantly affected by a soft failure in the method variable (sensitive to method variables)

The S parameter is calculated as the sum of all the severity values for each CQA multiplied by the factor weight for each CQA divided by the sum of the factor weights, see Equation 1. 13. from Part III.

Table 4. 6 P criteria.

P	Frequency of occurrence
1	Highly improbable
2	Remote Possibility
3	Occasional
4	Probably
5	Frequent

Table 4. 7 D criteria.

D	Probability of detection
1	Very high
2	High
3	Medium
4	Low
5	Very low

Risk situations were considered for two different scenarios:

- Where the combination between W*S was equal to or greater than 105. This means that 2 of the 7 CQAs contain the highest values of W and S or the intermediate values of S and the highest values of W.
- RPN≥35, where the combination between W*S was 90 or greater. This means that at least 1 of the 7 CQAs contains the highest values of W and S, and the values of P and D were equal to or greater than 3.

The ones considered critical were marked in red coloration in the risk assessment table. The risk was considered significant and the proposed mitigation actions were established to, as far as possible, eliminate, reduce, or control possible errors. See Table 4. 8.

Some mitigation actions require experimental studies that are summarized in chapter 4.2.2, including agent solution solubility, stability of agent solution, TOC consumables ratio, black influence, and optimization of extraction conditions after swabbing sampling strategy.

Table 4. 8 FMEA risk assessment for the analytical TOC procedure.

CQAs of the TOC analytical method					**Selectivity**	**Precision**	**Accuracy**	**LoD**	**LoQ**	**Linearity**	**Robustness**	**S**	**Causes of the error**	**P**	**D**	**RPN**	**Proposed preventive actions**
				W	1	5	5	1	3	3	5						
Analytical Procedure		**Potential error***		**Error effect****	1: Null 3: Medium 9: High												
Experimental procedure to obtain the solution to be evaluated	Preparation of the solution to be evaluated	Solubilization of the agent to be evaluated in the solvent solution	Non-solubilization of the agent in the solvent chosen to perform the TOC measurement	Obtaining false negatives. Low repeatability of the TOC measurement	1	9	9	3	3	3	9	**6.8**	Erroneous selection of the solvent solution	**3**	**2**	**41**	Solubility study of the agent to be evaluated (experimental and bibliographic)
		Evaporation time	Low time	Low substance concentration	3	9	3	3	3	3	9	**5.6**	Evaporation time is not enough. The possible effect of the solvent solution on the TOC measurement	**3**	**1**	**17**	Blank TOC measurement for the same experimental conditions and verification by visual inspection
				High substance concentration	3	9	3	1	1	3	9	**5.3**		**3**	**1**	**16**	
			High time	Low substance concentration	3	9	1	3	3	3	9	**5.2**	A possible influence of external factors such as environmental conditions	**3**	**3**	**47**	Control, as far as possible, the environmental and working conditions, as far as possible
				High substance concentration	3	3	1	1	1	3	3	**2.2**		**3**	**3**	**20**	
		Homogenization of the solution for TOC measurement	Low time	Low substance concentration	1	9	3	3	3	3	9	**5.5**	Not enough time for the complete homogenization of the solution	**3**	**3**	**50**	Evaluate the experimental design for the selection of the optimal conditions
				High substance concentration	1	9	1	1	1	3	3	**3.4**		**3**	**3**	**31**	

			CQAs of the TOC analytical method	**Selectivity**	**Precision**	**Accuracy**	**LoD**	**LoQ**	**Linearity**	**Robustness**	**S**	**Causes of the error**	**P**	**D**	**RPN**	**Proposed preventive actions**
			W	1	5	5	1	3	3	5						
Analytical Procedure		**Potential error***		**Error effect****	1: Null 3: Medium 9: High												
Experimental procedure to obtain the solution to be evaluated	Preparation of the solution to be evaluated	Sample preparation by dilution	An experimental process with many dilutions	Obtaining false positives	3	3	1	3	3	3	9	**3.9**	Error in the choice of the experimental procedure. A possible influence of external factors	3	3	**35**	Statistical evaluation of TOC measurements for different experimental procedures
		Solution stability	Unstable sample during normal working conditions	Preparation of the solution every time you want to evaluate. Obtaining false positives	3	9	3	3	3	3	9	**5.6**	Instability of the TOC measurement due to the influence of the stability of the solution and/or the influence of environmental conditions	2	3	**34**	Establish an experimental procedure for the evaluation of the stability of the solution in different working conditions
	Blank preparation	According to the methodology used in the preparation of the solution to be evaluated	Low concentration	Obtaining false positives and/or false negatives regarding the identification and quantification of the studied substance	9	9	3	9	9	3	9	**6.9**	Influence of external factors: environmental conditions, analytical procedure, solvent solution, etc., on the TOC measurement	**3**	**3**	**62**	Repeatability evaluation of the blank and its influence on the TOC measurement of the substance
			High concentration		3	9	3	3	3	3	9	**5.6**		**3**	**3**	**50**	
	Swabbing	Effect of the surface used to perform the test	Low concentration	False positives	9	3	9	9	9	3	9	**6.9**	Influence of environmental conditions	**4**	**3**	**83**	Control of environmental and work conditions, as far as possible
			High concentration		3	3	3	1	1	3	3	**2.7**		**4**	**3**	**32**	

CQAs of the TOC analytical method				**Selectivity**	**Precision**	**Accuracy**	**LoD**	**LoQ**	**Linearity**	**Robustness**	**S**	**Causes of the error**	**P**	**D**	**RPN**	**Proposed preventive actions**
W				1	5	5	1	3	3	5						
Analytical Procedure		**Potential error***		**Error effect****	1: Null 3: Medium 9: High											
Experimental procedure to obtain the solution to be evaluated	Swabbing	Evaporation time	Low time – Low substance concentration	False positive	9	9	9	9	9	3	9	**8.2**	Insufficient evaporation time. Influence of the solvent or environmental conditions on the TOC measurement	**2**	**2**	**33**	Visual inspection and verification by TOC measurement. Planning standardized experimental procedure
			Low time – High substance concentration		9	3	9	1	1	3	9	**5.5**		**2**	**2**	**22**	
			High time – Low substance concentration		3	3	3	3	3	3	9	**3.0**	Influence of the environmental conditions	**3**	**4**	**36**	Control of environmental and working conditions, as far as possible
			High time – High substance concentration		3	3	3	1	1	3	3	**2.7**		**3**	**4**	**32**	
		Swabbing procedure	Saturation of the swab with the solvent	Obtaining false negatives	3	9	9	1	1	1	9	**6.3**	Soaking the swab with the solvent solution for too long. Do not drag the entire evaporated sample with the swab	**2**	**1**	**13**	Establish an experimental procedure for sampling by swab following the internal instruction of LMSA
			Lack of solvent on the swab		1	9	9	1	1	1	9	**6.2**	Not immersing the swab sufficiently in the solvent solution	**2**	**1**	**12**	

CQAs of the TOC analytical method					**Selectivity**	**Precision**	**Accuracy**	**LoD**	**LoQ**	**Linearity**	**Robustness**	**S**	**Causes of the error**	**P**	**D**	**RPN**	**Proposed preventive actions**
W					1	5	5	1	3	3	5						
Analytical Procedure		**Potential error***		**Error effect****	1: Null 3: Medium 9: High												
Experimental procedure to obtain the solution to be evaluated	Swabbing	Extraction time	Low time – Low substance concentration	Low repeatability of the TOC measurement. Obtaining false negatives	9	9	3	9	9	9	9	**7.8**	There is no extraction of the substance to analyze the swab. Incorrect selection of extraction conditions	**3**	**3**	**69**	Optimization of the extraction time through the construction of an experimental design
			Low time – High substance concentration		3	9	3	1	1	3	9	**5.2**		**3**	**3**	**47**	
			High time – Low substance concentration	Low repeatability of the TOC measurement. Obtaining false positives	3	3	3	3	3	3	3	**3.0**	A possible effect of the hyssop. Influence of environmental conditions. Incorrect selection of extraction conditions	**3**	**3**	**27**	
			High time – High substance concentration		1	3	1	1	1	1	3	**1.9**		**3**	**3**	**17**	
		Agitation mode	Insufficient agitation, incomplete substance extraction	Low repeatability of the TOC measurement. Obtaining false negatives	1	3	9	3	3	3	9	**5.6**	Incorrect selection of the agitation mode	**3**	**4**	**67**	Optimization of extraction mode, using experimental design construction
	Solution preparation	TOC vial	Low substance concentration	Low repeatability of the TOC measurement. Obtaining false negatives or/and positives	9	9	9	3	9	9	9	**8.7**	Reuse of TOC vials	**3**	**3**	**79**	Use new TOC vials. Certification of analysis <10 ppb TOC
			High substance concentration		3	3	3	1	3	3	9	**4.3**		**3**	**3**	**39**	

				CQAs of the TOC analytical method	**Selectivity**	**Precision**	**Accuracy**	**LoD**	**LoQ**	**Linearity**	**Robustness**	**S**	**Causes of the error**	**P**	**D**	**RPN**	**Proposed preventive actions**
				W	1	5	5	1	3	3	5						
Analytical Procedure		**Potential error***		**Error effect****	1: Null 3: Medium 9: High												
Measurement acquisition	Reagents involved during TOC measurement	Oxidant flow	Low flow	Obtaining false negatives. At high substance concentration, insufficient flow to oxidize the entire sample. Not reproducible	1	9	3	3	3	3	9	5.5	Incorrect selection of measurement conditions	**3**	**2**	**33**	Design of experiments for the selection of the flow of acid and oxidant. The specifications described in the internal work instruction by LMSA are considered
			High flow	Incorrect operating conditions	1	1	1	1	1	1	1	**1.0**		**2**	**2**	**4.0**	
	Reagents involved during TOC measurement	Acid flow	Low flow	At high concentration of substance insufficient flow to oxidize the entire sample. Obtaining false negatives	1	9	3	3	3	3	9	5.5	Incorrect selection of measurement conditions	**3**	**2**	**33**	DoE for the selection of the flow of acid and oxidant. The specifications described in the internal work instruction by LMSA are considered
			High flow	At high substance concentration, excessive flux for complete oxidation	1	1	1	1	1	1	1	**1.0**		**2**	**2**	**4.0**	

				CQAs of the TOC analytical method	**Selectivity**	**Precision**	**Accuracy**	**LoD**	**LoQ**	**Linearity**	**Robustness**	**S**	**Causes of the error**	**P**	**D**	**RPN**	**Proposed preventive actions**
				W	1	5	5	1	3	3	5						
Analytical Procedure		**Potential error***		**Error effect****	1: Null 3: Medium 9: High												
Measurement acquisition	Time between measurements	Measurement time established according to internal LMSA instruction. Time established by LMSA to carry out the measurements of 1 hour	High wait time between replicates	Obtaining false positives	1	3	3	1	1	3	3	**2.6**	Influence of ambient conditions	**2**	**1**	**5.1**	Organization of the experimental process, grouping the replicates in the measurement interval of 1 hour
	The working environment while taking TOC measurement	Ambient light	Presence of ambient light	Obtaining false positives and/or negatives	9	3	1	1	1	3	9	**3.8**	Influence of environmental conditions	**2**	**1**	**7.6**	Control of working conditions, as far as possible
		Organic solutions	Presence of organic solutions	Obtaining false positives	9	9	3	3	3	3	9	**5.8**		**2**	**1**	**12**	Control of environmental and working conditions, as far as possible

*Agent is understood, in this case, as the set of evaluated APIs.

**False positives and/or negatives, refers to obtaining a TOC measurement higher or lower than the theoretical one calculated, and therefore the one prepared experimentally.

4.2.2. Method Optimization

This chapter describes the factors evaluated for method optimization for the DKP-T contaminant based on the preliminary risk assessment.

4.2.2.1. Agent Solution Solubility

DKP-T is a high-water soluble agent. The solubility of the agent was checked with a homogeneous 7.60 ppm TOC solution.

4.2.2.2. Stability of Stock Solution

The stock solution at the higher TOC concentration (20% above the UAL TOC, 7.60 ppm) was demonstrated to be stable during 6 hours under normal working storage conditions (24ºC at 65% HR). Based on a two-way ANOVA statistical test, the stability of the solution was demonstrated considering sample preparation and day as critical factors. The statistical *p*-value obtained, in both cases, was higher than the specification (significance level of 0.05); 0.41, and 0.29, respectively.

Moreover, based on a two-way ANOVA statistical test, the stability of stock solution was demonstrated to be stable during 4 days under 4ºC storage conditions. The statistical *p*-value obtained for sample preparation and day factors was higher than the specification (significance level of 0.05); 0.16, and 0.11, respectively.

4.2.2.3. TOC Consumables Ratio

The acid and the oxidant flow rate were specifically selected depending on the concentration range evaluated. The student solution was considered at the higher TOC concentration (20% above the maximum acceptable TOC value, 7.60 ppm) and prepared as a stock solution.

The acid flow rate was established at 1.00 µL/min as a supplier recommendation.

The acceptance criterion to achieve an optimal oxidant flow rate was established with recovery results≥80% and an RSD≤1%.

The oxidant flow rate was selected based on a quadratic DoE model including 5 experiments and 3 central points (by duplicates) variating within the specific range proposed by the supplier (from 0.7 to 2.8 µL/min). The acid flow rate was established as a constant variable.

Because the data does not follow a linear trend (see Figure 4. 7 and Table 4. 9), it was not possible to extrapolate the data and thus choose the optimal oxidant flow. Therefore, considering the R data obtained (described in Equation 4. 4), the first one closer to 100% and with a value of RSD<1% was

considered the optimal oxidant flow value (1.8 µL/min). It was verified that if a lower oxidant flow had been chosen, it would not have a sufficient oxidizing capacity for the concentration studied.

$$R\% = \frac{Corrected\ TOC\ value}{Theoric\ TOC\ value} * 100$$

Equation 4. 4.

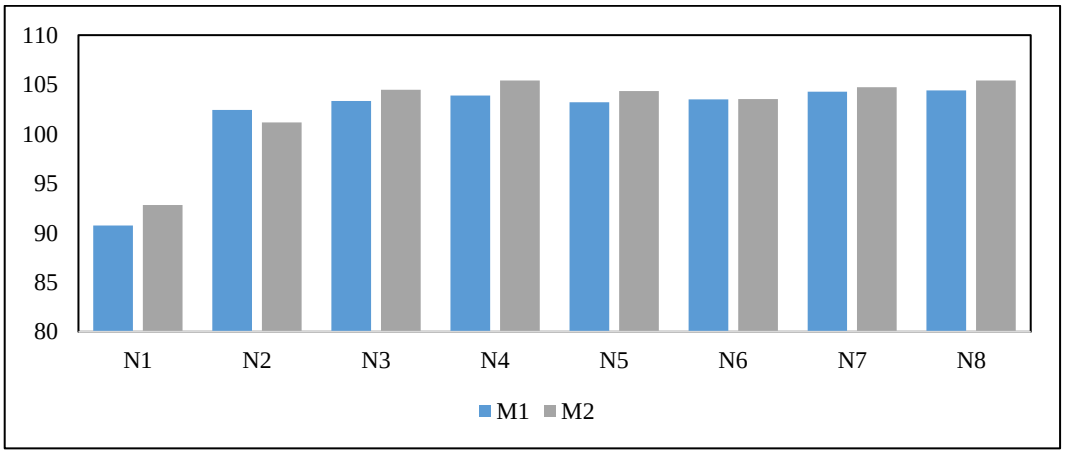


Figure 4. 7 N. º of experiments vs. R (%) for the development of the TOC method of the DKP-T (N1 to N6 experimental conditions; M1 and M2 each replicate).

Table 4. 9 Conditions of acid and oxidant flow rate for each experiment.

N. º of experiment	Acid flow rate (µL/min)	Oxidant flow rate (µL/min)	R (%)	RSD (%)
N1	1	0.700	91.8	8.49
N2	1	1.30	102	5.00
N3	1	1.80	104	0.745
N4	1	2.30	105	0.570
N5	1	2.80	104	0.660
N6	1	1.80	104	0.702

4.2.2.4. Blank Influence on the Evaluated Sample

The blank contribution was shown to be significant during the measurement of the agent in the solvent solution[16].

The blank was prepared from purified water from the production plant. Greater representativeness and reproducibility were observed for the working concentration studied when the preparation of the working samples was carried out with this same water.

Moreover, the experimental evaluation of the effect of the blank was carried out with the preparation of 6 independent blanks, considering the same preparation as that carried out for the LAL of TOC for the set of samples to be prepared (reference sample, rinse samples and swabbing samples), using

purified water as solvent. This experiment was repeated over two days. The acid and oxidant flow rates for the TOC measurement set were 1.0 and 1.8 µL/min, respectively.

A two-factor ANOVA statistical test was performed with a single sample per group and significant differences were obtained between the factors evaluated (p-value<0.05) (including the factors: day, samples, and experimental procedures). As a result, the preparation of a blank for each set of experiments was performed before each studied sample preparation.

4.2.2.5. Optimization of Extraction Conditions After Swabbing

A full factorial DoE was developed including different time and agitation modes (vortex and ultrasonic agitation) and time within 1 to 10 min (see Table 4. 10).

Table 4. 10 DoE experimental conditions.

Type of agitation	Vortex	R (%)	RSD (%)	Ultrasonic	R (%)	RSD (%)
Time 1 (min)	1	98.2	1.33	1	100	0.640
	1	98.8	2.47	1	102	0.890
	1	101	3.29	1	95.4	0.290
Time 2 (min)	5	101	0.950	5	98.9	1.20
	5	98.8	1.40	5	96.5	1.93
	5	96.1	0.780	5	94.0	0.830
Time 3 (min)	10	100	1.33	10	99.4	0.660
	10	101	1.23	10	99.4	0.550
	10	97.9	0.220	10	96.1	0.340

Based on a two-way ANOVA statistical test with various samples per group (three), there were no significant differences between samples for any of the modes of agitation (including as critical factors: sample, time, and agitation mode) obtaining a p-value>0.05. The acceptance criterion will be marked by the lowest RSD value (%) within experiments, to demonstrate the repeatability of the data. The selected condition was 10 min by ultrasonic agitation to achieve an homogeneous mixture with a simple agitation mode and with less interaction of the analyst, since the %R factor was >80% in any of the evaluated agitation methods.

4.3. Method Validation

The ICH Q2 (R1) and (R2) guidelines[15,17], and the specific TOC guideline[7] refer to the requirements to validate the TOC methodology. The acceptance criterion is summarized in Table 4. 11.

The detailed results for each validated parameter are included in this section.

Table 4. 11 Table of merit for DKP-T TOC method validation.

Parameters	Acceptance criterion	Result
Specificity	Without significant differences between techniques for the purified water TOC value (HPLC-TOC)	PASS
Linearity	Response vs Concentration. Linear relationship.	PASS
	y= ax +b	PASS
	r y $r^2 \geq 0.990$	PASS
	Y-interception (b)	PASS
	Confidence Interval (CI) including 0 with a probability of 95%.	PASS
LoD	$\frac{3.3\sigma}{S}$ lower than the LAL of TOC.	PASS
LoQ	$\frac{10\sigma}{S}$	PASS
	p-value>0.05	PASS
	IC (p=0.05) includes the calculated concentration	PASS
Accuracy	R≥ 80 %	PASS
	IC *(p=0.05)*	PASS
	Statistical test t-Student *p-value*>0.05 (95% confidence level)	PASS
	$F_{Calculated}<F_{Tabulated}$ /*p-value*>0.05	PASS
R%	R≥50 %	PASS
	IC (p=0.05, *n*-1) (%)	PASS
	$F_{Calculated}<F_{Tabulated}$ / *p-value*>0.05	PASS
Precision. Repeatability	Overall RSD ≤10.0%	PASS
	RSD for each concentration level~30%	PASS
Precision. Intermediate Precision	Without significant differences between days for each concentration *p-value*>0.05	PASS
Robustness	Paired *t*-test: $t_{calculated}<t_{tabulated}$	PASS

4.3.1. Specificity

The specificity evaluation was developed for two different concentration levels, 80 and 120 percent of the highest value for the UAL TOC, 6.20 ppm TOC (see Table 4. 2). The prepared samples were analyzed by the TOC method developed, and the one currently authorized as a reference method.

Based on a one-way ANOVA statistical test non-significant difference was obtained between methods (the probability (*p-value)* was greater than the significance level (α=0.05) with a confidence

level of 95 percent, and the variance between samples was lower than the variation within samples ($F_{calculated}<F_{tabulated}$)). Moreover, the RSD in percent between both methods was lower than 1 (see Table 4. 12).

Table 4. 12 Result table for the detection of organic carbon with authorized and TOC methods.

Method	Sample (%)	Measured value (µg)
TOC	80	185
	120	283
	Blank	4.30
Authorized HPLC method	80	188
	120	281
	Blank	Not detected
RSD (%) between techniques<1		
One-way ANOVA (method as a critical factor): *p-value>0.05 (0.24)*; $F_{calculated}<F_{tabulated}$ *(2.66<18.51)*		

4.3.2. Detection and Quantitation Limits

The LoD and LoQ parameters were calculated considering the included variability in the set of samples prepared for the regression line due to the criticality of the sample preparation and the evaluated blank effect during sample measurements.

Five concentration levels were included along the studied range being an estimated LoD concentration, the LAL of TOC for DKP-T for the overall sample methodologies. This concentration was expanded around ±20% (60, 80, 100, 120, and 140% from the specified lower acceptable concentration). Three replicates were analyzed for each concentration level. The validated parameters were calculated following Equation 4. 5.

$$LoD = \frac{3.3\sigma}{S}; LoQ = \frac{10\sigma}{S}$$

Where:

σ: standard deviation associated with the y-intersection for the evaluated regression.

S: standard deviation associated with the slope.

Equation 4. 5.

The regression line is presented in Figure 4. 8.

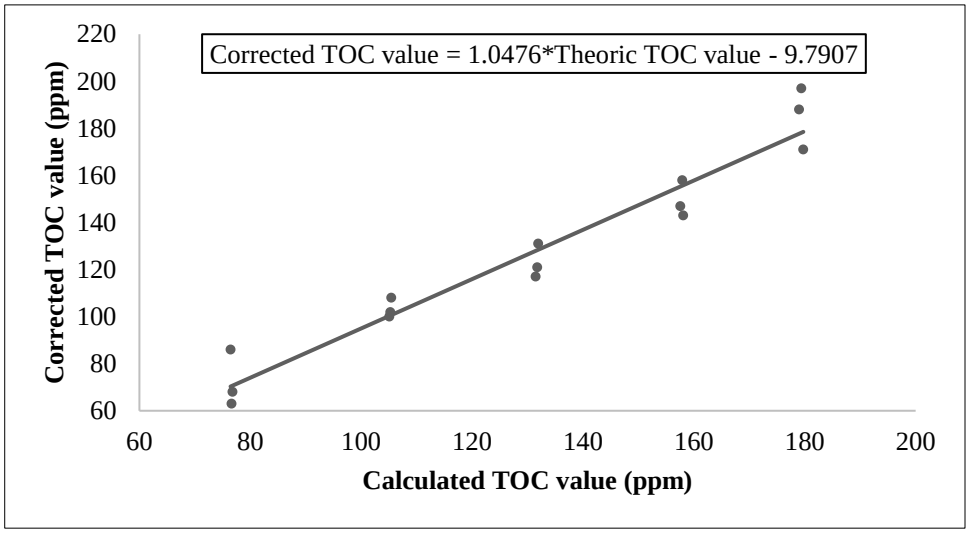


Figure 4. 8 Regression line for the LoD and LoQ evaluation.

The *y*-interception obtained was -9.79 and 10.01 was the standard deviation associated with the slope value. An LoD of 31.5 ppb TOC and an LoQ of 95.5 ppb TOC were calculated.

The results were validated for both parameters with 6 replicates for each[17] (see Table 4. 13).

Table 4. 13 LoD and LoQ validation.

Nº of sample	**Corrected TOC value (ppb)**	**R (%)**
LoQ-1	75.2	77.8
LoQ-2	90.2	93.4
LoQ-3	70.2	72.7
LoQ-4	111	115
LoQ-5	89.2	92.3
LoQ-6	160	166
IC confidence level 95% (ppb)	[66.9-139]	
Statistic t-student p-value (confidence level 95%, α=0.05)	0.866>0.05	
Calculated TOC value (ppb)	97.0	
Sample	**Corrected TOC value (ppb)**	
LoD-1	30.5	
LoD-2	57.0	
LoD-3	16.5	
LoD-4	33.2	
LoD-5	41.2	
LoD-6	55.2	
Calculated TOC value (ppb TOC)	*32.3*	

4.3.3. Linearity

The linearity was evaluated for six concentration levels[17] including the 12.5, 25, 50, 100, and 120% from the UAL TOC value and the 80% from the LAL of TOC, including both sampling methodologies (see Table 4. 2).

Reference sample preparation was followed by spiking the stock solution prepared at the specified working concentration to the calculated volume for the solvent solution, directly into the TOC measuring vial.

The linear regression is shown in Figure 4. 9. Based on a preliminary visual inspection, a linear trend was observed. The calculated R and the R^2 were 0.9999 and 0.9996, respectively, obtaining a good model fitting. Moreover, the CI included 0 (-0.062:0.029).

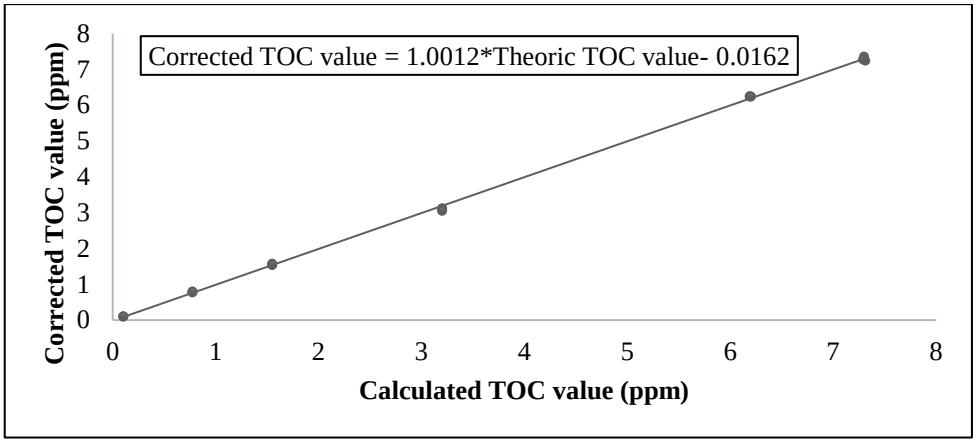


Figure 4. 9 Linear regression evaluation.

4.3.4. Accuracy

The accuracy was evaluated after the linearity was achieved. Three concentration levels were analyzed, for the UAL TOC values selected as the worst-case concentration, being 80, 100, and 120 percent, including both sampling methodologies (see Table 4. 2). Each concentration level was analyzed in triplicate for the reference sample[17]. Moreover, for both sampling methodologies, the three different surfaces used for the equipment involved during CV were studied including stainless steel, plastic, and epoxy resin, to simulate manufacturing equipment.

One of the critical parameters evaluated during accuracy validation was the R%. The reported R values were the average of three individual samples expressed as shown in Equation 4. 6. It was calculated based on the measured TOC values of the working solution and the reference solution. The subtraction of the TOC reading of the blank solution was considered[5].

$$\%Recovery = \frac{TOC_W}{TOC_R} x100$$

Where:

TOC_R: TOC reading solution for the reference sample subtracting the TOC blank measure.

TOC_W: TOC reading solution for the studied sample subtracting the TOC blank measure.

Equation 4. 6.

The obtained results are shown in sections 4.3.4.1, 4.3.4.2, and 4.3.4.3, being accepted results as the criterion established in Table 4. 11.

4.3.4.1.Accuracy Test for the Reference Sample

The accuracy results for the reference samples are presented in Table 4. 14. All results meet the criteria specified in Table 4. 11, without significant differences between concentration levels for each acceptable limit evaluated.

Table 4. 14. Accuracy validation results for calculated acceptable TOC limits for the reference sample.

120		**100**		**80**	
UAL (6.21 ppm TOC)					
$\bar{x}$ (ppm)	7.92	$\bar{x}$ (ppm)	7.07	$\bar{x}$ (ppm)	4.97
s (ppb)	0.60	*s* (ppb)	0.46	*s* (ppb)	0.28
IC (ppm) [6.10-9.08]		IC (ppm) [5.33-7.60]		IC (ppm) [4.29-5.66]	
p-value (95% confidence level) 0.506		*p-value* (95% confidence level) 0.420		*p-value* (95% confidence level) 0.695	
$\%\bar{R}$	104	$\%\bar{R}$	104	$\%\bar{R}$	102
CI (%R) [84.0-124]		CI (%R) [86.6-122]		CI (%R) [88.01-115]	
One-way ANOVA (concentrations): $F_{calculated} < F_{tabulated}$ (0.14<5.14); *p*-value>0.05 (0.87)					
LAL (130 ppb TOC)					
$\bar{x}$ (ppb)	155	$\bar{x}$ (ppb)	131	$\bar{x}$ (ppb)	104
s (ppb)	11.1	*s* (ppb)	2.56	*s* (ppb)	1.42
IC (ppb) [127-182]		IC (ppb) [124-137]		IC (ppb) [100-107]	
p-value (95% confidence level) 0.756		*p-value* (95% confidence level) 0.777		*p-value* (95% confidence level) 0.556	
$\%\bar{R}$	98.6	$\%\bar{R}$	99.9	$\%\bar{R}$	99.7
CI (%R) [81.1-116]		CI (%R) [94.8-105]		CI (%R) [96.1-103]	
One-way ANOVA (concentrations): $F_{calculated} < F_{tabulated}$ (0.68<5.14); *p*-value>0.05 (0.54)					

$\bar{x}$: mean value for the three replicates, for each concentration level, from the corrected TOC value.
s: standard deviation for the three replicates, for each concentration level.
CI: Confidence Interval.
%R: Corrected TOC value against Calculated TOC value, in percentage.

4.3.4.2.Recovery Test for Rinse Sampling

Recovery values for the UAL TOC involving rinse sampling are shown in Table 4. 15. All results meet the criteria specified in Table 4. 11, without significant differences between concentration levels for each surface material. The lowest recovery value for each surface studied was selected as the worst case for future sample measurements.

Table 4. 15 Recovery results for rinse sampling including INOX, plastic, and epoxy resin plates for the UAL.

120		**100**		**80**	
INOX					
$\bar{x}$ (ppm)	7.33	$\bar{x}$ (ppm)	6.37	$\bar{x}$ (ppm)	4.92
s (ppb)	0.17	*s* (ppb)	0.10	*s* (ppb)	0.12
IC (ppm) [6.91-7.75]		IC (ppm) [6.11-6.61]		IC (ppm) [4.62-5.22]	
$\%\bar{R}$	99.7	$\%\bar{R}$	102	$\%\bar{R}$	101
CI (%R) [93.9-105]		CI (%R) [97.92-106]		CI (%R) [95.2-107]	
Individual worst-case R (%)= 98.00					
One-way ANOVA (concentrations): $F_{calculated} < F_{tabulated}$ (0.89<5.14); *p*-value>0.05 (0.46)					
PLASTIC					
$\bar{x}$ (ppm)	6.25	$\bar{x}$ (ppm)	5.37	$\bar{x}$ (ppm)	4.38
s (ppb)	0.990	*s* (ppb)	0.230	*s* (ppb)	0.540
IC (ppm) [3.80-8.71]		IC (ppm) [4.78-5.95]		IC (ppm) [3.02-5.73]	
$\%\bar{R}$	86.3	$\%\bar{R}$	85.9	$\%\bar{R}$	90.1
CI (%) [52.5-120]		CI (%) [76.6-95.2]		CI (%) [62.3-118]	
Individual worst-case R (%)= 71.0					
One-way ANOVA (concentrations): $F_{calculated} < F_{tabulated}$ (0.15<5.14); *p*-value>0.05 (0.86)					
EPOXY RESIN					
$\bar{x}$ (ppm)	4.35	$\bar{x}$ (ppm)	3.89	$\bar{x}$ (ppm)	3.32
s (ppb)	0.480	*s* (ppb)	0.350	*s* (ppb)	0.320
CI (ppm) [3.15-5.55]		CI (ppm) [3.03-4.75]		CI (ppm) [2.53-4.11]	
$\%\bar{R}$	60.0	$\%\bar{R}$	65.7	$\%\bar{R}$	70.0
CI (%) [43.4-76.6]		CI (%) [51.5-79.7]		CI (%) [53.8-86.2]	
Individual worst-case R (%)= 52.0					
One-way ANOVA (concentrations): $F_{calculated} < F_{tabulated}$ (2.39<5.14); *p*-value>0.05 (0.17)					

4.3.4.3.Recovery Test for Swabbing

Recovery values for the UAL TOC involving swabbing are shown in Table 4. 16. All results meet the criteria specified in Table 4. 11, without significant differences between concentration levels for each surface material. The lowest recovery value for each surface studied was selected as the worst case for future sample measurements.

Table 4. 16 Recovery results for swabbing including INOX, plastic, and epoxy resin plate for the UAL.

120		**100**		**80**	
INOX					
$\bar{x}$ (ppm)	3.09	$\bar{x}$ (ppm)	2.54	$\bar{x}$ (ppm)	1.98
s (ppb)	0.026	*s* (ppb)	0.0058	*s* (ppb)	0.026
CI (ppm) [3.02-3.15]		CI (ppm) [2.53-2.56]		CI (ppm) [1.91-2.04]	
$\%\bar{R}$	96.1	$\%\bar{R}$	96.1	$\%\bar{R}$	98.3
CI (%R) [94.1-98.2]		CI (%R) [95.6-96.7]		CI (%R) [94.9-102]	
Individual worst-case R (%)= 94.0					
One-way ANOVA (concentration): $F_{calculated} < F_{tabulated}$ (4.30<5.14); *p*-value>0.05 (0.069)					
PLASTIC					
$\bar{x}$ (ppm)	3.14	$\bar{x}$ (ppm)	2.58	$\bar{x}$ (ppm)	2.01
s (ppb)	0.030	*s* (ppb)	0.0058	*s* (ppb)	0.046
CI (ppm) [3.06-3.21]		CI (ppm) [2.57-2.60]		CI (ppm) [1.89-2.12]	
$\%\bar{R}$	97.7	$\%\bar{R}$	97.7	$\%\bar{R}$	99.7
CI (%R) [95.3-99.9]		CI (%R) [97.1-98.2]		CI (%R) [94.1-105]	
Individual worst-case R (%) = 92.0					
One-way ANOVA (concentration): $F_{calculated} < F_{tabulated}$ (2.13<5.14); *p*-value>0.05 (0.20)					
EPOXY RESIN					
$\bar{x}$ (ppm)	2.73	$\bar{x}$ (ppm)	2.43	$\bar{x}$ (ppm)	1.95
s (ppb)	0.18	*s* (ppb)	0.026	*s* (ppb)	0.044
CI (ppm) [2.28-3.18]		CI (ppm) [2.37-2.50]		CI (ppm) [1.85-2.06]	
$\%\bar{R}$	85.59	$\%\bar{R}$	92.18	$\%\bar{R}$	95.30
CI (%R) [71.5-99.7]		CI (%R) [86.7-94.7]		CI (%R) [90.0-101]	
Individual worst-case R (%)= 79.0					
One-way ANOVA (concentration): $F_{calculated} < F_{tabulated}$ (5.03<5.14); *p*-value>0.05 (0.052)					

4.3.5. Precision

4.3.5.1. Repeatability

Repeatability was performed for the reference sample and both sampling methodologies considered the UAL TOC, in both methodologies, as a worst-case concentration (Table 4. 2). For each sample type, three replicates were performed[17]. For rinse samples and swabbing, repeatability was developed involving the same strategy as in the accuracy section. The results in Table 4. 17 for reference sample show that the overall RSD (%) for the three concentrations was less than 10 %.

Table 4. 17 Reproducibility measurements for the reference sample.

UAL TOC for rinse sampling (6.21 ppm TOC)			
Concentration	**120%**	**100%**	**80%**
s	0.60	0.46	0.28
RSD (%)	7.92	7.07	5.56
Overall RSD (%)	5.14		
UAL TOC for swabbing (2.56 ppm TOC)			
Concentration	**120%**	**100%**	**80%**
s	0.068	0.075	0.040
RSD (%)	2.21	2.86	1.99
Overall RSD (%)	2.35		

The results in Table 4. 18 show that the individual RSD (%) for each concentration at each concentration level considering the three surfaces materials and the respective sampling methodologies, was less than 30%. Moreover, the overall RSD (%) was less than 10%. The results presented correspond to the worst case for the three repetitions of each studied concentration repeated on three different days (the mean of the three results at the same concentration level and the same experimental conditions with a higher individual RSD value (%)).

Table 4. 18 Reproducibility measurements for swabbing and rinse samples including INOX, plastic, and epoxy resin plates for their UAL TOC value respectively.

Parameter	**Swabbing**			**Rinse samples**		
INOX						
Concentration	**120%**	**100%**	**80%**	**120%**	**100%**	**80%**
s	0.085	0.062	0.040	0.16	0.066	0.059
RSD (%)	2.84	2.59	2.02	2.17	1.02	1.20
Overall RSD (%)	2.48			1.46		
Plastic						
Concentration	**120%**	**100%**	**80%**	**120%**	**100%**	**80%**
s	0.20	0.01	0.081	0.57	0.23	0.54
RSD (%)	6.25	0.37	3.79	8.47	4.37	12.44
Overall RSD (%)	3.47			8.43		
Epoxy Resin						
Concentration	**120%**	**100%**	**80%**	**120%**	**100%**	**80%**
s	0.18	0.03	0.04	0.48	0.16	0.32
RSD (%)	6.65	1.09	2.23	11.13	3.92	9.58
Overall RSD (%)	3.32			8.21		

4.3.5.2.Intermediate Precision

The influence of the day for each sampling technique was evaluated for its UAL of TOC, considered as the worst-case concentration. Before each analysis, one reference sample was analyzed first. Analyst influence was not performed because the sampling methodologies for routine TOC analysis are internally procedural. Before carrying out the sampling, the personnel involved carry out specific training. The values obtained for each concentration studied during three days, for the different surface materials, involving both specific sampling methodologies, are shown in Table 4. 19 and Table 4. 20.

Table 4. 19 Intermediate precision validation for rinse sampling including INOX, plastic, and epoxy resin surfaces.

Factor		**Day 1**			**Day2**			**Day3**		
Concentration level (%)		**120**	**100**	**80**	**120**	**100**	**80**	**120**	**100**	**80**
Material of the surface		**<u>INOX</u>**								
Parameter	**%$\bar{R}$**	102	104	101	102	104	103	99.7	102	102
	***s* (%R)**	3.09	1.06	1.25	2.25	1.37	0.72	2.33	1.63	2.50
	%CV$_{\%R}$	3.03	1.02	1.24	2.20	1.32	0.70	2.33	1.59	2.46
One-way ANOVA (day) for each studied concentration		**120%**			**100%**			**80%**		
		$F_{calculated}<F_{tabulated}$ (0.77<5.14) *p*-value>0.05 (0.50)			$F_{calculated}<F_{tabulated}$ (1.24<5.14) *p*-value>0.05 (0.36)			$F_{calculated}<F_{tabulated}$ (1.50<5.14) *p*-value>0.05 (0.29)		
Material of the surface		**<u>Plastic</u>**								
Parameter	**%$\bar{R}$**	86.3	85.9	90.1	92.6	91.4	93.0	91.7	89.4	95.2
	***s* (%R)**	13.6	3.75	11.2	7.84	3.92	1.02	6.41	2.71	2.15
	%CV$_{\%R}$	15.8	4.37	12.4	8.47	4.29	1.09	6.99	3.03	2.25
One-way ANOVA (day) for each studied concentration		$F_{calculated}<F_{tabulated}$ (0.34<5.14) *p*-value>0.05 (0.73)			$F_{calculated}<F_{tabulated}$ (2.06<5.14) *p*-value>0.05 (0.21)			$F_{calculated}<F_{tabulated}$ (0.47<5.14) *p*-value>0.05 (0.65)		
Material of the surface		**<u>Epoxy Resin</u>**								
Parameter	**%$\bar{R}$**	65.8	74.1	70.9	64.1	70.7	70.6	60.0	65.5	69.9
	***s* (%R)**	5.79	12.1	7.03	10.3	6.64	6.75	6.68	2.52	6.50
	%CV$_{\%R}$	8.80	16.29	9.91	16.1	9.39	9.55	11.1	3.87	9.29
One-way ANOVA (day) for each studied concentration		$F_{calculated}<F_{tabulated}$ (0.49<5.14) *p*-value>0.05 (0.63)			$F_{calculated}<F_{tabulated}$ (0.91<5.14) *p*-value>0.05 (0.45)			$F_{calculated}<F_{tabulated}$ (0.031<5.14) *p*-value>0.05 (0.97)		

Table 4. 20 Intermediate precision validation for swabbing including INOX, plastic, and epoxy resin surfaces.

Factor		**Day 1**			**Day2**			**Day3**		
Concentration level (%)		**120**	**100**	**80**	**120**	**100**	**80**	**120**	**100**	**80**
Material of the surface		**INOX**								
Parameter	$\%\bar{R}$	97.0	96.4	100	98.7	99.2	102	96.1	96.1	98.3
	s (%R)	3.00	2.58	2.00	1.53	2.10	1.53	0.82	0.27	1.53
	$\%CV_{\%R}$	3.09	2.68	2.00	1.55	2.12	1.50	0.86	0.24	1.55
One-way ANOVA (day) for each studied concentration		**120%**			**100%**			**80%**		
		$F_{calculated}<F_{tabulated}$ (1.27<5.14) *p*-value>0.05 (0.35)			$F_{calculated}<F_{tabulated}$ (2.28<5.14) *p*-value>0.05 (0.18)			$F_{calculated}<F_{tabulated}$ (2.88<5.14) *p*-value>0.05 (0.13)		
Material of the surface		**Plastic**								
Parameter	$\%\bar{R}$	98.7	101	97.8	97.7	101	99.7	104	99.4	95.2
	s (%R)	6.20	0.047	6.50	0.90	3.09	2.20	2.50	2.83	8.10
	$\%CV_{\%R}$	6.25	0.046	6.62	0.96	3.06	2.21	2.38	2.84	8.50
One-way ANOVA (day) for each studied concentration		$F_{calculated}<F_{tabulated}$ (1.93<5.14) *p*-value>0.05 (0.23)			$F_{calculated}<F_{tabulated}$ (0.47<5.14) *p*-value>0.05 (0.65)			$F_{calculated}<F_{tabulated}$ (0.41<5.14) *p*-value>0.05 (0.68)		
Material of the surface		**Epoxy Resin**								
Parameter	$\%\bar{R}$	85.6	92.2	95.3	91.1	96.5	99.3	89.1	93.7	99.7
	s (%R)	5.69	1.00	2.13	7.63	1.84	0.87	3.54	3.17	3.23
	$\%CV_{\%R}$	6.65	1.09	2.23	8.38	1.91	0.88	3.97	3.38	3.24
One-way ANOVA (day) for each studied concentration		$F_{calculated}<F_{tabulated}$ (0.68<5.14) *p*-value>0.05 (0.54)			$F_{calculated}<F_{tabulated}$ (2.92<5.14) *p*-value>0.05 (0.13)			$F_{calculated}<F_{tabulated}$ (3.40<5.14) *p*-value>0.05 (0.10)		

4.3.6. Robustness

During the robustness tests, the effect of the different experimental parameters on the analysis of the reference sample was evaluated. The agent was dissolved in the working water solution and a commercial acid phosphoric 0.072% solution. The influence of the TOC response on the conductivity variations and the pH was evaluated following the recommendations of the equipment supplier. The studied concentrations were 120% of the UAL of TOC and 80% of the LAL of TOC, considering both sampling strategies, to include the overall DKP-T concentration range. The results are shown in Table 4. 21.

Based on a statistical procedure paired sample t-test, no significant differences in the TOC response depending on the solvent solutions were obtained.

Table 4. 21 Validation results for 120% from the upper and 80% from the LAL of TOC.

Acid Phosphoric 0.072 %						
120% of the UAL of TOC				**80% of the LAL of TOC**		
Sample replicate	**Corrected TOC (ppm)**	**Conductivity (mS/cm)**	**pH**	**Corrected TOC (ppm)**	**Conductivity (mS/cm)**	**pH**
S1	7.27	1440.1	2.35	0.104	1492.0	2.36
S2	7.28	1457.6	2.36	0.107	1491.5	2.35
S3	7.27	1455.7	2.37	0.114	1553.4	2.36
S4	7.29	1457.0	2.36	0.120	1542.3	2.36
S5	7.33	1454.1	2.36	0.112	1552.4	2.37
S6	7.39	1454.5	2.36	0.115	1551.7	2.37
$\bar{x}$ Corrected TOC	7.30			0.112		
RSD (%) Corrected TOC	0.65			5.0		
H_2O						
Sample replicate	**Corrected TOC (ppm)**	**Conductivity (mS/cm)**	**pH**	**Corrected TOC (ppm)**	**Conductivity (mS/cm)**	**pH**
S1	7.25	4.13	6.99	0.111	3.10	6.98
S2	7.37	4.50	6.95	0.11	2.80	6.97
S3	7.41	4.37	6.85	0.119	3.57	7.06
S4	7.39	4.07	6.96	0.115	3.33	6.98
S5	7.32	4.60	6.89	0.117	3.47	6.94
S6	7.40	4.73	6.91	0.121	2.70	6.86
$\bar{x}$ Corrected TOC	7.36			0.116		
RSD (%) Corrected TOC	0.83			3.78		
Solvent Results Comparison						
$\bar{x}$ Corrected TOC result	7.31			0.114		
Overall RSD (%) Corrected TOC	0.99			4.40		
Paired t-test two tails	$T_{Calculated}<T_{Tabulated}$ (2.24<2.57); *p-value*>0.05 (0.076)			$T_{Calculated}<T_{Tabulated}$ (1.58<2.57); *p-value*>0.05 (0.17)		

4.4. Risk Assessment for an Online TOC Equipment

Obtaining a representative result of the sample to be analyzed is essential. This section aims to define the main risks to obtaining an OOS value after external actions are performed on the TOC M9 Sievers online equipment.

The methodology followed for the risk assessment is based on an Ishikawa diagram as a risk analysis and a cause-and-effect diagram as a risk evaluation strategy[14].

The main effects of TOC equipment are investigated to find probable risk factors that might cause a measurement equipment failure for TOC. The critical criterion is shown in Table 4. 22. The severity ranking proposed is shown in Table 4. 23.

Table 4. 22 Critical criteria for the cause-and-effect diagram.

High	6	Incorrect TOC results detected during technical support
Medium	3	Incorrect TOC results detected during the measurement time
Low	1	None effect

The Ishikawa diagram was made as a starting point for the risk analysis to detect possible factors that may affect the normal operation of the TOC equipment. The different factors were classified, including online measurements, offline measurements, and water plant management. It is shown in Figure 4. 10.

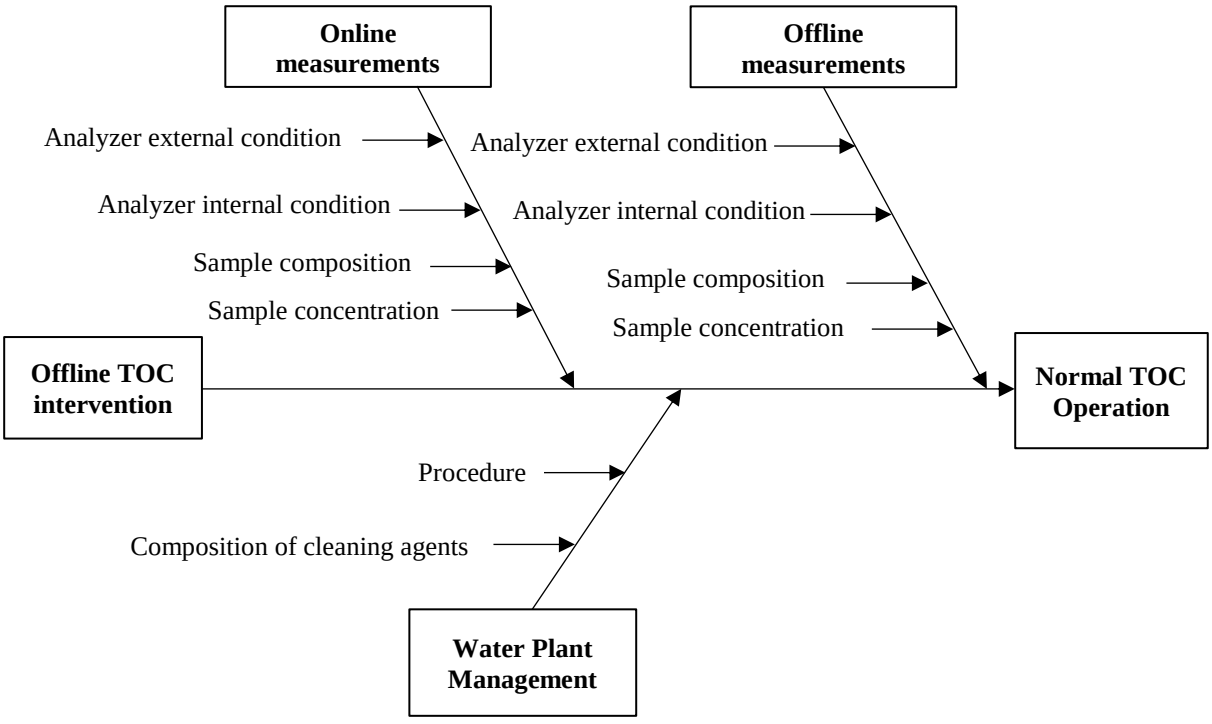


Figure 4. 10 Ishikawa diagram for an Online TOC equipment.

Table 4. 23 Cause and effect diagram for the TOC instrument.

Cause	Definition	Rank	Effect	Mitigation
Online Measurement				
Analyzer external conditions	Evaluate each external part of the TOC equipment, including the condition of the water inlet, the pH of the waste sample stream (confirming that the pH of the waste stream is less than 4), and the condition of the reagent syringe (confirming that there are no bubbles inside)	3	Wrong TOC measurements. Lack of measurements	Periodic internal inspection. Periodic maintenance review
Analyzer internal conditions	Evaluate each refill part for the TOC equipment, including acid, and oxidizer containers, reagent syringes, UV lamp, sample pumps, valves, and internal water reservoir	3	Wrong TOC measurements. Lack of measurements	Periodic external inspection. Periodic maintenance review
Sample composition	The solvent solution properties and the agent included	1	Affect the operation of the TC or IC membrane, and may cause the equipment to break	Commonly purified water is used
Sample concentration	The water concentration must be equal or less than 500 ppb TOC. This parameter must be under control to guarantee the acceptance specification	3	Affects the operation of the TC or IC membrane, and may cause the equipment to break	Commonly, the measured concentration is less than 500 ppb TOC (specification). If the water concentration is higher than the specification, the water composition will be evaluated, investigating the effect on the TOC equipment. A syringe cleaning procedure was applied for a concentration higher than the specified
Offline Measurement				
Analyzer external condition	Evaluate the related external parts of the TOC equipment, including the condition of the water inlet, and the status of the reagent syringe (confirming that there are no bubbles inside)	3	Wrong TOC measurements. Lack of measurements	Periodic internal inspection, before each offline measurement. Periodic maintenance review. Make sure the DI water reservoir is full, especially when running high TOC or high salt samples, to rinse the analyzer by running low TOC water through the internal sample loop

Table 4. 23 Cause and effect diagram for the TOC instrument.

Cause	Definition	Rank	Effect	Mitigation
Analyzer internal condition	Evaluate each spare part for the TOC equipment, including acid and oxidizer containers, reagent syringes, UV lamp, sample pumps, valves, and internal water reservoir	3	Wrong TOC measurements. Lack of measurements	Periodic internal inspection, before each offline measurement. Periodic maintenance review
Sample composition	Different pharmaceutical products are manufactured at LMSA including different APIs and excipients. Furthermore, detergent components might be part of the measured sample. The solvent solution properties and the agent included	3	Could affect the correct operation of the TC or IC membrane, and may cause the equipment to break	Evaluate sample composition before sample measurement. if there is evidence that it is a difficult sample to analyze for TOC, rinse with Milli-Q water, and if normal TOC values are not obtained, clean with a dilute acid solution
Sample concentration	Different API concentrations are present in the measured cleaning validation samples. The sample concentration must be within the equipment specifications to guarantee a correct quantification	3	Affect the operation of the TC or IC membrane as well as obtain possible positive out specifications during online TOC measurements	Rinse the analyzer with Milli-Q water after any offline measurement with a TOC concentration greater than 1 ppm to reduce the result to normal values
Water Plant Management				
Procedure	TOC equipment is installed in the water plant and different cleaning procedures are applied to clean each part	6	Affect the operation of the TC or IC membrane. Modification of the normal TOC values	Isolate the TOC equipment during any procedure that may activate the entry of water into the equipment. Be sure to properly purge the system from the plant before starting the online TOC readings
Composition of the cleaning agent	The only cleaning agent that could interact with the flow sample for the TOC equipment is the Divosan Activ VT5. It is a cleaning agent with peracetic acid 5%	6	Affect the operation of the TC or IC membrane. Non-representative TOC values	Investigate the source of the reagent and assess the TOC effect

The TOC equipment runs self-test activities automatically when starts the analysis, and provides a summary of faults for review generating an automatic error report. This section includes common failures and troubleshooting tips. After confirming that there are no warning or error messages indicating a specific problem, proceed to perform a visual inspection (external and internal) of the analyzer while running a test to identify any obvious symptoms or physical problems.

During external operations, the proper functioning of the TOC equipment may be altered, resulting in incorrect TOC, IC, and/or TC values. Mitigation actions based on the submitted risk assessment were implemented to ensure the quality of the online and offline TOC results.

4.5. Transferability Study Between an Online and Offline TOC Equipment

The pharmaceutical company LMSA purchased a Sievers M9 Laboratory Analyzer for laboratory TOC measurements. An integrated instrument with a laboratory data management system for a paperless workflow.

The current and future process flow is described in Figure 4. 11.

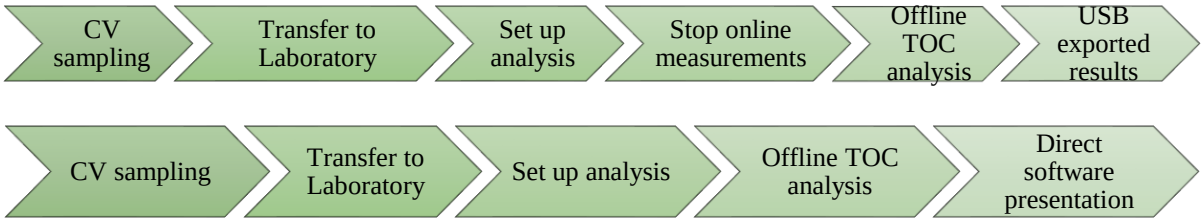


Figure 4. 11 Current (up) and future (down) process flow for TOC analysis.

The strategy followed to evaluate the transferability between TOC equipment[18] is based on precision, accuracy, and linearity verification[19].

The accepted criteria is based on a standard procedure by the equipment provider detailed in Table 4. 24.

Table 4. 24 Table of merit for method transferability.

Parameter	Acceptance criteria	Result
Linearity	Response vs. Concentration. Linear relationship.	PASS
	y=ax+b	PASS
	R and R^2>0.990	PASS
	y-interception (b); IC includes 0 with a probability of 95%	PASS
Precision. Repeatability	RSD <5 (%)	PASS
Precision. Intermediate Precision	One-way ANOVA (equipment): Without significant difference between equipment *p-value>0.05*	PASS
Accuracy	R≥85	PASS
	Statistical test *t*-Student *p-value>0.05* (95% confidence level)	PASS
	$F_{Calculated}<F_{Tabulated}$/ *p-value>0.05*	PASS

4.5.1. Linearity

Certified TOC standards were analyzed with the offline equipment at five different concentrations for certified TOC standards including 250, 500, 1000, 5000, and 10000 ppb TOC. Moreover, a certified blank solution was considered before the sequence measurement.

The linear regression is shown in Figure 4. 12. Based on a preliminary visual inspection, a linear trend was observed. The calculated R and the R^2 were 0.9999, obtaining a good model fitting. Moreover, the CI included 0 (-8.69:27.9).

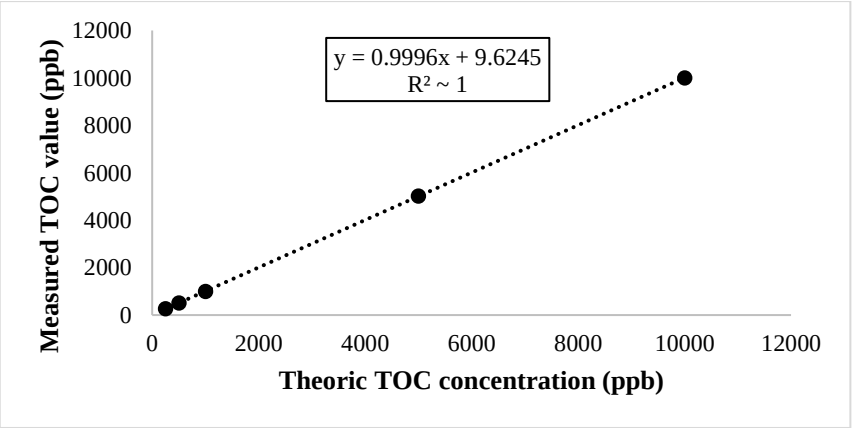


Figure 4. 12 Linear regression for the Offline TOC equipment.

4.5.2. Precision

4.5.2.1. Repeatability

The certified TOC standard at 500 ppb concentration was analyzed with the offline equipment at six times. The results in Table 4. 25 show that the overall RSD (%) for the studied concentration was less than 5%.

Table 4. 25 Repeatability evaluation for the 500 ppb TOC certified standard with the Offline TOC equipment.

Replicate	Measured ppb TOC M9 Offline
1	513
2	515
3	517
4	518
5	520
6	526
$\bar{x}$	**518**
s	**4.54**
RSD (%)	**0.875**
Blank (ppb)	**14.7**
R (%)	**101**

4.5.2.2. Intermediate Precision

The influence of the equipment on the certified TOC standard at 500 ppb concentration was evaluated. The values obtained for both equipment are shown in Table 4. 25 and Table 4. 26.

Table 4. 26 Measurements for the 500 ppb TOC certified standard with the Online TOC equipment.

Replicate	Measured ppb TOC M9 Online
1	515
2	519
3	523
4	517
5	521
6	515
$\bar{x}$	**518**
s	**3.26**
RSD (%)	**0.630**
R (%)	**101**

The one-way ANOVA test result is shown in Table 4. 27.

Table 4. 27 One-way ANOVA statistical results considering the factor equipment for the 500 ppb certified TOC standard.

Factor	*SS*	*DF*	*MS*	$F_{Calculated}$	*p-value*	$F_{Tabulated}$
Between groups	0.0833	1	0.0833	0.00534	0.943	4.96
Within groups	156	10	15.6			
Total	156	11				

Moreover, the statistical parameter evaluated (one-way ANOVA test) demonstrate the precision parameter including the fifth certified TOC standards involved for linearity, including the blank standard. See Table 4. 28.

Table 4. 28 One-way ANOVA statistical results considering the factor equipment for the linearity certified TOC standards.

Factor	*SS*	*DF*	*MS*	$F_{Calculated}$	*p-value*	$F_{Tabulated}$
Between groups	$3*10^4$	1	$3*10^4$	0.00212	0.964	4.96
Within groups	$16*10^8$	10	$16*10^8$			
Total	$16*10^8$	11				

4.5.3. Accuracy

The same certified TOC standards involved in the precision evaluation were used for the accuracy test based on the statistical *t*-test evaluation. The obtained results are shown in Table 4. 29.

Table 4. 29 Statistical t-test for equipment measurement comparison considering 500 ppb certified TOC standard.

Parameter	**Measured TOC M9 Online**	**Measured TOC M9 Offline**
$\bar{x}$	518	518
Number of measurements	6	6
Hypothetical mean difference	0	
Degrees of freedom	5	
$t_{Calculated}$	0.0683	
p-value	0.948	
$t_{Tabulated}$ (two tiles)	2.57	

In addition, no significant differences were detected, based on the t-test for the certified TOC standards involved in the linearity evaluation. See Table 4. 30.

Table 4. 30 Statistical t-test for equipment measurement comparison considering linearity certified TOC standards.

Parameter	**Measured TOC M9 Online**	**Measured TOC M9 Offline**
$\bar{x}$	2800	2910
Number of measurements	6	6
Hypothetical mean difference	0	
Degrees of freedom	5	
$t_{Calculated}$	1.67	
p-value	0.155	
$t_{Tabulated}$ (two tiles)	2.57	

4.6. Conclusions

This study was initially planned for an agent from the corticosteroid family. It was initially selected because the cleaning of the equipment involved during its manufacture is critical and necessary after each of the MB. However, different experiments demonstrated, on the one hand, the low solubility of the agent in water, being the optimal solvent for the use of the TOC technique, and, on the other hand, the lack ability of the technique, due to the experimental procedure required for the preparation of the samples, to reproducibly quantify the calculated TOC acceptable lower limit value.

The summarized results demonstrate the ability to correctly detect the presence of DKP-Tl after the CV process with a reliable and rapid TOC method for routine residual analysis involving AQbD. The FMEA strategy was successfully involved to detect the optimized method parameters, including the target agent solubility and stability in the solvent solution, the blank influence on the evaluated target agent concentration, and the extraction condition optimization for swabbing. After method optimization, all the requirements for method validation were accomplished including selectivity, linearity, accuracy, LoD, LoQ, precision, and robustness. Additionally, the compound was successfully recovered from stainless steel, plastic, and epoxy resin for rinse sampling and swabbing techniques during TOC analysis for the previously calculated worst-case scenarios.

During external operations, the proper running of the TOC equipment may be altered, generating incorrect TOC, IC, and/or TC results. Involving risk assessment, the criticalities during TOC measurements were detected, and specific mitigation actions were implemented to ensure the quality of the TOC measurements.

Finally, the transferability between TOC equipment was demonstrated in terms of linearity, precision, and accuracy. As a result, the validated DKP-T TOC method could be implemented indistinctly.

4.7. Bibliography

(1) Parenteral Drug Association. *Points to Consider for Cleaning Validation. N^{o} 29.*; PDA, 2012.

(2) European Medicines Agency. Annex 15: Qualification and Validation. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 30, 2015, pp 1–16.

(3) Reid, G. L.; Morgado Kimber Barnett, J.; Harrington Jian Wang, B.; Harwood David Fortin Pfizer, J. Analytical Quality by Design (AQbD) in Pharmaceutical Development. *Am Pharm Rev* **2013**.

(4) Peraman, R.; Bhadraya, K.; Padmanabha Reddy, Y. Analytical Quality by Design: A Tool for Regulatory Flexibility and Robust Analytics. *Int J Anal Chem* **2015**, *2015*.

(5) Food and Drug Administration. Guide to Inspections Validation of Cleaning Processes. 1993, pp 1–6.

(6) World Health Organization. *Annex 4. Supplementary Guidelines on Good Manufacturing Practices: Validation*; 2006.

(7) Active Pharmaceutical Ingredients Committee. Guidance on Aspects of Cleaning Validation in Active Pharmaceutical Ingredient Plants. May 2014.

(8) European Medicines Agency. Guideline on Setting Health Based Exposure Limits for Use in Risk Identification in the Manufacture of Different Medicinal Products in Shared Facilities. London, UK, November 20, 2014, pp 1–11. www.ema.europa.eu/contact.

(9) Suez Water Technologies. *Total Organic Carbon for Cleaning Validation Programs; Laboratory, Online & at-Line Analysis.*; American Pharmaceutical Review, Ed.; 2020.

(10) European Pharmacopoeia. 2.2.44. TOC in Water for Pharmaceutical Use. *Ph. Eur. 11th*. Strasbourg, France January 2023, pp 85–86.

(11) The United States Pharmacopoeia. 643. Total Organic Carbon; 2021; pp 266–267.

(12) The Ministry of Health, L. and W. Quality Control of Water for Pharmaceutical Use. In *The Japanese Pharmacopoeia*; 2021; pp 1736–1740.

(13) European Medicines Agency. ICH Guideline Q9 on Quality Risk Management. *EMA/CHMP/ICH/24235/2006*. London, UK, September 2015.

(14) Dr. Christopher Burgess. Laboratory Data Management Guidance: Analytical Procedure Lifecycle Management (APLM) Guideline. ECA Analytical Quality Control Group June 28, 2018, pp 1–80.

(15) European Medicines Agency. ICH Q2 (R1). Validation of Analytical Procedures: Text and Methodology. *CPMP/ICH/381/95*. London, UK, June 1995.

(16) Bisutti, I.; Hilke, I.; Raessler, M. Determination of Total Organic Carbon- An Overview of Current Methods. *TrAC - Trends in Analytical Chemistry* **2004**, *23* (10–11), 716–726.

(17) European Medicines Agency. ICH Guideline Q2(R2) on Validation of Analytical Procedures. *EMA/CHMP/ICH/82072/2006*. London, UK, March 31, 2022.

(18) European Medicines Agency. Chapter 6: Quality Control. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 28, 2014, pp 1–8.

(19) GE Power Water & process Technologies. *Combined TOC and Conductivity Method Transfer Validation Support Package*; 2016.

5. STATISTICS APPLIED TO STABILITY STUDIES ACROSS THE DRUG PRODUCT LIFECYCLE

Stability studies ensure the quality, safety, and efficacy of pharmaceuticals[1]. The specific period through which the same characteristics and properties of a drug product are retained after having been manufactured in a specific container or closure, variating environmental factors using qualified and calibrated climatic chambers, is carried out in a continuous program[2]. This concept is technically expressed as the shelf-life or commonly as the expiration period under specific storage conditions. This parameter is a CQA for all dosage forms contemplated as a requisite for final product approval.

The stability of the pharmaceutical drug depends on [2]:

- The environmental factors including: light, room temperature, and humidity.
- Factors related to the product, including its composition and the dosage form, the physic-chemical properties of the API, and the excipients.
- The packaging material characteristics.
- The pharmaceutical manufacturing processes.
- The character of the container or closure.

Those studies have to be developed following the issued guidelines. In the 1980s, the first guidelines were issued. Later, the ICH harmonized guidelines for marketing and registering products in other countries. The guidance documents detailed the different test conditions and requirements for formulations, excipients, APIs, or pharmaceuticals. Those are detailed in Table 5. 1.

Table 5. 1 ICH guidelines related to stability testing.[3–11]

Code	Title
ICH Q1A(R2)	*Scientific Guideline. Stability Testing of New Drug Substances and Products*
ICH Q1B	*Scientific Guideline. Stability Testing: Photostability Testing of New Drug Substances and Products*
ICH Q1C	*Scientific Guideline. Stability testing of New Dosage Forms*
ICH Q1D	*Scientific Guideline. Bracketing and Matrixing Designs for Stability Testing of Drug Substances and Products*
ICH Q1E	*Scientific Guideline. Evaluation of Stability Data*
ICH Q1F	*Scientific Guideline. Stability Data Package for Registration Applications in Climatic Zones III and IV*
ICH Q5C	*Scientific Guideline. Stability testing of Biotechnological/Biological Products*

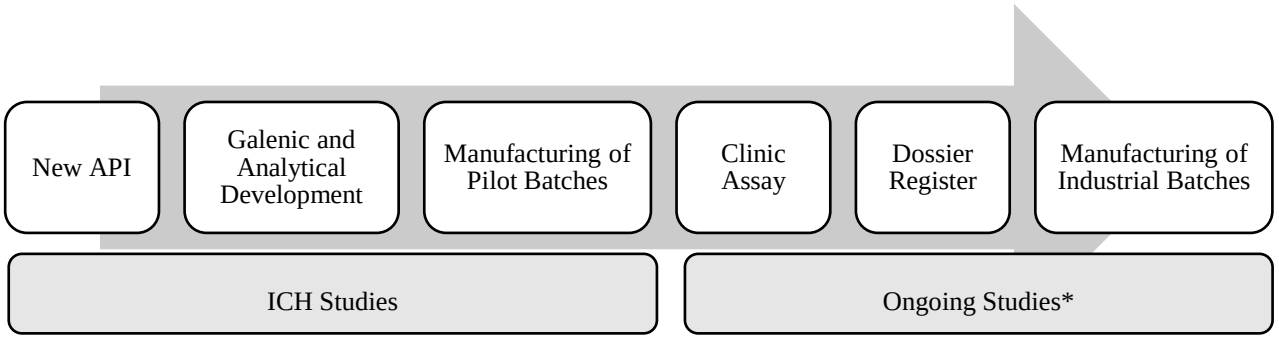
In addition, at the end of the 19th century, the FDA introduced the 21 CFR Part 11[12] standard based on the cGMP, establishing requirements related to its stability tests and its expiration date[13].

Moreover, the Committee for Proprietary Medicinal Products (from now on CPMP) with the EMA issued the guidelines detailed in Table 5. 2, assisting with the authorization in the European Union of pharmaceuticals.

Table 5. 2 Guidance on stability testing by EMEA.[4,14–19]

Title	Code	Current version (year)
Note for Guidance on Maximum Shelf-life for Sterile Products for Human use After First Opening or Following Reconstruction	CPMP/QWP/159/96	1998
Note for Guidance on Start of shelf-life of the Finished Dosage Form	CPMP/QWP/072/96	2001
Note for Guidance on In-Use Stability Testing of Human Medicinal Products	CPMP/QWP/2934/99	2001
Guideline on Stability Testing of Existing Active Substances and Related Finished Products	CPMP/QWP/122/02/Rev.1 Corr.	2003
Guideline on Declaration of Storage Conditions	CPMP/QWP/609/96/Rev.2	2007
Guideline on Stability Testing for Applications for Variations to a Marketing Authorization	EMA/CHMP/CVMP/QWP/44107 1/2011-Rev.2	2014

The different stability programs involved during the manufacturing processes are summarized in Figure 5. 1.



*Ongoing studies: stability studies developed after marketing authorization to detect a different stability outcome[20].

Figure 5. 1 Stability programs involved during manufacturing pharmaceutical processes.

The different stability parameters are reviewed in Table 5. 3.

Table 5. 3 Stability parameters details for the general case[22,1].

Climate Zone	Study Conditions	Storage Condition	Period Covered (month)
I. Temperate	Long-term*	21°C±2°C and 45% RH±5% RH	12
II. Subtropical and Mediterranean	Long-term	25°C±2°C and 60% RH±5% RH	12
III. Hot and dry	Long-term	30°C±2°C/35% RH±5% RH	12
	Intermediate	30°C±2°C/65% RH±5% RH	6
	Accelerated**	40°C±2°C/25% RH±5% RH	6
IVa. Hot and humid	Long-term	30°C±2°C/65% RH±5% RH	12
	Accelerated	40°C±2°C/75% RH±5% RH	6
IVb. Hot and higher humid	Long-term	30°C±2°C/75% RH±5% RH	12
	Accelerated	40°C±2°C/75% RH±5% RH	6
Refrigerated	Long-term	5ºC±3ºC	12
	Accelerated	25°C±2°C and 60% RH±5% RH	6
Frozen	Long-term	-20ºC±5ºC	12
	Accelerated	5ºC±3ºC	6

*Long-term: real stability study. A study is carried out during the proposed efficacy period and under the conditions determined according to the nature of the product.

**Accelerated: stability study involving severe storage conditions. It is laid out to cause physical changes or the product's chemical degradation by applying severe storage conditions.

The degradation time of a pharmaceutical regarding the long-term stability studies, including multiple MB, allows the calculation of its shelf-life. The obtained data can be statistically analyzed independently for each storage condition[5].

The strategy followed is detailed in Figure 5. 2.

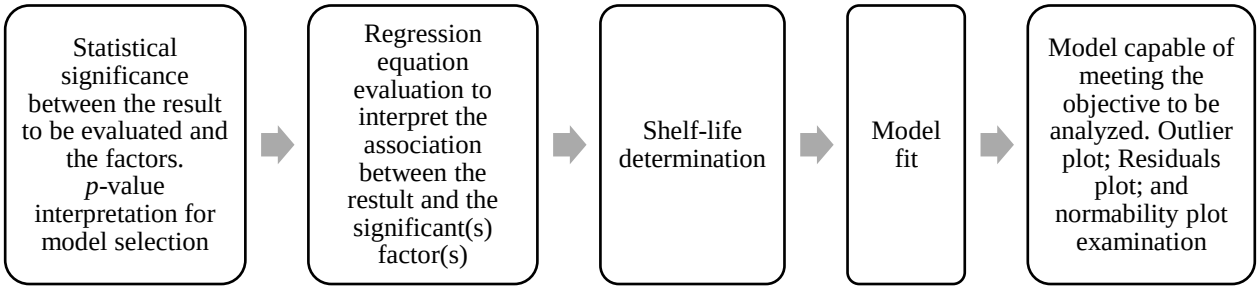


Figure 5. 2 The strategy followed for the statistical analysis of the stability study.

This chapter aims to calculate the shelf-life of two different oral dosage forms based on a multiple MBs approach:

- Liquid dosage form (multivitamin mixture): ensure the CQA in the percentage of total impurities, being the specification value at the end of the product's shelf-life, ≤1.5%. See chapter 5.2.1.
- Solid dosage form (coated tablet): ensure the CQA assay of the APIs that compound the product (in percent), with its specification value, at the end of the product's shelf-life, ≥90%. See chapter 5.2.2.

5.1. Statistics

The software involved in the statistics calculations was Minitab. The following sections described the statistic involved to achieve the shelf-life calculation for stability studies[5,22].

5.1.1. Method Calculation

Before method calculation, the analysis option for the stability study is involved considering the batches included in the study as a fixed or random factor. The fixed factor refers to those not randomly selected from the manufacturing process to carry out the shelf-life stability study. The random factor is the opposite, being a randomized representative sample. In addition, it must be established if there is a linear relationship between the value under study and time, with the kinetics of the degradation reaction of order 0 (involving a linear regression method based on the integration velocity equation). See Equation 5. 1.

$$[Y]=[Y]_0-kt$$

Where:

k: kinetic constant.

t: time.

Y: evaluated response.

Y_0: evaluated response at time 0.

Equation 5. 1.

The shelf-life is calculated based on the percent response that is within the proposed specifications (commonly 50 percent selection) with a 95 % confidence.

The *F*-distribution is selected to test the different models and is detailed in Equation 5. 2.

$$\frac{DF\ for\ the\ evaluated\ term}{DF\ for\ the\ error}$$

DF: Degrees of freedom

Equation 5. 2.

The key result is assigned to the *p*-value and is proportional to the calculated *F*-distribution. See Equation 5. 3. If it is lower than the specified value (0.05), the evaluated result and the factor are statistically significant and the factor remains in the model. If it is higher, the factor is excluded considering the reverse reason.

p-value=1-*F*-distribution (for each term)

Equation 5. 3.

The rest of the sections include the statistical description considering the batches as a fixed factor due to all the stability studies included in a specific period of time being included for the shelf-life calculation.

5.1.1.1.Analysis of Variance for Fixed Batches

The ANOVA is evaluated for the selected significant factors[23].

The equations for the different statistics are presented in Table 5. 4.

Table 5. 4 ANOVA table.

Source	**DF**	**SS**	**MS**	**F-distribution**	***p*-value** [5]
Significant factor	*p*-1	$\sum(\bar{y}_i - \bar{y})^2$	$\frac{SS_{significant\ factor}}{p-1}$	$\frac{MS_{significant\ factor}}{MS_{error}}$	α
Error	*n-p*	$\sum(y_{ij} - \bar{y}_i)^2$	$\frac{SS_{error}}{n-p}$		
Total	*n*-1	$\sum(y_{ij} - \bar{y})^2$	$\frac{SS_{total}}{n-1}$		

n: observations included in the model.

p: coefficient's number in the model.

X: stability design values.

$\bar{y}$: mean response.

$\bar{y}_i$: adjusted response.

y_{ij}: observed response.

SS: sum of squares. $SS_{significant\ factor}$ means the quantity of variation explained by the significant terms in the model. SS_{error} means the quantity of variation that could not be explained by the predictors. SS_{total} means the total quantity of variability included in the model.

MS: mean squares; the variance related to the adjusted values.

The only factors included in the ANOVA statistics are the ones with a statistical significance *p*-value (selected as 0.25).

5.1.1.2.Regression Model Calculation for Fixed Batches

The general formula is detailed in Equation 5. 4.

$$Y=X\beta+\varepsilon$$

Where:

β: significant factor.

ε: normal random independent variables.

X: stability design values.

Equation 5. 4.

5.1.1.3.Model Fitness for Fixed Batches

The determination of the model fitness is based on the examination of two statistical parameters:

- R-square is defined as the amount of variation explained in the selected model. It is calculated as the difference between 1 and the ratio of the SS_{error} and SS_{total}.
- R-square adjusted is defined as the amount of variation explained in the selected model considering the number of responses. It is calculated as the difference between 1 and the ratio of the MS_{error} and MS_{total}.

Both parameters have to be as near as close to 100%.

5.1.2. Shelf-life Calculation for Fixed Batches

The studies presented include the calculation of the shelf-life based on a lower or upper specification limit at the end of the actually authorized shelf-life of the pharmaceuticals for the evaluated response.

Depending on the factors included in the model based on the method calculation, the statistic followed for the shelf-life calculation varies:

- Batch, time, and the interaction between them obtaining a regression equation independent for each included batch (different slope and intercept).
- Batch, and time, obtaining a regression equation independent for each included batch (the same slope and different intercept).
- Time, obtaining a unique regression equation for all included batches (the same slope and intercept).

In all scenarios, the formula described at Equation 5. 4 is detailed as shown in Equation 5. 5, adapted depending on which one was obtained during method calculation.

$$\bar{y}_{ij}=\hat{\beta}x_{ij}+\widehat{\alpha_i};$$

$$\hat{\beta}x_{ij}+\widehat{\alpha_i}=L \text{ or } \hat{\beta}x_{ij}+\widehat{\alpha_i}=U$$

Where:

$\widehat{\alpha_i}$: the intercept considering the *i* batches included in the model.

$\hat{\beta}$: slope considering the *i* batches included in the model.

i: a general index for the batch that the shelf-life is estimated.

j: a general index for the time.

L: the lower specification limit.

U: the upper specification limit.

$\bar{y}_{ij}$: fitted value.

Equation 5. 5.

The software evaluates two conditions to calculate a significant shelf-life:

- Whether the mean response is greater than the L or U at time 0.
- Whether the mean response decrease over time.

If both conditions are true, the shelf-life is calculated based on a quadratic equation where the composition of the terms is different according to the number of significant factors included in the model. See Equation 5. 6.

$$\text{Shelf-life} = \frac{-b-\sqrt{b^2-4ac}}{2a}$$

Where:

Batch, time, and batch*time interaction	**Batch*time**	**Time**
$a=\hat{\beta}_i^2-t^2_{cl,n-2l\sigma l+i,l+i}$	$a=\hat{\beta}^2-t^2_{cl,n-l-1\sigma l+1,l+1}$	$a=\hat{\beta}^2-t^2_{cl,n-2\sigma 2,2}$
$b=2[\hat{\beta}_i(\widehat{\alpha_i}-S)-t^2_{cl,n-2l\sigma i,l+i}]$	$b=2[\hat{\beta}(\widehat{\alpha_i}-S)-t^2_{cl,n-l-1\sigma i,l+1}]$	$b=2[\hat{\beta}(\hat{\alpha}-S)-t^2_{cl,n-2\sigma 1,2}]$
$c=(\widehat{\alpha_i}-S)^2-t^2_{cl,df\sigma j,j}$	$c=(\widehat{\alpha_i}-S)^2-t^2_{cl,n-l-1\sigma i,1}$	$c=(\hat{\alpha}-S)^2-t^2_{cl,n-2\sigma 1,1}$
σ: the variance. l: the number of levels in the batch factor. n: the total number of response values. S: L or U depending on the specification limit established. $t_{cl,df}$: the t statistic for the confidence interval (CI) and the degrees of freedom (df).		

Equation 5. 6.

5.1.3. Residuals Plot Interpretation

After concluding the study, different residual plots should be evaluated, including the normal probability plot, the histogram representation, the residuals vs. fits, and the residuals vs. time order for the model observations.

5.1.3.1. Normal Probability Plot

The normal distribution of the residuals has to be ensured by approximating the graph to a straight line. The key result is assigned to an accurate *p*-value higher than 0.05 (with 95 % of confidence).

5.1.3.2. Histogram Plot

The residual distribution of the included observations is presented in the histogram plot. Detection of outliers or skewed data can be evaluated.

5.1.3.3. Residuals vs. Fit

The randomized distribution across the central line for all the residuals of the fit values with the selection model has to be ensured. An outlier or nonconstant variance can be identified.

The poolability was demonstrated by a t-test with a confidence of 95% and a significance level of α=0.05. If the *p*-value >0.05 the null hypothesis is accepted assuming homogeneous coefficients.

5.1.3.4. Residuals vs. Time Order Observations

The residuals for the data collection in time order are presented. The independence between residuals must be ensured without any pattern or trend showing possible correlation.

5.2. Results

5.2.1. Liquid Dosage Form

The statistics calculations to determine the shelf-life of the multivitamin product subjected to the long-term stability study for the climatic zone IVa are compiled. The data included contains the storage condition in which an OOS value has been obtained for the CQA assay (in percent) of the APIs present in the product formulation, being ≥90% of the value of the lower specification limit at the end of the product's shelf-life currently authorized (24 months). The product formulation is detailed in Table 5. 5.

Table 5. 5 Multivitamin product API formulation.

Description	Composition (%, w/w)
Ascorbic Acid	7.00
Alpha-tocopherol acetate	1.15
Biotin	0.02
Cholecalciferol	1.73
Nicotinamide	1.44
Pyridoxine Hydrochloride	0.21
Retinol Palmitate	0.22
Riboflavin	0.34

The condition of the stability study and the packaging material studied are detailed in Table 5. 6.

Table 5. 6 Stability conditions for the evaluated pharmaceutical product.

Pharmaceutical specialty	Stability condition	Packaging material	Colour for the packaging material
Liquid multivitamin dosage form	30ºC±2ºC /65%±5% RH	Glass flask	Opaque

For the present study, the data was obtained from both types of stability studies the ICH and Ongoing. According to the internal criteria, for the ongoing stability studies, the frequency of batch introduction was 1 recently manufactured batch per current year.

The shelf-life calculation was developed achieving a linear model that represents the association among the time and the response variable, and the batch factor if considered. A full factorial design was involved (the month selected 0, 12, and 24 were evaluated for all included MBs). The results for each evaluated API are detailed in the following sections.

Table 5. 7 shows the results obtained for the shelf-life considering the storage condition of 30ºC±2ºC/ 65%±5% RH, based on the assay value as a percentage of the vitamins that have resulted in an OOS value at the end of the validity period of the multivitamin product, these being: Nicotinamide and Retinol Palmitate.

Table 5. 7 Calculated shelf-life for the set of MBs included in the stability study, for the environmental condition 30ºC±2ºC /65%±5% RH, considering the assay (%) as a critical value.

API evaluated	MBs included	Kinetic order for the degradation reaction*	Shelf-life calculated (adjusted time-month)
Nicotinamide	9	0	20
Retinol Palmitate	9	0	17
*Linear relation between the evaluated attributed and the factor time.			

As a result, the shelf-life should be shortened to 17 months to avoid the presence of assay-related OOS for both evaluated APIs.

5.2.1.1.Shelf-life Calculation Based on Nicotinamide Assay Content

The distribution of the data was initially evaluated, being necessary to transform them by means of a mathematical calculation based on the Johnson Transformation since it did not comply with a normal distribution. See Figure 5. 3.

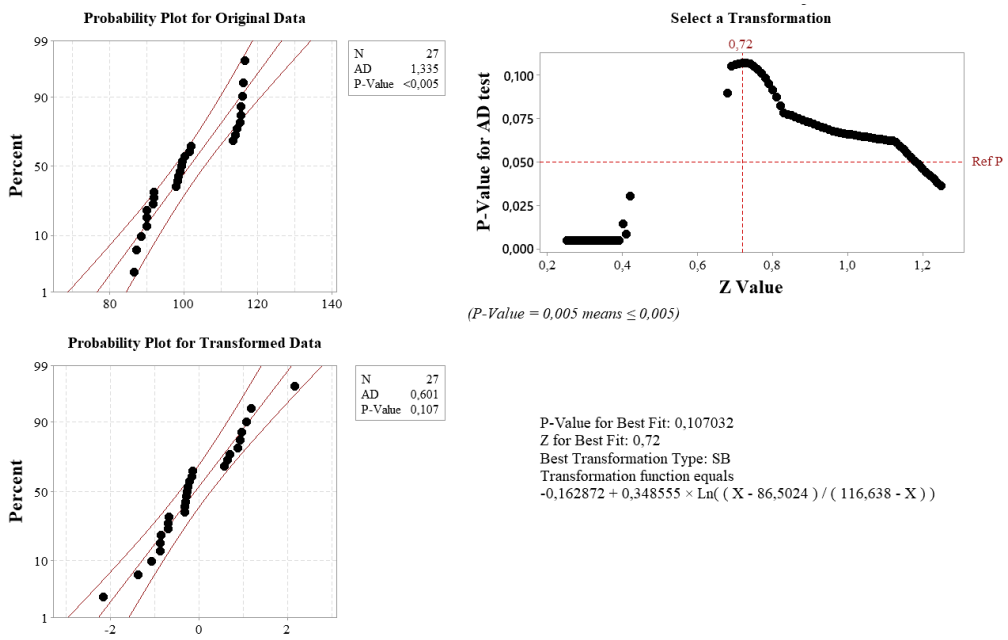


Figure 5. 3 Johnson mathematical transformation for Nicotinamide assay (%) values.

A possible outlier value was observed. This was accepted as part of the set of data variability after performing a Grubbs statistical test obtaining a *p*-value>0.05 (0.302). See Figure 5. 4.

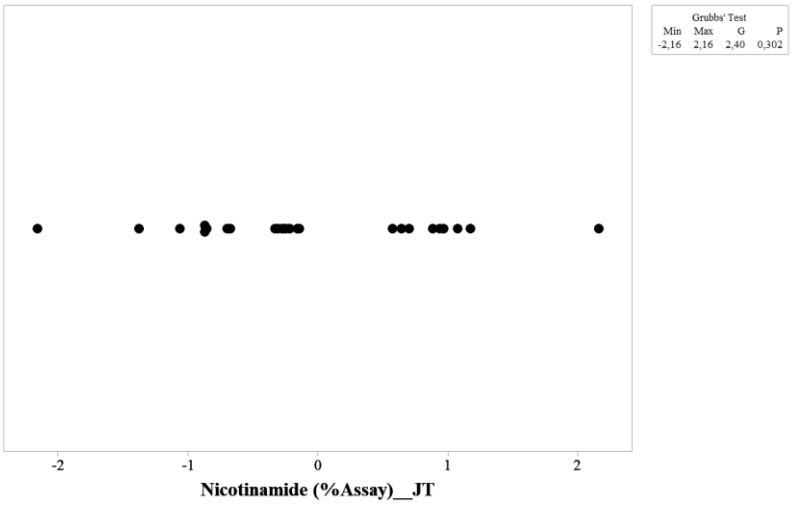


Figure 5. 4 Outlier plot based on Grubbs test for Nicotinamide assay.

Before model calculation, the factor batch was established as a fixer factor, because all the MBs summited to the stability study of interest were included. Furthermore, the kinetic order of the degradation equation was calculated being 0, obtaining a linear relationship between the CQA and the time. See Figure 5. 5.

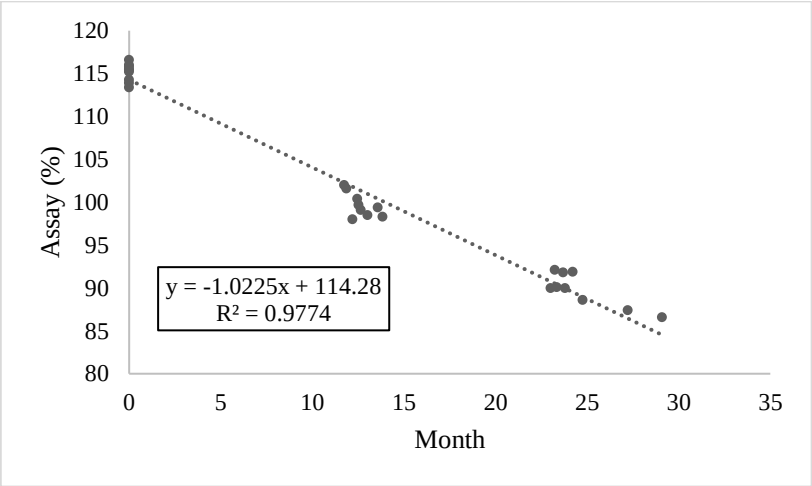


Figure 5. 5 Linear regression for Nicotinamide assay (%) results at the end of the authorized product's shelf-life.

The linear model includes only the moth as a factor, due to the association within the evaluated response (assay) and the factor was statistically significant. For the established significant level of 0.25[5], a *p*-value lower than this specified value was obtained (month factor *p*-value=0.000). In this scenario, the batch factor and the interaction between month and batch are not statistically significant by reverse reasoning (*p*-value batch=0.261, *p*-value month*batch=0.316).

As a result, only one regression equation for the set of MBs is included due to the unique significance of the month factor as shown in Equation 5. 7.

$$Nicotinamide\ Assay\ (\%) = 0.961 - 0.08442 * Month$$

Equation 5. 7.

The shelf-life period was calculated as 20.0325 months (20 months as adjusted time) guaranteeing, with 95% confidence, that 50% of the responses obtained are below the specified upper limit for the assay. See Figure 5. 6.

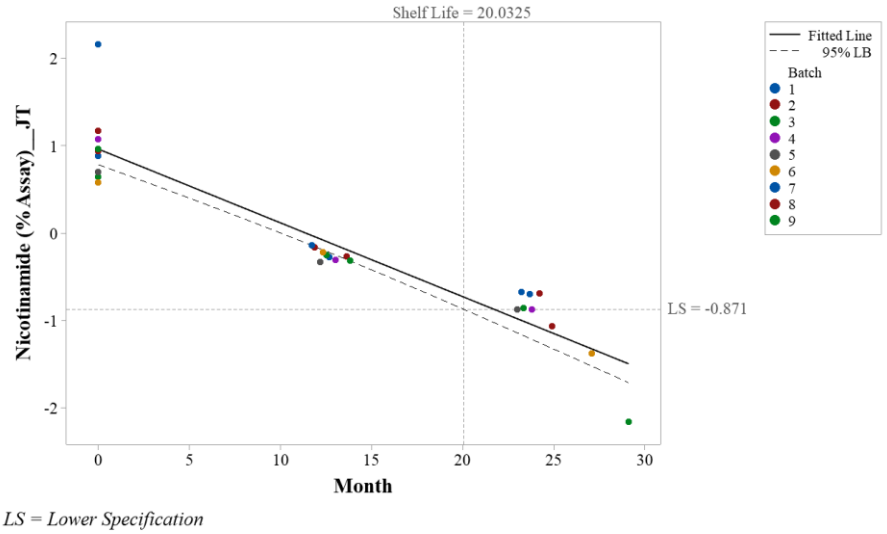


Figure 5. 6 Shelf-life plot for all batches included for Nicotinamide assay quantification. LS referred to the Johnson transformation results.

The determination coefficient and the adjusted (R^2: 87.17, R^2_{adj}: 86.65) were near 100, adjusting the data correctly to the presented model.

Moreover, the residuals are normally and randomly distributed across 0. There was no evidence to conclude that the mean differs from 0 at the 0.05 level of significance (*p*-value>0.05, for the calculated *t*-test (0.620). See Figure 5. 7.

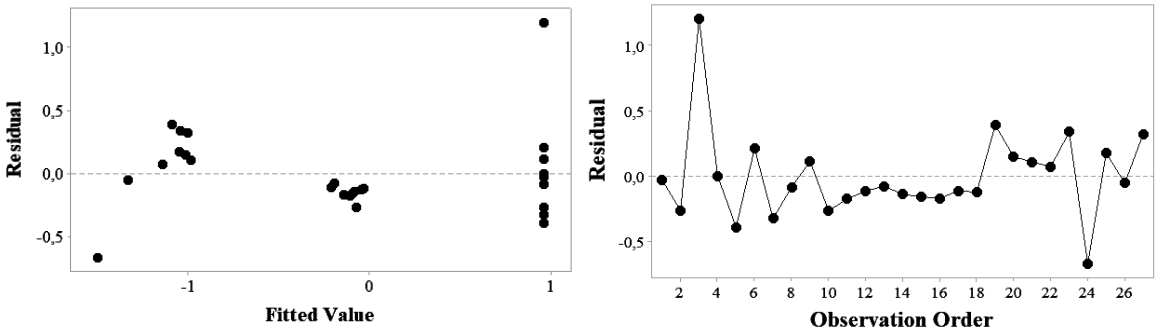


Figure 5. 7 Residuals plot against the fitted values and the observation in time order for Nicotinamide.

5.2.1.2.Shelf-life Calculation Based on Retinol Palmitate Assay Content

The distribution of the data was initially evaluated, being necessary to transform them by means of a mathematical calculation based on the Johnson Transformation since it did not comply with a normal distribution. See Figure 5. 8.

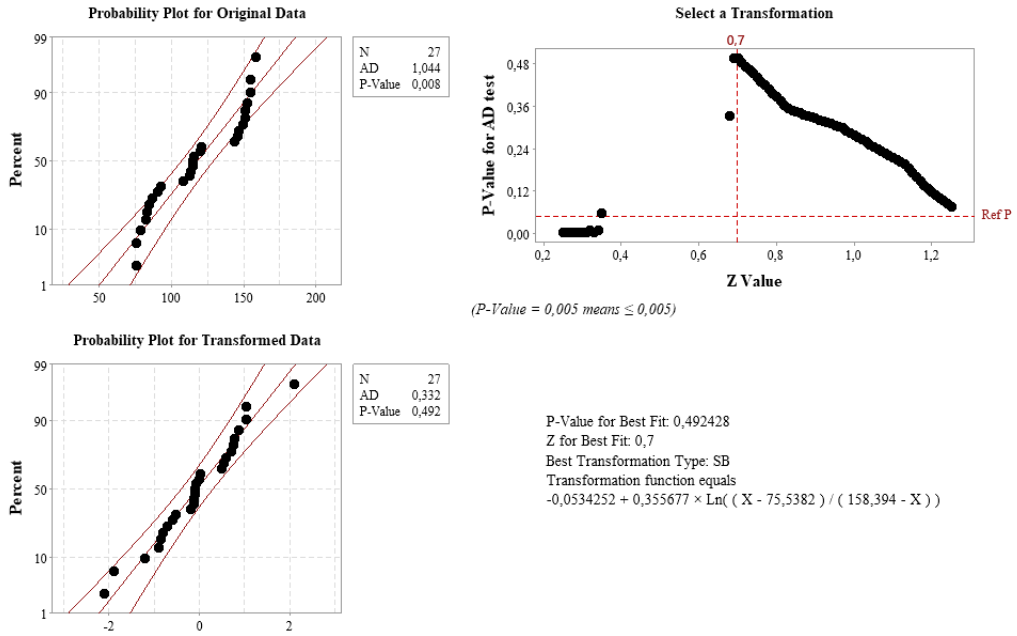


Figure 5. 8 Johnson mathematical transformation for Retinol Palmitate assay (%) values.

A possible outlier value was observed. This was accepted as part of the set of data variability after performing a Grubbs statistical test obtaining a *p*-value>0.05 (0.426). See Figure 5. 9.

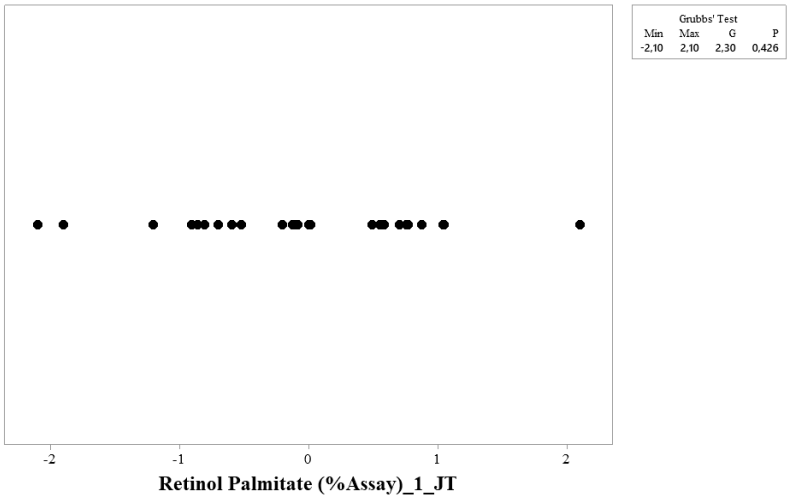


Figure 5. 9 Outlier plot based on Grubbs test for Retinol Palmitate assay.

Before model calculation, the factor batch was established as a fixer factor, because all the MBs summited to the stability study of interest were included. Furthermore, the kinetic order of the degradation equation was calculated being 0, obtaining a linear relationship between the CQA and the time. See Figure 5. 10.

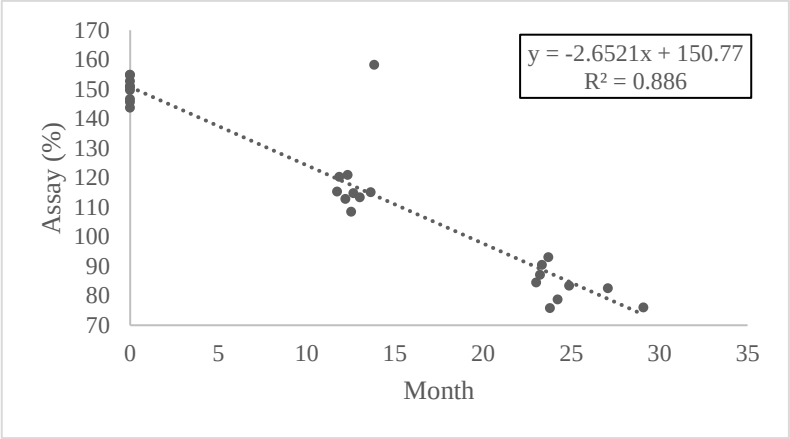


Figure 5. 10 Linear regression for Retinol Palmitate assay (%) results at the end of the authorized product shelf-life.

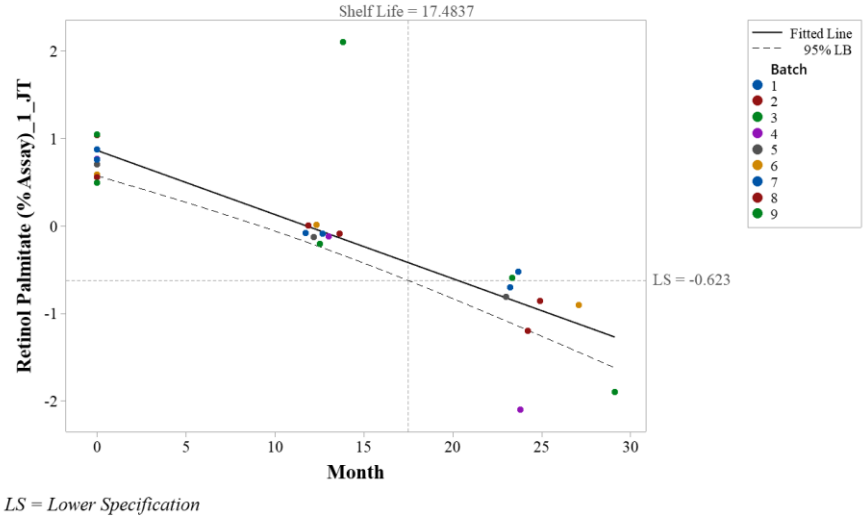
The linear model includes only the moth as a factor, due to the association within the evaluated response (assay) and the factor was statistically significant. For the established significant level of 0.25^{5}, a *p*-value lower than this specified value was obtained (month *p*-value=0.001). In this scenario, the batch factor and the interaction between month and batch are not statistically significant by reverse reasoning (*p*-value batch=0.970, *p*-value month*batch=0.977).

As a result, only one regression equation for the set of MBs is included due to the unique significance of the month factor as shown in Equation 5. 8.

$$Retinol\ Palmitate\ Assay\ (\%) = 0.861 - 0.0732 * Month$$

Equation 5. 8.

The shelf-life period was calculated as 17.4837 months (17 months as adjusted time) guaranteeing, with 95% confidence, that 50% of the responses obtained are below the specified upper limit for the assay. See Figure 5. 11.



LS = Lower Specification

Figure 5. 11 Shelf-life plot for all batches included for Retinol Palmitate assay (%) quantification. LS referred to the Johnson transformation results.

The coefficient of determination and the adjusted one (R^2: 65.81, R^2_{adj}: 64.44) were accepted despite presenting a certain deficiency in the fit of the data to the model presented. The normal distribution was only achieved by Johnson's mathematic transformation.

Moreover, the residuals are normally and randomly distributed across 0. There was no evidence to conclude that the mean differs from 0 at the 0.05 level of significance (*p*-value>0.05, for the calculated *t*-test (0.780). See Figure 5. 12.

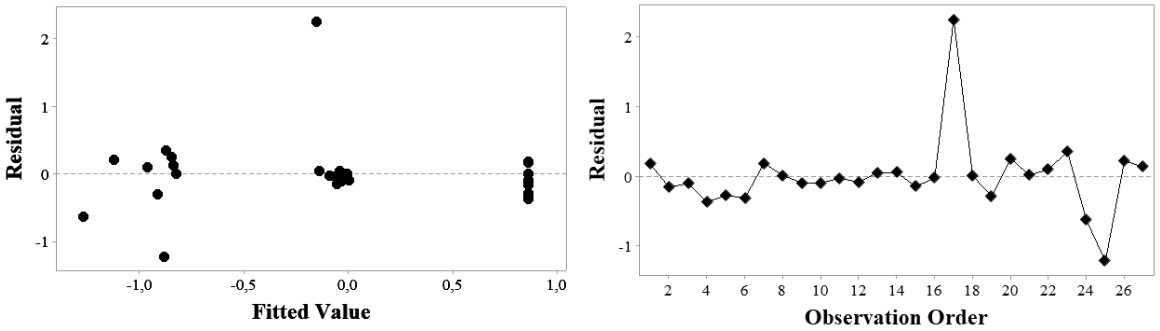


Figure 5. 12 Residuals plot against the fitted value and the observation in time order for Retinol Palmitate.

5.2.2. Solid Dosage Form

The statistics calculations to determine the shelf-life of the coated tablet product subjected to the long-term stability study for the stability condition are compiled. The stability conditions included were those in which an OOS value was obtained for the CQA total impurities (in percentage), being ≤1.5 % of the value of the upper specification limit evaluated at the end of the product's shelf-life currently authorized (18 or 24 months depending on the product specialty).

The condition of the stability study and the packaging materials studied are detailed in Table 5. 8.

Table 5. 8 Stability condition for the evaluated pharmaceutical product.

Stability conditions	Packaging material	Colour for the packaging material
30ºC±2ºC /65%±5% RH	PVC+PVdC120g/Alu	Transparent
		White
	PVC/Alu	Transparent
PCV: polyvinyl chloride. PVdC: poly-Vinylidene chloride. Alu: Aluminium.		

For the present study, the data was obtained from both types of stability studies, the ICH and ongoing, including MB for different product specialties. According to internal criteria, for the ongoing stability studies, the frequency of batch introduction was 1, being a recently manufactured batch per current year.

The shelf-life calculation was developed achieving a linear model that represents the association among the time and the response variable, and the batch factor if considered. A full factorial design was involved (the month selected 0, 12, and 24 were evaluated for all included MBs). The results for each evaluated packaging material are detailed below.

The results obtained for the shelf-life calculation including the storage condition of 30ºC±2ºC/ 65%±5% RH, based on the total impurities value as a percentage that has resulted in an OOS value at the end of the validity period of the coated tablet product, are shown in Table 5. 9. The number of MB and the kinetic order for the degradation reaction are included.

Table 5. 9 Calculated shelf-life for the set of MBs included in the stability study for the environmental condition 30ºC±2ºC /65%±5% RH, considering the total impurities (%) as a critical value.

30ºC±2ºC /65%±5% RH				
Packaging material	**Colour for the packaging material**	**MB included**	**Kinetic order for the degradation reaction**	**Shelf-life calculation (Adjusted time-month)**
PVC+PVdC120g/Alu-Transparent	Transparent	8	0	21
PVC+PVdC120g/Alu-White	White	6	0	25
PVC/Alu	Transparent	3	0	19

As a result, for the stability conditions evaluated, specific packaging materials should be selected to avoid the presence of total impurities-related OOS at the end of the product's shelf-life currently authorized. See Table 5. 10.

Table 5. 10 Summary of the set of optimal packaging materials for the authorized shelf-life of 18 and 24 months.

Packaging material	**30ºC ± 2ºC /65% ± 5% RH**	
	18 months	24 months
PVC+PVdC120g/Alu-Transparent	X	*
PVC+PVdC120g/Alu-White	X	X
PVC/Alu- Transparent	X	*
X: optimal packaging materials for the authorized shelf-life. *: requires the reduction of the shelf-life of the evaluated specialty.		

5.2.2.1.Shelf-life Calculation for the Stability Condition 30ºC ± 2ºC /65% ± 5% RH

- PVC+PVdC120g/Alu-Transparent

The plot of the residuals against their expected values was normally distributed, obtaining a *p*-value>0.05 (0.058). See Figure 5. 13.

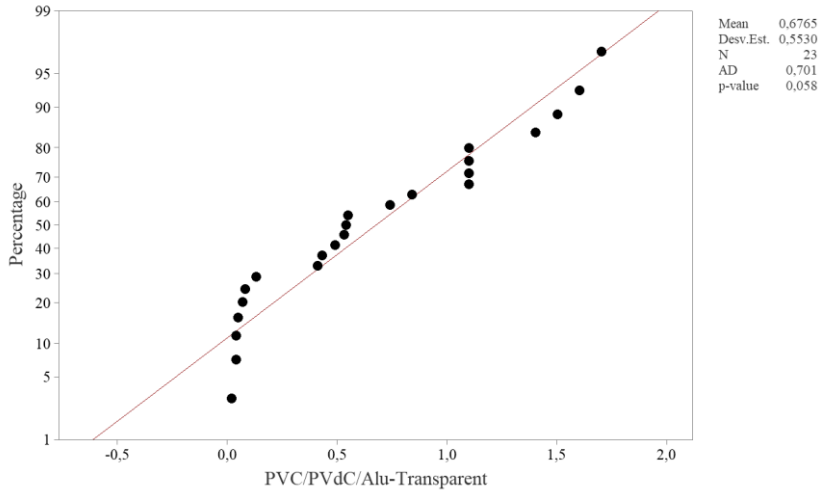


Figure 5. 13 Probability plot for total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

No outliers were detected. All data included was accepted as part of the set of data variability after performing a Grubbs statistical test obtaining a *p*-value>0.05 (1.0). See Figure 5. 14.

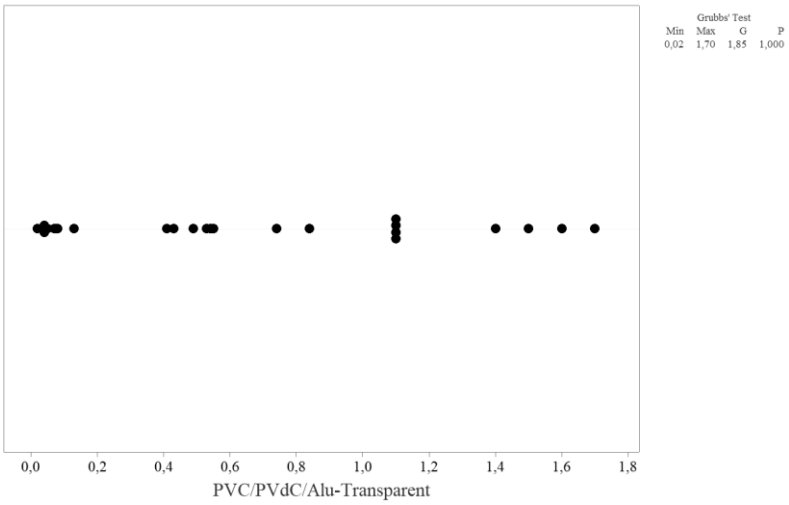


Figure 5. 14 Outlier plot based on Grubbs test or total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

Before model calculation, the factor batch was established as a fixer factor, because all the MBs summited to the stability study of interest were included. Furthermore, the kinetic order of the degradation equation was calculated being 0, obtaining a linear relationship between the CQA and the time. In this example, the degradation of the total impurities (%) implies an increase in the value, leading to the obtaining of a positive slope in the regression line. See Figure 5. 15.

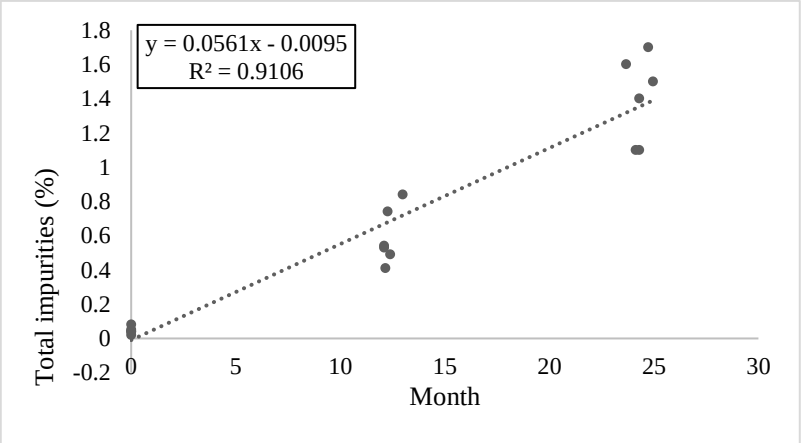


Figure 5. 15 Linear regression for total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

The linear model includes month and batch as factors, due to the association within the evaluated response (total impurities (%)), and the factors were statistically significant. For the established significant level of 0.25[5], a *p*-value lower than this specified value was obtained (month *p*-value=0.000 and batch *p*-value=0.207). In this scenario, the interaction between month and batch was not statistically significant by reverse reasoning (*p*-value month*batch=0.530).

As a result, there was one regression equation for each included MB as is shown in Equation 5. 9.

Batch	
1	*PVC/PVdC/Alu-Transparent= -0,0030 + 0,05746 Month*
2	*PVC/PVdC/Alu-Transparent= -0,0884 + 0,05746 Month*
4	*PVC/PVdC/Alu-Transparent= -0,1562 + 0,05746 Month*
5	*PVC/PVdC/Alu-Transparent= 0,1377 + 0,05746 Month*
6	*PVC/PVdC/Alu-Transparent= 0,1084 + 0,05746 Month*
7	*PVC/PVdC/Alu-Transparent= 0,0972 + 0,05746 Month*
8	*PVC/PVdC/Alu-Transparent= 0,0049 + 0,05746 Month*

Equation 5. 9.

The shelf-life period was 21.1906 months (21 months as adjusted time) guaranteeing, with 95% confidence, that 50% of the responses obtained are below the specified upper limit for the assay. See Figure 5. 16.

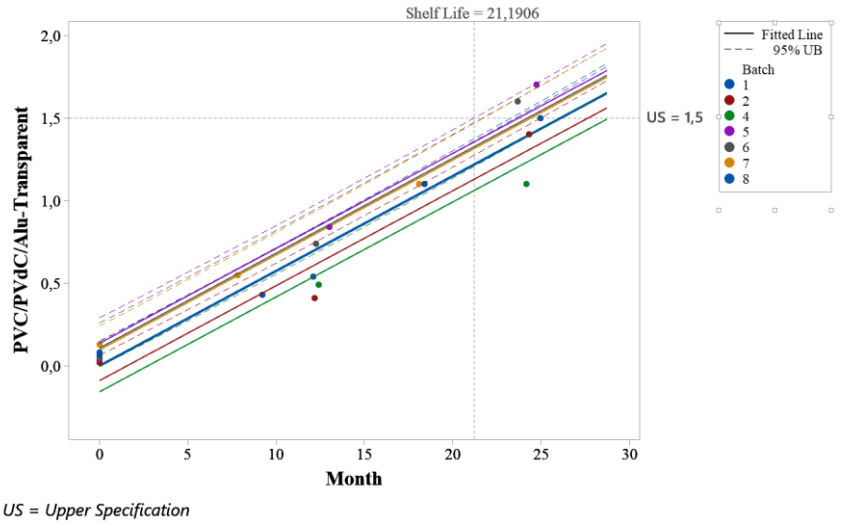


Figure 5. 16 Shel-life calculation based on the total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

The determination coefficient and the adjusted (R^2: 96.45, R^2_{adj}: 94.54) were near 100, adjusting the data correctly to the presented model.

Moreover, the residuals are normally and randomly distributed across 0. There was no evidence to conclude that the mean differs from 0 at the 0.05 level of significance (*p*-value>0.05, for the calculated *t*-test (0.100). See Figure 5. 17.

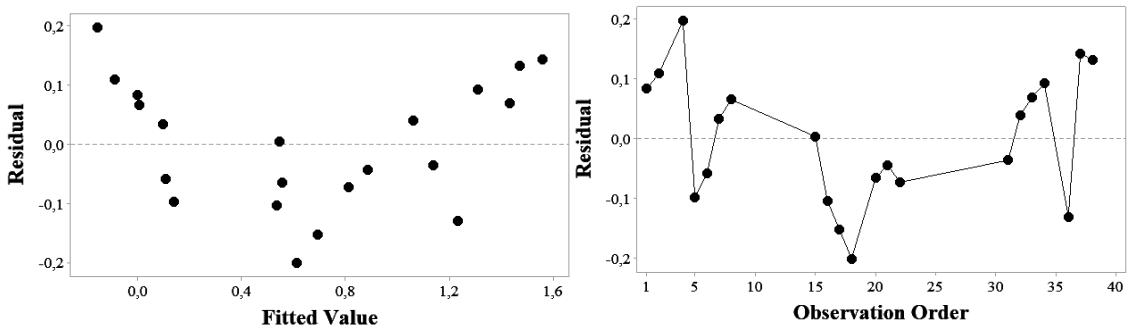


Figure 5. 17 Residuals plot against the fitted value and the observation in time order for the total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

- PVC+PVdC120g/Alu-White

The plot of the residuals against their expected values was normally distributed, obtaining a *p*-value>0.05 (0.081). See Figure 5. 18.

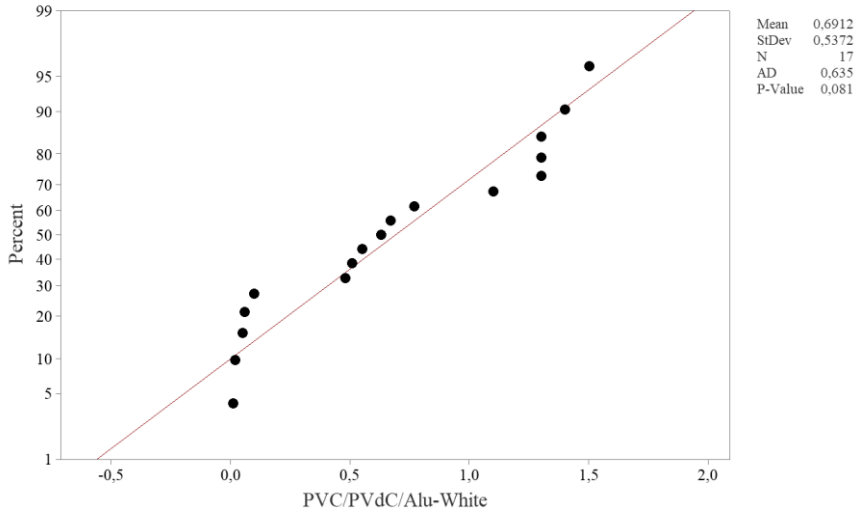


Figure 5. 18 Probability plot for total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-White packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

No outliers were detected. All data included was accepted as part of the set of data variability after performing a Grubbs statistical test obtaining a *p*-value>0.05 (1.0). See Figure 5. 19.

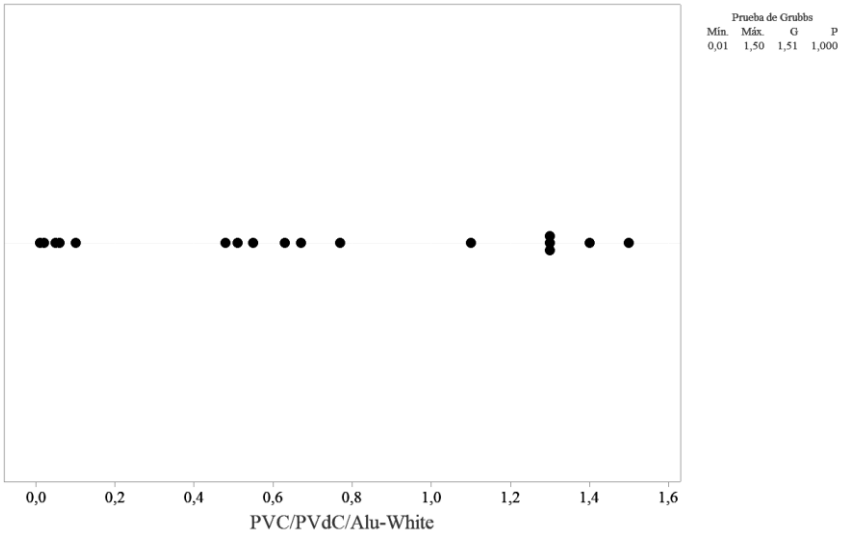


Figure 5. 19 Outlier plot based on Grubbs test or total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-White packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

Before model calculation, the factor batch was established as a fixer factor, because all the MBs summited to the stability study of interest were included. Furthermore, the kinetic order of the degradation equation was calculated being 0, obtaining a linear relationship between the CQA and the time. In this example, the degradation of the total impurities (%) implies an increase in the value, leading to the obtaining of a positive slope in the regression line. See Figure 5. 20.

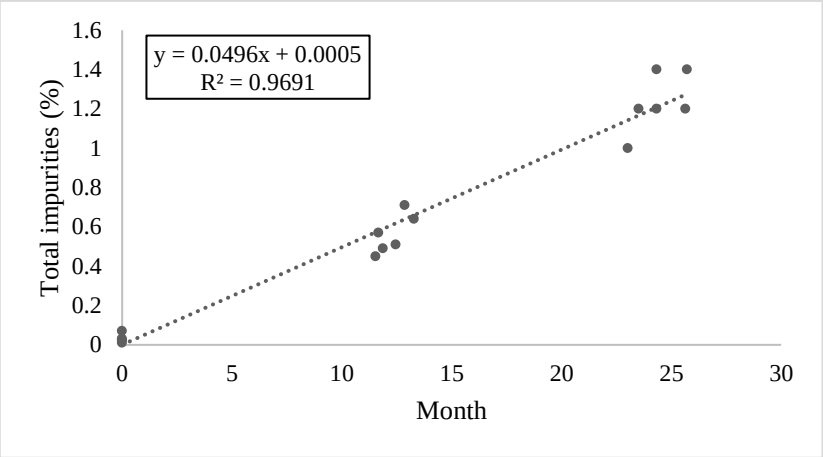


Figure 5. 20 Linear regression for total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Al-White packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

The linear model includes the moth and batch as factors, due to the association within the evaluated response (total impurities) and the factors were statistically significant. For the established significant level of 0.25[5], a *p*-value lower than this specified value was obtained (month *p*-value=0.000 and batch *p*-value=0.105). In this scenario, the interaction between month and batch was not statistically significant by reverse reasoning (*p*-value month*batch=0.643).

As a result, there was one regression equation for each included MB as is shown in Equation 5. 10.

Batch	
1	*PVC/PVdC/Alu-White= -0.0097 + 0.05219Month*
2	*PVC/PVdC/Alu-White= 0.0931 + 0.05219 Month*
4	*PVC/PVdC/Alu-White= 0.0966 + 0.05219Month*
5	*PVC/PVdC/Alu-White= -0.0504 + 0.05219Month*
6	*PVC/PVdC/Alu-White= -0.0459 + 0.05219Month*

Equation 5. 10.

The shelf-life period was 25.074 months (25 months as adjusted time) guaranteeing, with 95% confidence, that 50% of the responses obtained are below the specified upper limit for the assay. See Figure 5. 21.

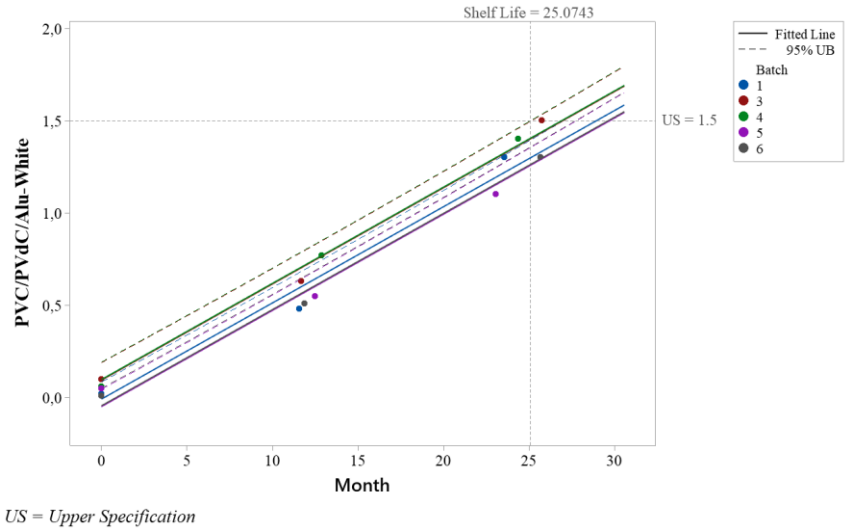


Figure 5. 21 Shel-life calculation based on the total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-White packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

The determination coefficient and the adjusted (R^2: 98.70, R^2_{adj}: 97.98) were near 100, adjusting the data correctly to the presented model.

Moreover, the residuals are normally and randomly distributed across 0. There was no evidence to conclude that the mean differs from 0 at the 0.05 level of significance (*p*-value>0.05, for the calculated *t*-test (0.354). See Figure 5. 22.

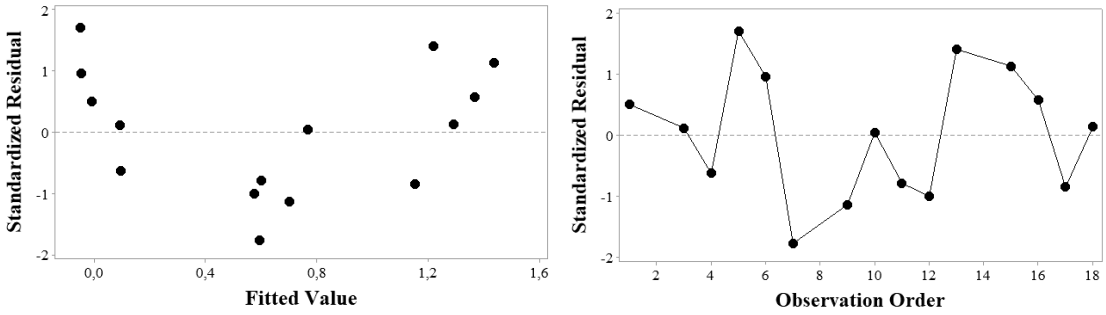


Figure 5. 22 Residuals plot against the fitted value and the observation in time order for the total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC+PVdC120g/Alu-White packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

- PVC-Alu-Transparent

The plot of the residuals against their expected values was normally distributed, obtaining a *p*-value>0.05 (0.097). See Figure 5. 23.

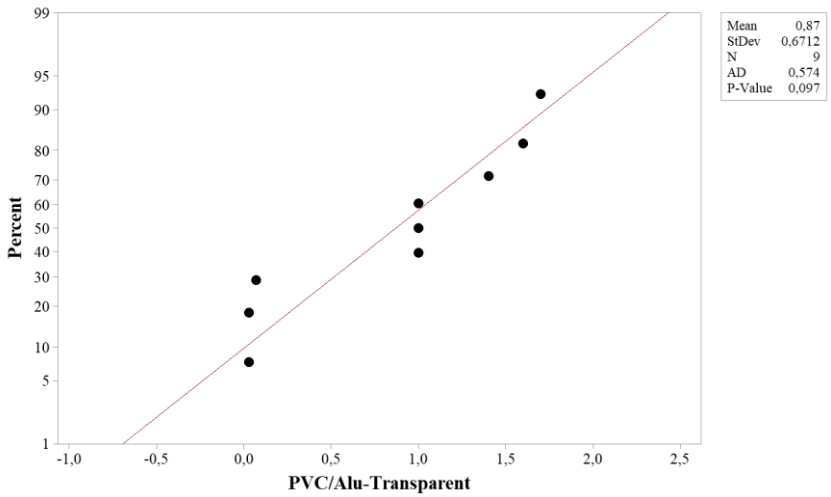


Figure 5. 23 Probability plot for total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC-Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

No outliers were detected. All data included was accepted as part of the set of data variability after performing a Grubbs statistical test obtaining a *p*-value>0.05 (1.0). See Figure 5. 24.

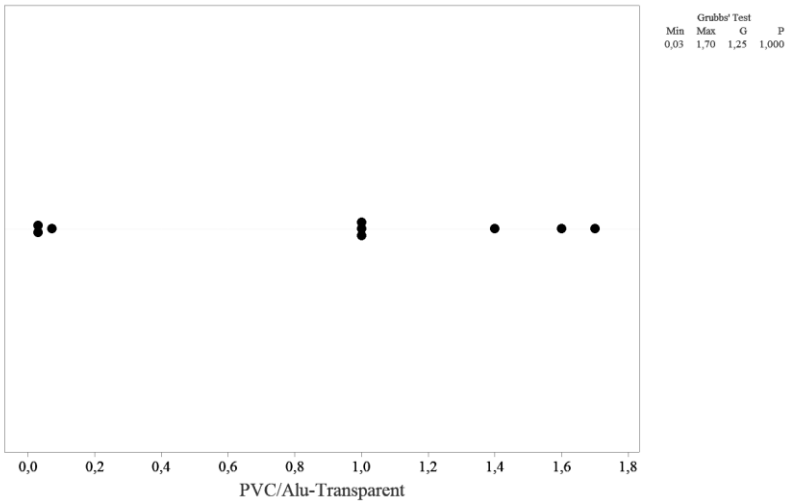


Figure 5. 24 Outlier plot based on Grubbs test or total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC-Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

Before model calculation, the factor batch was established as a fixer factor, because all the MBs summited to the stability study of interest were included. Furthermore, the kinetic order of the degradation equation was calculated being 0, obtaining a linear relationship between the CQA and the time. In this example, the degradation of the total impurities (%) implies an increase in the value, leading to the obtaining of a positive slope in the regression line. See Figure 5. 25.

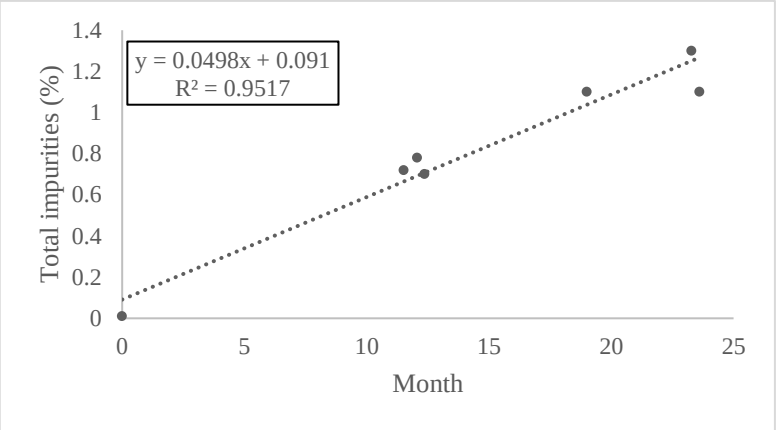


Figure 5. 25 Linear regression for total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC-Al-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

The linear model includes only the moth as a factor, due to the association within the evaluated response (total impurities), and the factor was statistically significant. For the established significant level of 0.25[5], a *p*-value lower than this specified value was obtained (month *p*-value=0.000). In this scenario, the factor batch and the interaction between month and batch were not statistically significant by reverse reasoning (*p*-value batch=0.591 and *p*-value month*batch=0.816).

As a result, there was one regression equation as is shown in Equation 5. 11.

$$PVC\text{-}Al=0.0848+0.06940*Month$$

Equation 5. 11.

The shelf-life period was 19.4207 months (19 months as adjusted time) guaranteeing, with 95% confidence, that 50% of the responses obtained are below the specified upper limit for the assay. See Figure 5. 26.

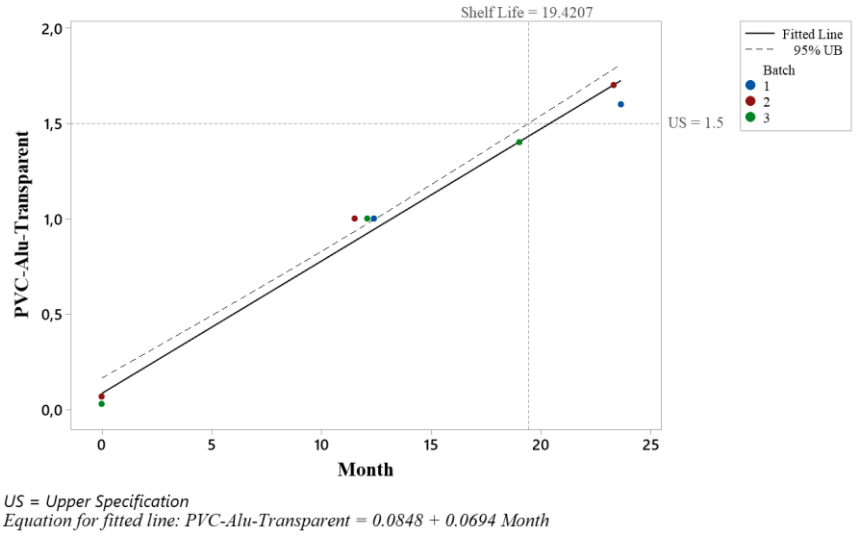


Figure 5. 26 Shel-life calculation based on the total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC-Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

The determination coefficient and the adjusted (R^2: 98.77, R^2_{adj}: 98.60) were near 100, adjusting the data correctly to the presented model.

Moreover, the residuals are normally and randomly distributed across 0. There was no evidence to conclude that the mean differs from 0 at the 0.05 level of significance (*p*-value>0.05, for the calculated *t*-test (0.100). See Figure 5. 27.

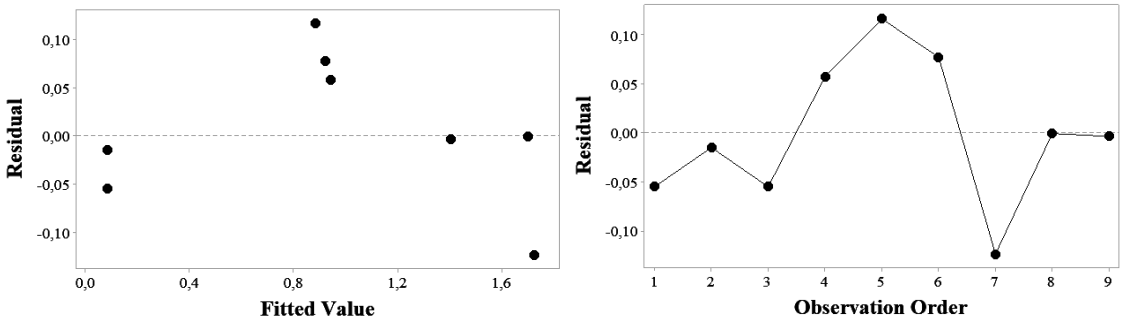


Figure 5. 27 Residuals plot against the fitted value and the observation in time order for the total impurities (%) results at the end of the authorized DKP-T product shelf-life for the PVC-Alu-Transparent packaging material at the stability conditions 30ºC ± 2ºC /65% ± 5% RH.

5.3. Conclusions

The applicability of the statistic to the calculation of the shelf-life has been demonstrated, guaranteeing the period in which the medicine meets the specifications for the CQAs of interest. A representative data set is necessary to ensure the evaluated response by a linear model regression.

Long-term stability studies were involved for MBs including the sampling points at 0, 12, and 24 months.

For both case studies, the proposed shelf life allowed the occurrence of OOS to be eliminated.

In the first case, there was necessary to reduce the product's shelf-life currently authorized to maintain the evaluated CQA inside the specification limits for all the APIs.

In the second case, only one packaging material was demonstrated to be feasible for the higher product's shelf-life currently authorized (24 months). However, all evaluated packaging materials were useful for 18 months of shelf-life authorized.

This strategy makes it possible to guarantee the quality of the pharmaceutical product during the calculated shelf-life.

5.4. Bibliography

(1) European Medicines Agency. Chapter 6: Quality Control. *Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use*. Brussels, Belgium, March 28, 2014, pp 1–8.

(2) World Health Organization. Annex 5. Guidelines for Stability Testing of Pharmaceutical Products Containing Well Established Drug Substances in Conventional Dosage Forms. 1996, pp 65–80.

(3) International Council for Harmonisation. *Quality Guidelines*. https://www.ich.org/page/quality-guidelines (accessed 2022-12-17).

(4) Narayan, S.; Manupriya, C. A Review on Stability Studies of Pharmaceutical Products. *International journal of Applied Pharmaceutical and Biological Research* **2017**, *2* (3), 67–75.

(5) European Medicines Agency. ICH Q1E Evaluation of Stability Data. London, UK, August 2003, pp 1–17.

(6) European Medicines Agency. ICH Q1A (R2) Stability Testing of New Drug Substances and Products Step 5. London, UK, August 2003, pp 1–20.

(7) European Medicines Agency. ICH Q1B Photostability Testing of New Active Substances and Medicinal Products Step 5. London, UK, January 1998, pp 1–9.

(8) European Medicines Agency. ICH Q1F Stability Data Package Registration Climatic Zones III and IV. London, UK, June 2006, pp 1–1.

(9) European Medicines Agency. ICH Q1C Stability Testing: Requirements for New Dosage Forms Step 5. London, UK, January 1998, pp 1–2.

(10) European Medicines Agency. ICH Q5C Quality of Biotechnological Products: Stability Testing of Biotechnological/Biological Products Step 5. London, UK, July 1996, pp 1–9.

(11) European Medicines Agency. ICH Q1D Bracketing and Matrixing Designs for Stability Testing of Drug Substances and Drug Products Step 5. London, UK, February 2002, pp 1–8.

(12) Food and Drug Administration. Part 11, Electronic Records; Electronic Signatures. August 2003, pp 1–12.

(13) Food and Drug Administration. Expiration Dating and Stability Testing for Human Drug Products. 2014.

(14) European Medicines Agency. Note for Guidance on Start of Shelf-Life of the Finiched Dosage Form. London, UK, May 31, 2001, pp 1–2.

(15) European Medicines Agency. Note for Guidance on In-Use Stability Testing of Human Medicinal Products. London, UK, March 1, 2001, pp 1–3.

(16) European Medicines Agency. Guideline on Declaration of Storage Conditions. London, UK, November 19, 2007, pp 1–4.

(17) European Medicines Agency. Guideline on Stability Testing: Stability Testing of Existing Active Substances and Related Finished Products. London, UK, December 17, 2003, pp 1–18.

(18) European Medicines Agency. Note for Guidance on Maximum Shelf-Life for Sterile Products for Human Use After First Opening or Following Reconstitution. London, UK, January 28, 1998, pp 1–2.

(19) Medicines Agency, E. Guideline on Stability Testing for Applications for Variations to a Marketing Authorisation. London, UK, March 21, 2014, pp 1–15. www.ema.europa.eu.

(20) World Health Organization. Annex 10. Stability Testing of Active Pharmaceutical Ingredients and Finished Pharmaceutical Products. 2009, pp 309–351.

(21) European Medicines Agency. VICH GL58 Stability Testing of New Veterinary Drug Substances and Medicinal Products in Climatic Zones III and IV. *EMA/CBMP/VICH/335918/2016*. 2019.

(22) Shein-Chung Chow. *Statistical Design and Analysis of Stability Studies*, 1st Edition.; Champan and Hall/CRC: New York, 2007. https://doi.org/https://doi.org/10.1201/9781584889069.

(23) Kutner, M. H.; Nachtsheim, C.; Neter, J.; Li, W. Chapter 16. Single-Factor Studies. In *Applied Linear Statistical Models*; McGraw-Hill Irwin: New York, 2005; pp 677–732.

6. FROM RESEARCH TO IMPLEMENTATION

Advanced technologies allow a better process and product understanding by the monitoring and control of the product; reduced the risk of poor product manufacturing improving their quality, safety, efficiency, and consistency.

Regulatory authorities recognize the ability of advanced technologies to reduce the cost of pharmaceutical manufacturing if implemented correctly[1,2]. The effectiveness of the technology application could be evaluated by measuring the cost associated to achieve a good quality product.

Quality product cost could be summarized in four types[3]:

- Prevention referred to the design, implementation, and maintenance of the quality system.
- Evaluation and audition of products and components ensuring conformance to quality standards and performance requirements.
- Internal failure of a product, component, or material that does not meet quality requirements.
- External failures as a result of shipping defective products to customers.

As a result, sometimes, the associated manufacturing cost per quality unit is not acceptable.

The implementation of advanced technologies has a long-term positive impact on the cost that depends on their integration into the quality and manufacturing system[4]:

- Analytical error reduction due to the low or null interaction of the analyst with the product when using advanced techniques.
- Reduction of materials and laboratory solvents and, therefore, reduction of their waste management.
- Analysis of time reduction and therefore the release of the product due to the real-time process monitoring.

It should be noted that the implementation of these techniques requires a high investment in time, resources, and personnel to guarantee their applicability.

The quality control analysis, according to the requirement of the country of sale, is carried out for each of the batches manufactured covering a large part of the time that is invested for the batch release. As an example, the manufacturing timeline for Sucramal product is presented in Figure 6. 28.

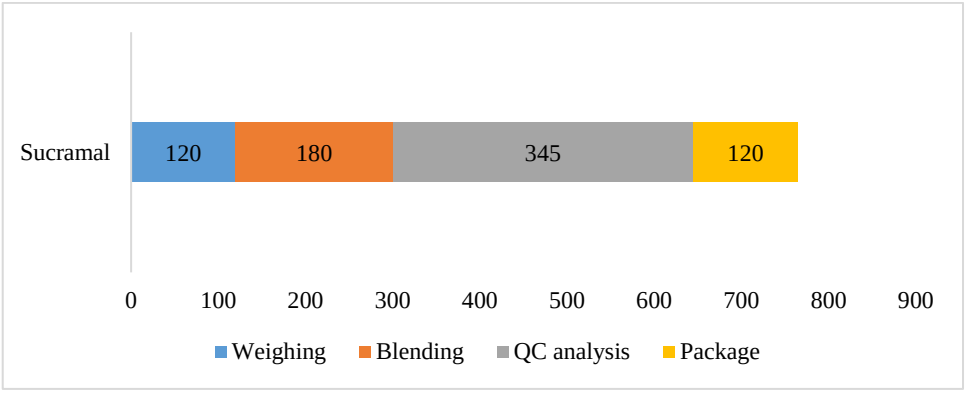
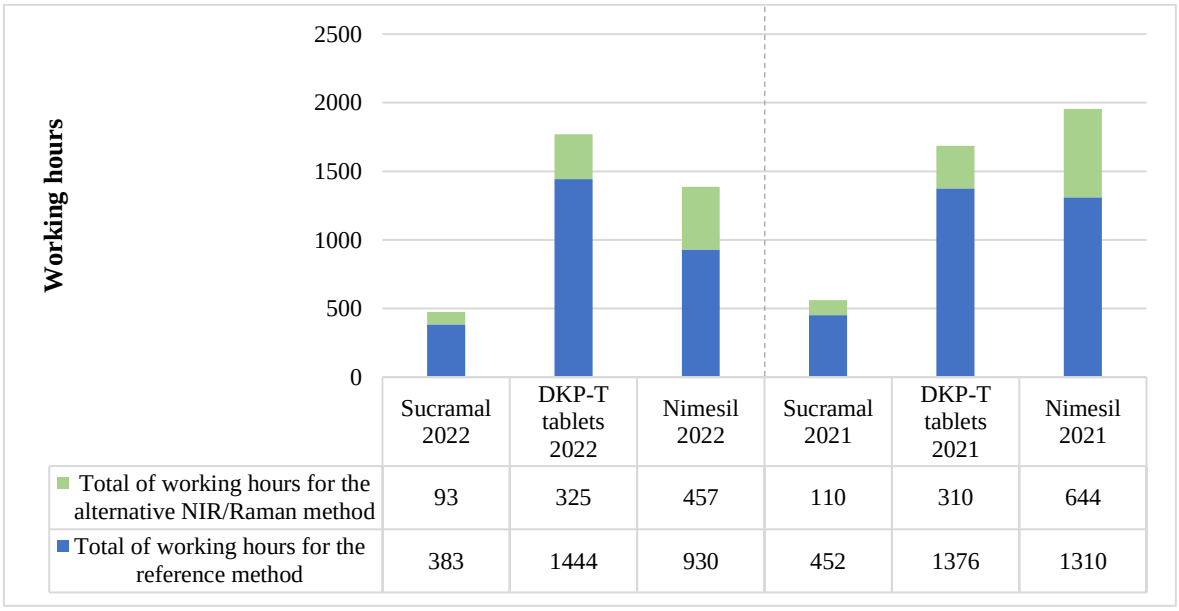


Figure 6. 28 Manufacturing timeline (in minutes) for Sucramal product.

The analysis of time reduction in terms of MBs per year and quality control analyzing time is detailed in Figure 6. 29 for the included researched products.



	Sucramal 2022	DKP-T tablets 2022	Nimesil 2022	Sucramal 2021	DKP-T tablets 2021	Nimesil 2021
Total of working hours for the alternative NIR/Raman method	93	325	457	110	310	644
Total of working hours for the reference method	383	1444	930	452	1376	1310

Considering the NIR/Raman method the control analysis in 1 out of 10 batches by both methods (reference and alternative).

Figure 6. 29 Comparison of hours invested between the reference method and the alternative method.

The analysis times in the determination of the CQAs of interest for the investigated products considering each one of the registered and alternative methods (by NIR or Raman spectroscopy) have been stipulated following the real working conditions in the QC laboratory. In 2021 527 batches of products in solid oral form were manufactured (86 Sucramal, 344 DKP-T tablets, and 131 for Nimesil products). As a result, considering the analysis times per manufactured batch (see Table 6. 11), 3138 hours were invested for quality control by the reference method application, compared to 1064 hours that could have been invested when implementing the alternative methods. Hence, the hiring of one person and 20 percent of the work of a second person could be saved from the time invested (considering 1752 working day hours).

In 2022, 561 batches of products in solid oral form were manufactured (73 Sucramal, 361 DKP-T tablets, and 93 for Nimesil products). As a result, 2757 hours were invested for quality control by the reference method application, compared to 875 hours that could have been invested when implementing the alternative methods. Hence, the hiring of one person and 7 percent of the work of a second person could be saved from the time invested (considering 1752 working day hours).

Table 6. 11 Work hours per batch manufactured for each product evaluated by the reference and alternative method.

Solid oral form product	**Working hours/batch involving the reference method**	**Working hours/batch involving the alternative spectroscopic method**
Sucramal	5.25 hours	0.75 hours
DKP-T tablet	4 hours	0.5 hours
Nimesil	10 hours	3.90 hours

As indicated in Figure 6. 29 the product with higher savings corresponds to DKP-T tablets due to the advanced technology that will be implemented for all MBs.

6.1. Bibliography

(1) Food and Drug Administration. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. 2004. http://www.fda.gov/cvm/guidance/published.html.

(2) Blanco, M.; Villarroya, I. NIR Spectroscopy: A Rapid-Response Analytical Tool. *Trends in Analytical Chemistry* **2002**, *21* (4), 240–250.

(3) Debbie Peru. PAT – Reducing the Cost of Quality in Consumer Products Manufacturing. *European Pharmaceutical Review* **2008**, No. 3.

(4) Panzitta, M.; Ponti, M.; Bruno, G.; Cois, G.; D'Arpino, A.; Minghetti, P.; Mendicino, F. R.; Perioli, L.; Ricci, M. The Strategic Relevance of Manufacturing Technology: An Overall Quality Concept to Promote Innovation Preventing Drug Shortage. *Int J Pharm* **2017**, *516* (1–2), 144–157. https://doi.org/10.1016/j.ijpharm.2016.11.018.

CONCLUSIONS

In conclusion, this thesis has discussed four key application areas for advanced analytical techniques in combination with statistics.

1. After new sources of variability were detected during the identity test for incoming SM involving NIR spectroscopy, the maintenance of the identification library has been developed and properly validated, in terms of specificity and robustness, to unambiguously identify and qualify 237 SM. Moreover, the direct transferability between NIR Bruker equipment (MPA and MPA II) has been demonstrated through the implementation of PCA as a mathematical data treatment and a specific diagram of decisions to visualize possible spectral significance differences. The risk assessment based on the FMEA analysis has been involved to evaluate the SM's criticality for transferability evaluation.
2. The development and validation of offline methods for the identification, qualification, and quantification of the CQA assay, and content uniformity if needed, for solid pharmaceutical forms are satisfactorily achieved. Moreover, the proper maintenance of an inline NIR methodology for identity, assay content, and moisture content determination after the final mixing unit operation has been achieved.
    - Sub-cascading spectral strategy has been used for identity test development. The quantification methods have been developed involving PLS analysis.
    - The set of methods have been satisfactorily validated based on the guidelines approved for this purpose in terms of specificity and robustness for the identity methods, and in terms of specificity, linearity, accuracy, precision, robustness, and range for the quantitative methods.
    - The experimental design for sample sets preparation and the sampling accessory and equipment measurement parameters have been critical for guaranteeing sampling representativeness and applicability throughout the product life cycle.
    - Due to satisfactory results, the NIR strategy is in the approval process as an alternative determination method for the release of batches manufactured in the registry of the product's authorized regions.
3. The validation and if applicable verification of the manufacturing equipment cleanliness has been achieved with the development and validation of methodology based on a rapid general organic carbon technique.

   The applicability of the TOC technique was demonstrated for a water-soluble contaminant.

   The AQbD strategy has been involved to achieve satisfactory results in terms of limit of detection and quantification, linearity, accuracy, precision, and robustness.

Moreover, satisfactory results have been achieved for the transferability study between offline and inline equipment, being able to guarantee the good indistinct use of both.

4. The application of statistics has allowed the optimization of the shelf-life of two specific pharmaceutical forms, solid and liquid, involving the evaluation of historical data for long-term stability studies, avoiding the presence of CQAs out of specification at the end of the useful life.

From these arguments, one can conclude that the methodologies developed in this thesis provide solutions to real needs used satisfactorily in the routine control of the quality of medicinal products in the pharmaceutical industry. Hence, the capacity of NIR and Raman spectroscopy, and TOC as rapid analysis techniques has been demonstrated. As a result, its correct implementation demonstrates a better understanding of the pharmaceutical process that allows to improve the quality, safety, and efficacy of the drug, also highlighting the cost savings associated with the drug manufacturing process.