Master’s Thesis

Anna Haukilahti

LAPPEENRANTA-LAHTI UNIVERSITY OF TECHNOLOGY LUT
School of Engineering Science
Chemical Engineering
Master's Programme in Chemical Engineering and Water Treatment
MASTER'S THESIS

Anna Haukilahti

IDENTIFYING DATA INTEGRITY RISKS IN REGULATED ENVIRONMENT USING PROCESS MAPPING

Lappeenranta 2020

Examiners: Professor Satu-Pia Reinikainen
M.Sc. (Tech.) Piia Poranen

Instructors: M.Sc. (Chem.) Tuija Happonen
M.Sc. (Pharm.) Christina Huntus

TIIVISTELMÄ

Lappeenrannan-Lahden teknillinen yliopisto LUT
School of Engineering Science
Kemiantekniikan koulutusohjelma

Anna Haukilahti

Tiedon eheys ja siihen liittyvien riskien tunnistaminen säädellyssä ympäristössä käyttäen prosessin kuvantamista

Diplomityö, 2020
Sivuja 78, Kuvia 25, Taulukoita 5, Liiteitä 2
Tarkastajat: Professori Satu-Pia Reinikainen ja DI Piia Poranen
Ohjaajat: FM Tuija Happonen ja Proviisori Christina Huntus

Hakusanat: Tiedon eheys, prosessi, prosessin kuvantaminen, GMP, riskinarvio

Tämän diplomityön tarkoituksena oli tutkia tiedon eheyteen liittyviä riskejä ja niiden paikantamista prosesseissa prosessin kuvantamisen avulla. Tutkimuksen teoreettisena taustana terminologian ja kokeellisessa osassa käytettyjen menetelmien tutkimiseen käytettiin pääasiassa kirjallisuutta ja määräyksiä, jotka liittyvät hyvien tuotantotapojen toteuttamiseen laboratorioympäristössä.

Diplomityön kokeellinen osio suoritettiin Oy Medfiles Ltd.:n inhalaatioanalytiikan laboratoriossa Kuopiossa. Tavoitteena oli tutkia, toimiiko prosessin kartoitus tehokkaasti ja tarvittavan tarkasti tiedon eheyteen liittyvien riskien havaitsemisessa. Tutkimuksen aikana laboratorion makroprosessista piirrettiin kartta työryhmässä, josta syntyi lisäksi kaksi pienempää alaprosessia. Näiden kolmen kartan avulla pystyttiin erilaisissa työryhmissä havaitsemaan tiedon eheyteen liittyvät riskit ja siirtämään ne riskien hallintaprosessiin, joka yrityksellä oli käytössä.

Prosessin kartoitus todettiin olevan äärimmäisen hyvä tekniikka tiedon eheyteen liittyvien riskien paikantamiseen yrityksen laboratorioprosesseissa ja sitä aiotaan soveltaa edelleen yrityksen muihinkin prosesseihin. Kartoilla huomattiin myös olevan muita mahdollisia tehokkaita sovellustarkoituksia esimerkiksi henkilökunnan koulutuksessa.

ABSTRACT

Lappeenranta-Lahti University of Technology LUT
School of Engineering Science
Chemical Engineering

Anna Haukilahti

Identifying Data Integrity Risks in Regulated Environment Using Process Mapping

Master's Thesis, 2020
Pages 78, Figures 25, Tables 5, Appendices 2
Examiners: Professor Satu-Pia Reinikainen and M.Sc. (Tech.) Piia Poranen
Instructors: M.Sc. (Chem.) Tuija Happonen and M.Sc. (Pharm.) Christina Huntus

Keywords: Data integrity, process, process mapping, GMP, risk assessment

The purpose of this thesis was to study the risks concerning the data integrity and the identification of these risks with the help of process mapping. In the theoretical part the main terminology and the methods used in the experimental part are studied based on literature and regulations concerning the good manufacturing practices.

The experimental part of the thesis was carried out in Oy Medfiles Ltd.'s inhalation analytics laboratory in Kuopio. The aim was to study if the process mapping method was efficient and thorough enough to identify the data integrity risks. During the study, the macro process of the laboratory was generated in a workshop and within it, two smaller sub processes. With the help of these three maps the different workgroups could identify the data integrity risks and then transfer them to the risk management process, which is used in the company.

Process mapping was concluded to be an efficient method of identifying the data integrity risks within the laboratory processes and it was noted that the method could be applied to company's other processes as well. The maps were noticed to have many different applications in addition to the data integrity risk identification. The training of the staff could use of the process maps for example.

ACKNOWLEDGEMENTS

This master's thesis was started in November 2019 with the best team I could ever hope to back me up. Now, approximately five months later the work has been done and I couldn't be more thankful and happy.

First I would like to thank Medfiles for giving me this special chance and trusting that I could make this project. It has been a pleasure to work with all of you and feel that the work I was doing mattered and will also be relevant in the future.

I'd also like to address thanks to my team: Professor Satu-Pia Reinikainen and instructors Christina Huntus and Tuija Happonen. Special thanks to my examiner from Medfiles, Piia Poranen, who was always prepared to help me and answer even my stupid questions without a doubt.

I am grateful for my sister and father. Even though you sometimes weren't so sure, that I will make it this far, you never stopped encouraging me to keep going.

Elina, Marianne and Tara. I hope you know that without you I wouldn't be at this point, here writing this text. Thank you for the unbelievable support, help and belief through these five years of university. With you in my life, these years have been the best so far.

Mother. I am grateful that you inspired me to study chemistry. I know you would be proud of me.

Kuopio, 26.03.2020

Anna Haukilahti

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ABBREVIATIONS

API	Active Pharmaceutical Ingredient
CAPA	Corrective and Preventive Actions
CDS	Chromatography Data System
cGMP	Current Good Manufacturing Practice
CRO	Contract Research Organization
ELN	Electronic Laboratory Notebook
EMA	European Medicines Agency
FDA	U.S. Food and Drug Administration
FMEA	Failure Mode Effects Analysis
GAMP	Good Automated Manufacturing Practice
GCP	Good Clinical Practice
GDP	Good Distribution Practice
GMP	Good Manufacturing Practice
GVP	Good Pharmacovigilance Practice
GxP	Good *(anything)* Practice
HPLC	High Performance Liquid Chromatograph
ISO	International Organization for Standardization
LES	Laboratory Execution System
LIMS	Laboratory Information Management System
MHRA	Medicines & Healthcare products Regulatory Agency
OOS	Out of Specification

OOX	Out of Specification, Expectation, Trend
PIC/S	Pharmaceutical Inspection Co-operation Scheme
PIF	Pharma Industry Finland
QA	Quality Assurance
RPN	Risk Probability Number
SOP	Standard Operating Procedure
UHPLC	Ultra-High Performance Liquid Chromatograph
WHO	World Health Organization

SYMBOLS

D	Detectability
P	Probability
S	Severity

1 INTRODUCTION

Before the Industrial Revolution in the 18th century, almost all products were provided by craftworkers. The products were done by one person from beginning to the end and there was no need for defining the process. The need for identifying the processes becomes useful when there are more than one person working in and around the process. (Stankiewicz and Moulijn, 2005)

The term *business process* became known widely all over the world in the early 1990s. Business processes were quickly linked to terms *re-engineering* and *process improvement*. This ground-breaking new way of thinking suddenly fell apart a few years later. The principles of re-engineering were understood and used wrong and it created huge losses for the organizations. The business process improvement was understood to mean laying off the staff and outsourcing, which created the negative tone to the re-engineering amongst the employees. (Sharp and McDermott, 2009)

From the beginning of the millennium, the interest in process improvement has begun to rise again and the tools for re-engineering are evolving quickly. Organizations want to try re-engineering and process improvement for various reasons. New interest can be explained by wide investigations of '*what went wrong in the 1990s'* and positive results. Numerous researches and explanations have been done and the negative tone from re-engineering has been successfully removed. (Sharp and McDermott, 2009)

Process mapping is done in different industries for example to make organizations' processes more transparent and to help visualise the process to analyse it. Process mapping is done for a variety of reasons, but almost always it is done to improve something. The maps can help to train new employees, make the risk management more efficient or increase quality. (Triaster, 2018)

Oy Medfiles Ltd., later referred to as Medfiles, has had a need for mapping of its processes to efficiently locate risks involving data integrity. Problems concerning *good data and record management practices* are increasing in the pharmaceutical industry since the fast development of the electronical systems. It is important to locate the data integrity risks

within organization's operations and try to reduce the consequences or remove the risks completely. (WHO, 2016)

According to the good manufacturing practices, the data collected in drug manufacturing has to be complete, consistent and accurate throughout its lifecycle. The increasing interest of the authorities regarding the data integrity is a result of the growing number of violations concerning the data integrity in the pharmaceutical industry. Data integrity assures the quality of the drugs and improves the patient safety by making the production transparent for authorities and the organization. (FDA, 2018)

1.1 Purpose and structure of the thesis

The aim of this master's thesis is to find efficient tools for process mapping. The purpose of the process mapping is to find risks regarding to data integrity in different processes. The experimental case will be carried out in Medfiles' inhalation analytics laboratory. The maps of the laboratory will be created based on the theory part of this thesis and the mapping will be conducted by the help of the so called *workshops*.

The main goal of this master's thesis can be defined in one question:

'What are the risks, concerning the data integrity in Medfiles' inhalation analytics laboratory and how to manage them?'

The theory part of this thesis defines the main terms regarding the subject and covers the aspects which need to be considered in the actual process mapping. Chapters 2-5 will cover the basics of the GMP and data integrity regulations, risk management and the process mapping process. The last Chapter of the literature review, Chapter 6 will cover tools and methods for process mapping and risk management.

In the experimental part the aim is to define and create a visual process map for Medfiles' inhalation analytics laboratory located in Kuopio. Chapters 7 and 8 will introduce the inhalation analytics laboratory of Medfiles. Chapter 8 will also cover sample flow in it according to Medfiles' standard operating procedures.

Chapter 9 will present the first actual part of the mapping process and the first process map of the laboratory is presented. In Chapter 10, risks concerning data integrity found with the help of the process map are presented and discussed. Chapter 11 is about the subprocesses of the macro map. There are two different subprocesses that were drawn in order to find the root causes of different data integrity risks. In Chapter 12, the risk management process is conducted. Risks are prioritized and FMEA is applied in the risk management process.

The process map of the inhalation analytics laboratory is drawn in Chapter 13 again in order to make it more accurate, and easy to use based on the subprocess maps and gained information during the mapping process. Chapter 14 discusses the needed changes in order to reduce or remove the found risks. In Chapter 15, the future and other uses for the maps are presented.

1.2 Oy Medfiles Ltd.

Medfiles is a contract research organization (CRO) offering a wide range of services. The services include consultation, development and laboratory services for pharmaceutical, food and nutrition and medical device industry. Medfiles was established in Kuopio in 1987 and is nowadays owned by a Japanese company WDB Holdings Co., Ltd.. Medfiles has three offices in Finland and three in Estonia, Latvia and Lithuania. (Oy Medfiles Ltd., 2019a)

This master's thesis' experimental part is made with Medfiles in the Kuopio laboratory unit and focuses on its inhalation analytics laboratory.

LITERATURE REVIEW

2 REGULATORY ENVIRONMENT AND GxP

In Medfiles' operations, they must follow the given guidelines by European Union. The guidelines are given by EMA, European Medicines Agency, and these guidelines apply to all medicine manufacturers that import and produce medicines to EU markets. EMA (2013, pp. 2) has stated that '*The holder of Manufacturing Authorisation must manufacture medicinal products so as to ensure that they are fit for their intended use, comply with the requirements of the Marketing Authorisation or Clinical Trial Authorisation, as appropriate and do not place patients at risk due to inadequate safety, quality or efficacy*.' This statement applies to all manufacturers and their distributors and suppliers. (EMA, 2013)

To achieve the quality objective stated by EMA the organization should have an efficient pharmaceutical quality system, which includes quality risk management and the current good manufacturing practices (cGMP). GMP sets the minimum standards for the production processes. GMP is considered also in quality control. The purpose of quality risk management is to evaluate and control possible risks in processes to ensure the quality of products. (EMA, 2013)

In addition to GMP, Medfiles also follows the guidelines of good pharmacovigilance practice (GVP), good distribution practice (GDP) and good clinical practice (GCP).

3 DATA INTEGRITY

Data integrity is a part of GMP and can be defined in different ways depending on the source. For example, U.S. Food and Drug Administration, FDA (2018, pp. 4) has defined the data integrity to be '*completeness, consistency and accuracy of data*'. FDA has also stated that

data should be attributable, legible, contemporaneously recorded, original and accurate (FDA, 2018).

The first data integrity guideline was published in 1963 by FDA (Rattan, 2018). FDA's definition of data integrity is defined by the term 'ALCOA', which is an acronym of five requirements for data to be integrated. World Health Organization, WHO has stated that in addition to ALCOA, data should be complete, consistent, enduring and available. These four requirements added to ALCOA form ALCOA+, which is presented in Figure 1. (McDowall, 2019a)

Figure 1. Extended ALCOA+ creates a good base for the data integrity. When followed, the risks regarding the data integrity should be minimized. (Rattan, 2018)

In Figure 1, all nine points of data integrity are presented. All these nine points affect the data quality and thus should be considered in the risk assessment in order to minimize the risks (Rattan, 2018). In Table I, the ALCOA+ terms are explained.

Table I. Definitions of ALCOA+ terms. (McDowall, 2019a)

Term	Definition
Attributable	Attributable means that the creator and possible modifier of the data can be identified. The data can be created by a person or for example computerised system. The attributable data has also the time and date of the creation and possible modification.
Legible	Legible data is unambiguous. This means that the data is fully understandable and the action it describes can be recreated without misunderstandings.
Contemporaneous	Contemporaneous means that the action or observation is recorded instantly.
Original	Data should be original. This means that data in the paper or file should be found in its first original form and the original version should be easily located. The term 'true copy' means that the copy of the original record is verified and accurate. Original data cannot be changed or deleted without leaving a mark.
Accurate	Accurate data is error free. The data should not be edited, but if it is necessary, the original data should be available. The editing event should have time and date and the editor should be identified.
Complete	Complete data has all obtained data from the analysis. In hybrid systems, where there are electronic systems and paper data in use together, the data should be linked together. The completeness of the electronic system can be assured with thorough the audit trail.
Consistent	Consistent means that the events are in order. This means that dates and possibly times of actions are recorded.
Enduing	Enduring data is recorded in official form. These official forms are for example identified worksheets or books that can be located.
Available	Available means that the data is easily accessed and can be reviewed. The data should be available for its defined lifetime.

3.1 Data governance

Data governance is defined by Pharmaceutical Inspection Co-operation Scheme, PIC/S (2018, pp. 5) as follows: '*Data governance is the sum total of arrangements which provide assurance of data integrity.*' This means the actions that need to be taken in order to ensure the integrity of data throughout its entire lifecycle. The cycle starts from the generation of data and ends to the discarding of it. (PIC/S, 2018)

The systems regarding to data governance should be part of the pharmaceutical quality system and follow the quality risk management philosophies. The data governance system is recorded in Quality Management System (QMS) and audited frequently. The supervision of the systems should be the responsibility of the management in order to remove the risks concerning the data integrity. According to PIC/S, pharmaceutical manufacturers and analytical laboratories should have a risk management system regarding the data integrity, that assures the standard integrity of the data. The systems should be completely documented. (PIC/S, 2018)

3.2 Electronic archiving

In the hybrid laboratory system, the data is archived physically and electronical data electronically. There are guidelines for the electronical archiving, because the electronical archiving may cause severe risks to the integrity of the data. (PIC/S, 2018) When archiving electronically, the data must contain the metadata and complete original data. Metadata must contain audit trails and the data has to be archived accordingly by the validated process. (WHO, 2016)

Data that is stored electronically should be accessible during its whole lifecycle just as the data that is on paper. The electronical data should also be possible to print out from the archiving system. This concerns all archived data, including metadata. The archiving system should have written procedures to follow and the disposal of the data should also be defined. (PIC/S, 2018)

4 RISK MANAGEMENT

International Organization for Standardization (ISO) has stated (2018, pp. 6) that risk is '*the effect of uncertainty on objectives*'. Risks occur in every process and can be assessed. Risks that are tolerable can be noted and accepted and risks that are unacceptable will be processed. All risks in processes need to be located in order to manage them. The risk management system controls all acceptable and unacceptable risks and its purpose is to create a tolerable environment to the process. (Nichols, 2011)

In general, risk management is for example a way to improve organization strategy or targets and help in decision making. Risk assessment is a process which includes the identification of the risk, risk analysis and the assessment of the significance of the risk. After the assessment, it is important to act in order to remove the risk or try to make it less significant for example by dividing the risk or changing the consequences of the risk. The monitoring and controlling of the risks should be part of the whole process in every phase. It is important to record the observations in order to make the risk management system efficient. (ISO, 2018)

Medfiles' risk management system follows the '*ICH guideline Q9 on quality risk management'* (2015) by EMA. The purpose of efficient quality risk management system in pharmaceutical industry is to ensure the high quality of the pharmaceuticals and thus improve the safety of the patient. (EMA, 2015)

According to the ICH guideline Q9 by EMA (2015, pp. 4), in quality risk management there are two principles that should be prioritized:

- *'The evaluation of the risk to quality should be based on scientific knowledge and ultimately link to the protection of the patient; and'*
- *'The level of effort, formality and documentation of the quality risk management process should be commensurate with the level of risk.'*

The quality risk management system should be a process that is controlled by a team that has members with different expertise. This team's responsibility is to assess the risks in different tasks within the organization based on the mentioned principles. They should also make sure that the system is defined and used adequately. (EMA, 2015)

4.1 Risk management process

The first part of the risk management process consists of risk identification, risk analysis and risk evaluation. This part is called risk assessment. After the assessment, the risks need to be controlled and this part contains risk control, -reduction and -acceptance. At the end of the risk management process, the risks should be reviewed and the process can be started again. During the process, the risk communication is important. The risk management process is presented in Figure 2. (EMA, 2015)

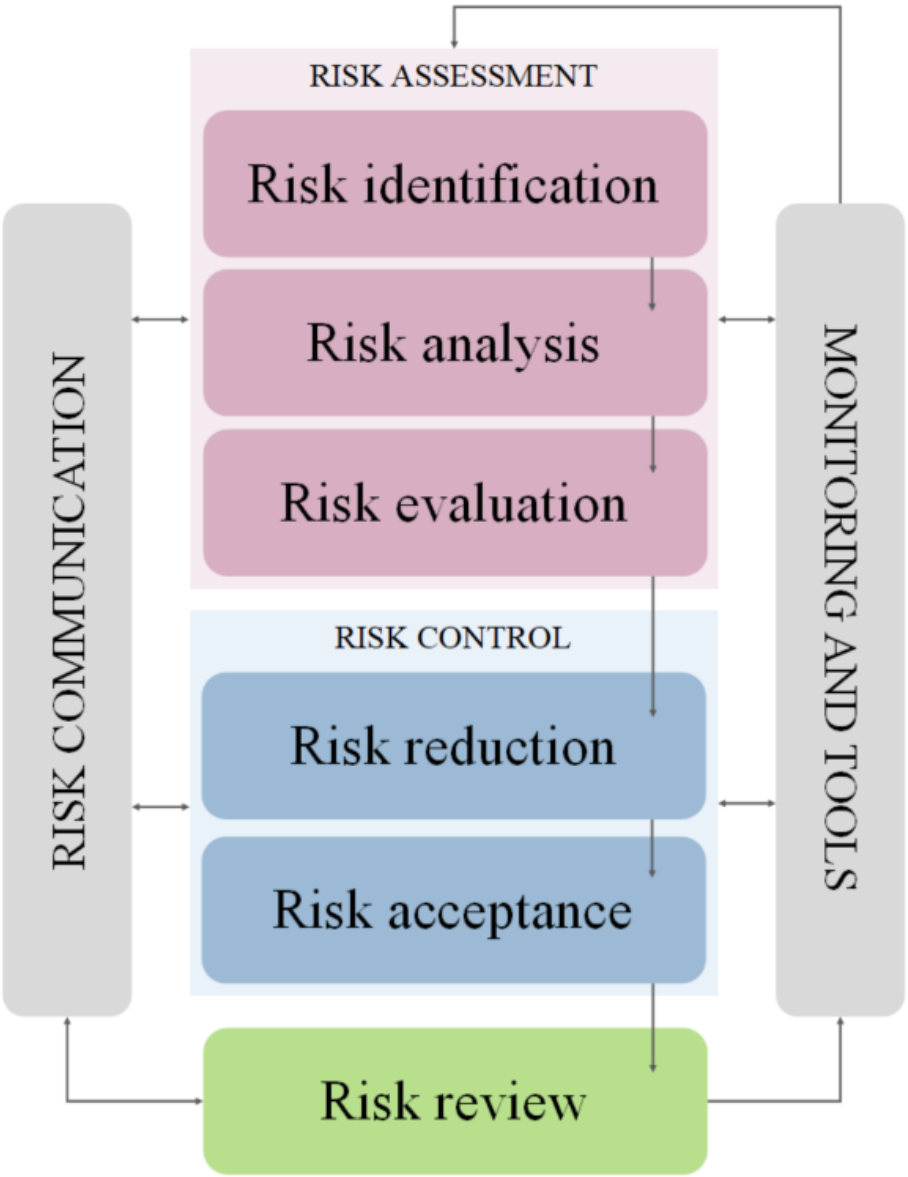

Figure 2. The risk management process consists of risk assessment, risk control and risk review. Risk communication and monitoring with applicable tools are done during the process several times. (EMA, 2015)

The risk management process starts by finding the risks. This is the first part of the risk assessment. All the risks found within the organization processes should be identified even though the root causes of the risks are not in the control of the organization (ISO, 2018). The identified risks need to be analysed. This means that the probability and the severity of the consequences of each risk are assessed. After the analysis, the risks are evaluated. The result of the risk assessment can be a qualitative representation of the scale of the risk or a quantitative assessment of the risk. (EMA, 2015)

The second part of the risk management process is risk control and it includes risk reduction and acceptance. The risk control is done to make the risk as ‘harmless’ as possible in order to accept the risk. When the risk reduction is done, it is possible that it creates new risks. If the reduction cannot be done efficiently enough, for example the risk analysis can be done again. The risk acceptance is just an acceptance decision of the risk. (EMA, 2015)

The communication and monitoring during the risk management process is important. The communication should happen between the quality risk management team and employees that are concerned with the risk and subjects of it. The communication should be done thoroughly at the review of the risk in order to receive more understanding and expertise concerning the risk. The risk should be monitored also after the process. (EMA, 2015)

Risks can be managed and assessed with tools of a different kind that are effective for each organization. These can be for example risk ranking and filtering or failure mode effects analysis (FMEA), which is used at Medfiles. (EMA, 2015)

4.2 Failure mode and effects analysis

Failure mode and effects analysis, or shorter FMEA has been used in the product quality assessment widely in different industries since the 1960s, when it was implemented by NASA. Different variations of FMEA are used also in Medfiles’ risk management processes. FMEA contains the prioritization of the risks by three different properties: severity (S), probability (P) and detectability (D), and needed actions in order to reduce the risk or remove it completely from the process. The model of FMEA, that will be used in this thesis is presented in Appendix I. (Liu, 2016)

In FMEA the risks concerning each process step are listed into a model and different reasons and consequences are written down. After this the S, P and D values are given to each risk according to Table II.

Table II. Severity, probability and detectability values table used in this thesis.

	Severity (S)	**Probability (P)**	**Detectability (D)**
1	Low risk	Unlikely to happen	Can be detected easily
2	Medium risk	Happens probably	Detectable, but unlikely
3	High risk	Happens almost certainly	Cannot be detected

The range for the values of S, P and D can vary from three to even ten, but when the count of different stages is increased, the evaluation gets harder. After the three different values are defined, the RPN (Risk Probability Number) is calculated with equation 1 below. (Liu, 2016)

$$RPN = S \cdot P \cdot D \quad (1)$$

When RPN is calculated, it helps to evaluate, what to do for the risk. If the result is low, for example 10% from the maximum score of the chosen range of S, P and D, there is no need for actions. If the result is below 50% from the maximum value, but over 10%, the risk has to be reduced or accepted and if the result is over 50% from the maximum value, the risk is not acceptable and has to be reduced. The percentages can vary depending on the organization. In Table III, the scale for RPN number is defined when the chosen range of risk levels is 1-3 as presented in Table II. (Kiran, 2017)

Table III. Risk probability number action plan.

RPN	**Action**	
≤ 3	No need for actions (green)	10 % from the maximum value
4 - 14	Risk has to be reduced or accepted (yellow)	-
≥ 15	Risk has to be reduced (red)	50 % from the maximum value

5 PROCESS MAPPING

Process mapping has become an important topic within organizations in short time. The development has been extremely fast after the terms *business process* and *re-engineering* became a phenomenon in the 1990s, even though they took a few steps back along the way. The point of process mapping was to understand and visualise the end-to-end process of the organization. The end-to-end process is the whole process from the suppliers to the customers and can exceed organizational boundaries for example to the suppliers' suppliers or to the customers' customers if needed. (Sharp and McDermott, 2009)

A great motivation for the process mapping was and still is to intensify the processes. When intensified, the production could become faster, cheaper and could even improve the organization image. (Stankiewicz and Moulijn, 2005) Mapping helps to visualise the processes and see the possible problems within it. Processes can be divided in four levels which are main process or macro process, subprocess, activity and task. The hierarchy of the process is presented in Figure 3. (Kalman, 2002)

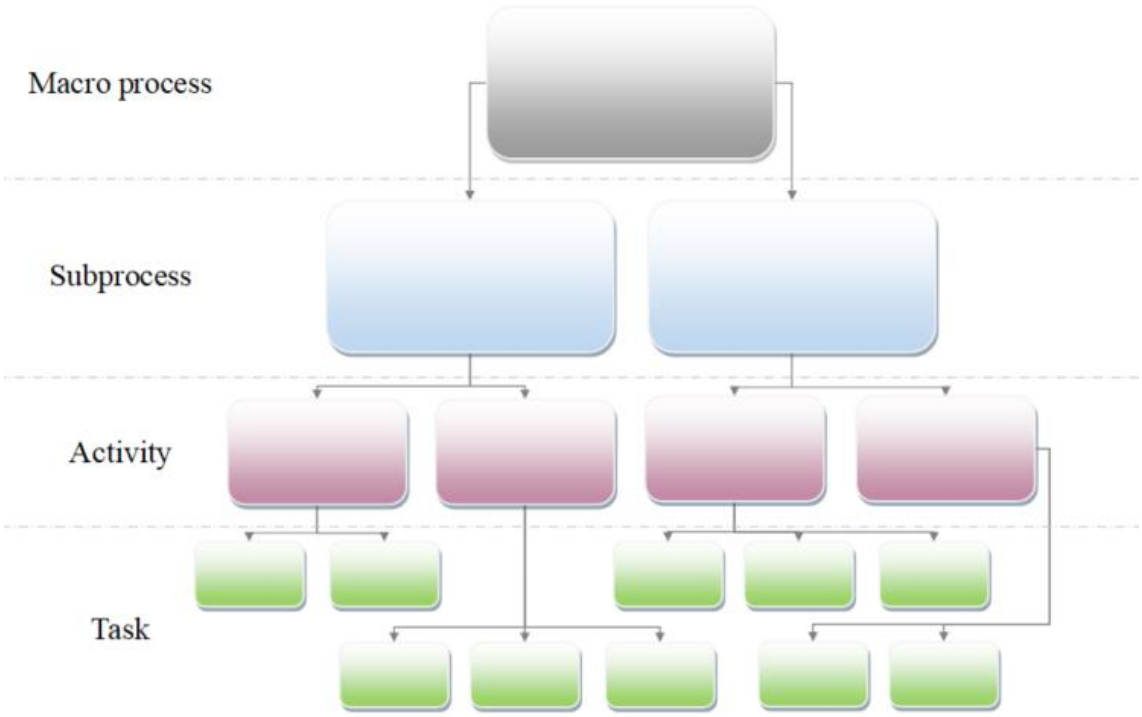

Figure 3. The macro process consists of subprocesses, which consists of activities, which consists of individual tasks. (Kalman, 2002)

The whole process, as presented in Figure 3 must be considered in process mapping, not just the macro process. Processes need to be understood properly, before the mapping can be done. It is important that the person creating the process maps has wide enough understanding of the processes in question. The process mapping is beneficial to organization but also to employees. Process mapping, when done right, can make the work more significant to individuals, workflow becomes more effortless and quality improves. These are just few benefits of process mapping and the improvements depend on the organization. Almost always the mapping decreases costs, which is a great motivation to the organizations. (Kalman, 2002)

In the GMP, or any GxP environment, where the data integrity problems and risks are extremely important to be identified the process mapping is a very efficient way to do it. The mapping that is done in order to find data integrity risks is called *data process mapping* and with it the areas where the data integrity is vulnerable can be located and mapped further. (McDowall, 2019b)

Medicines & Healthcare products Regulatory Agency (MHRA) states (2018, pp. 4) the principles of data integrity and declares that *'Organisations are expected to implement, design and operate a documented system that provides an acceptable state of control based on the data integrity risk with supporting rationale. An example of a suitable approach is to perform a data integrity risk assessment where the processes that produce data or where data is obtained are mapped out and each of the formats and their controls are identified and the data criticality and inherent risks documented.'* The important part is that the organization need to have a clear understanding about each process, in where the data integrity is crucial. The map should cover the whole process and consider all the aspects, for example employees, regulations and organizations quality systems. (MHRA, 2018)

5.1 Process mapping method

Process mapping can at the simplest be conducted with steps in order to take all aspects into consideration. Different sources present these steps a somewhat differently, but the main flow of the process mapping process is alike. (Kalman, 2002) Kalman (2002, pp. 62-67) describes the process in seven steps, which are presented in Figure 4.

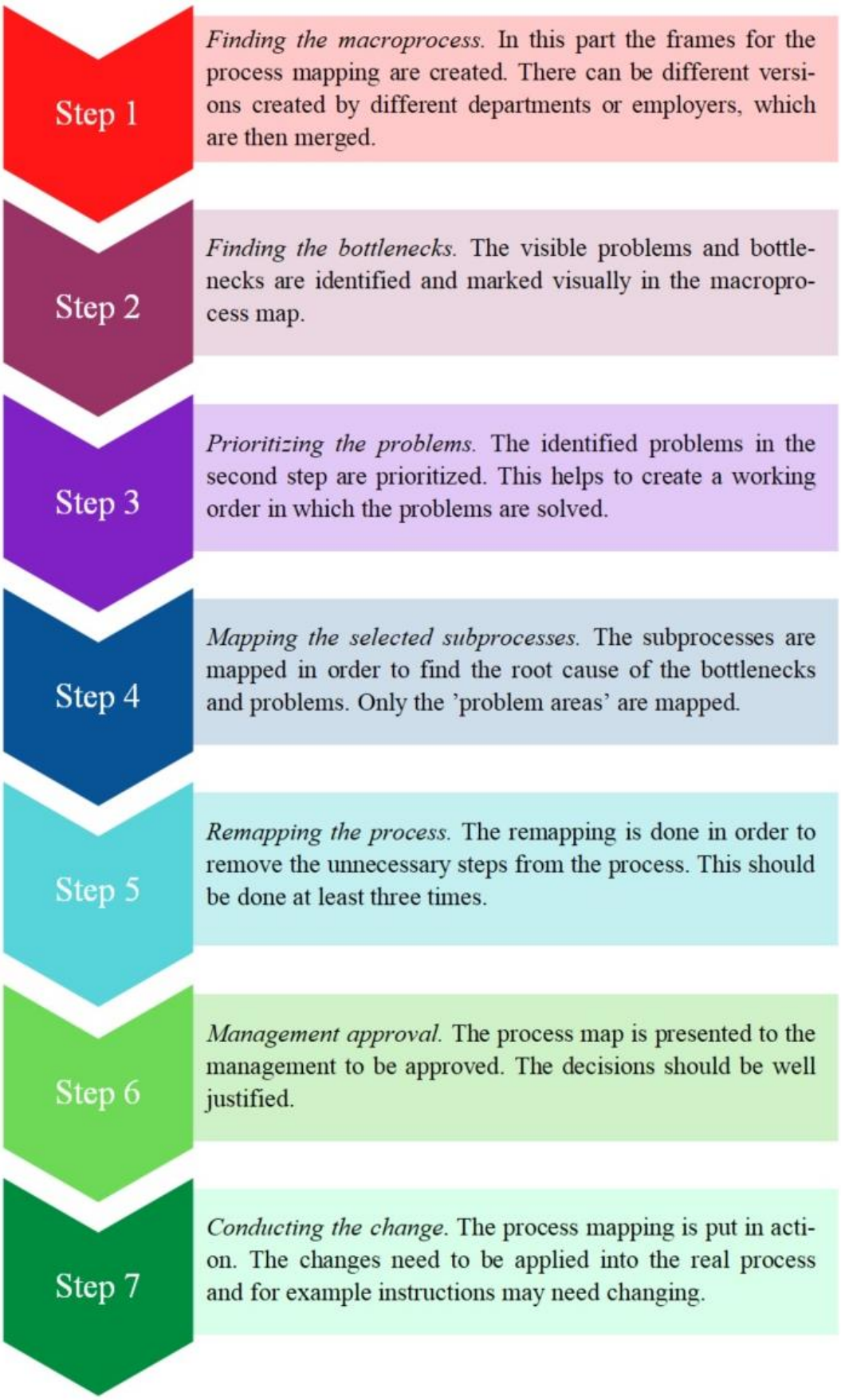

Figure 4. Mapping process by Kalman.

Before starting the mapping process presented in Figure 4, it is extremely important to have the authorization and support of the management. Without the support, the project can be expected to fail. The management will also define the processes that needs to be mapped. The mapping work group should have a variety of members that have different skills and titles but work around the same process. The group size should be less than 12 people but can consist of smaller groups that have for example different perspective. (Kalman, 2002)

In the first step the macro process is drawn. It is important to visualise the process. If more than one map is generated in the process, the maps are merged. The structure of the map should be simple and the number of phases should be as minimal as can be. (Madison, 2005)

With the help of the macro process map the second step can be started and the problematic bottlenecks and other issues are identified. The third phase is to prioritize the problems to define the order in which the problems are solved. (Kalman, 2002)

The subprocesses are drawn only in the parts, where the problems are found. This is done in the step four. It is a waste of time to draw all the subprocesses. It makes the process map easily unreadable. The subprocesses can be constructed as far as into activities and tasks depending on the root cause and how hard it is to track. After the root causes have been mapped, the process map should be redrawn in order to locate all problems and bottlenecks. The mapping procedure should be done two to three times during the mapping process to be efficient enough. (Kalman, 2002)

After the mapping has been done completely and all the unnecessary phases, risks and bottlenecks have been processed, the plans should be presented to the management approval. This is the part where the management support comes crucial. In the presentation, the actions should be planned and all the decisions justified properly. After the approval, the last step is to put the plans in action and do all the necessary work in order the changes to fit into the existing process. (Kalman, 2002)

5.2 Problems and mistakes

Process mapping, while being an efficient technique to improve organization productivity or make the processes visual, is criticized for being a slow and expensive technique. The reasoning behind this is that the mapping is easily ‘overdone’. The greatest potential of process mapping can be achieved by being aware of the commonly done costly mistakes and time consuming drawbacks of the mapping process. (Rosemann, 2006)

There are an infinite number of aspects to consider in the mapping process that may make the process inefficient, but it is impossible to take all into account. According to Rosemann (2006, pp. 251) there are a few typical mistakes that need to be avoided in order to make the process mapping successful. (Rosemann, 2006) The mistakes are presented in Figure 5.

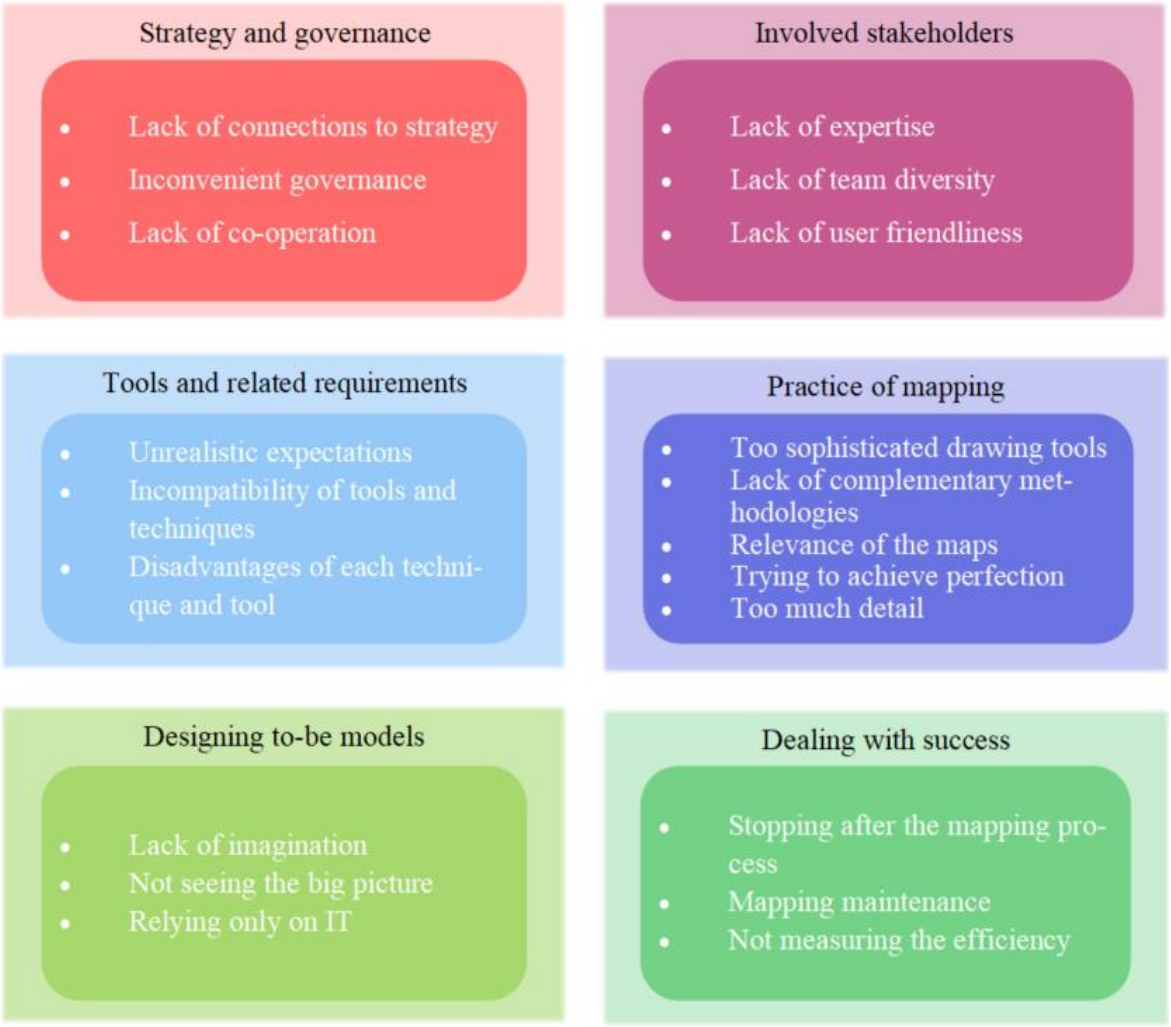

Figure 5. The 'pitfalls' of process mapping according to Rosemann (2006).

Rosemann presents so many 'pitfalls' to be avoided that sometimes it is not possible to avoid all of these. The important thing is to know what kind of mistakes can be disastrous to process mapping. In Figure 5, the mistakes are presented in six different categories, which cover the mapping process from the planning to the maintenance of the maps. (Rosemann, 2006)

Before the mapping process is started, the mapping team should have full support from the managers, whose responsibility is to choose the right processes to be mapped. The team should be formed to be as diverse as can be. There should be experts from each department working around the mapping process and the team should have enough time to work on the mapping process. There should also be someone who has expertise in the mapping process.

With the help from the management, the organization strategies can be linked to the process mapping if wanted. (Rosemann, 2006)

When the mapping process is started, many different things can go wrong. The tools used for the project should be well studied and different possibilities considered. It is not always the best idea to use old ways but on the other hand, something that works for other companies, does not necessary work for us. The tools chosen should be applicable to the organization processes and working ways properly. It should be also remembered that the tools have their own disadvantages. (Rosemann, 2006)

When the maps are created, it is easy to forget that the users of the maps should be able to use them for the required purpose. The maps need to be easy to understand and it is not necessary to make the maps perfect by small detailing. This can make the maps unreadable. The mapping team should have the same goals and expectations and be creative in the process. It should also be remembered that all the problems cannot be solved with new software or updates. (Rosemann, 2006)

After the mapping is done the governance and maintenance are important. The project may be at the end, but the process is ongoing till the process maps are outdated. There should be tools or ways for measuring the wanted efficiencies of the process and the maps should be updated when necessary. (Rosemann, 2006)

5.3 Process owners

A process owner refers to a person that as the name suggests, owns the process. A process owner is responsible for various things depending on the organization, but in general according to Fleischmann *et al.* (pp. 53, 2012) '*He is responsible for accepting the process model and is in charge of its implementation. During operation, process change requests must be approved by the process owner. He takes care of regular monitoring of the process and its optimization, if necessary'.* Process models mentioned here are same as process maps. (Fleischmann et al., 2012)

Medfiles (2020, pp. 2) defines the process owner followingly by GAMP (Good Automated Manufacturing Practice): '*The person ultimately responsible for the business process or process to be managed.*' This definition is somewhat more straightforward, but is generally the same as the definition by Fleischmann *et al.*. (Oy Medfiles Ltd., 2020)

Process owners are important to be nominated in order to maintain the processes. The process owner is a responsible person but can delegate different tasks to capable people, for example experts or responsible persons. (Fleischmann et al., 2012)

6 TOOLS AND METHODS

There are a wide range of different techniques to conduct the process mapping. In this Chapter, the different tools and methods used in this thesis are presented.

6.1 Microsoft Visio

Microsoft Visio is widely recommended and used tool in process mapping. There are many different styles and designs in Visio to create process maps and the choosing of the style should not be hurried. All styles have their strengths and it is important to find the one that works for the organization's processes. (Helmers, 2018)

When choosing the proper style for the process map the first thing is to choose the symbols that will be used. The symbols should be understandable by the mappers and the users of the maps. In Figure 6 the basic symbols that are used in process mapping. (Helmers, 2018)

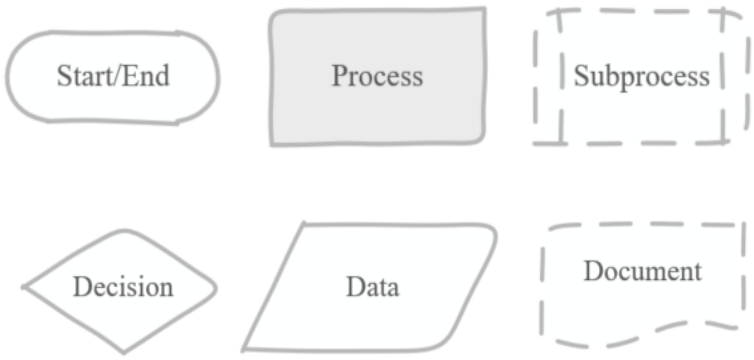

Figure 6. Basic flowchart symbols used in process mapping.

In Figure 6 there are presented the general meanings of each symbol but the meanings can be changed if needed. It does not matter what symbols are chosen for process, decision, data etc., but as mentioned, they should be consistent throughout all organizations maps and defined to all who read the maps. There should be few different symbols in order to make the maps easy to read. (Helmers, 2018) The symbols presented in the process maps should all be connected to other symbols with a connector, which usually is a simple arrow. The decision box should have at least two arrows away from it, because there should be something to choose from. (Abubakker, 2012)

The map will be conducted as cross-functional flowchart or so called ‘swimlane’ process map. The swimlane map is presented below in Figure 7. The purpose of a map like this is to see the data flow of the process and visualise for example the department, a responsible person or a software where or by whom the process step is done. (Visual Paradigm Online, 2020)

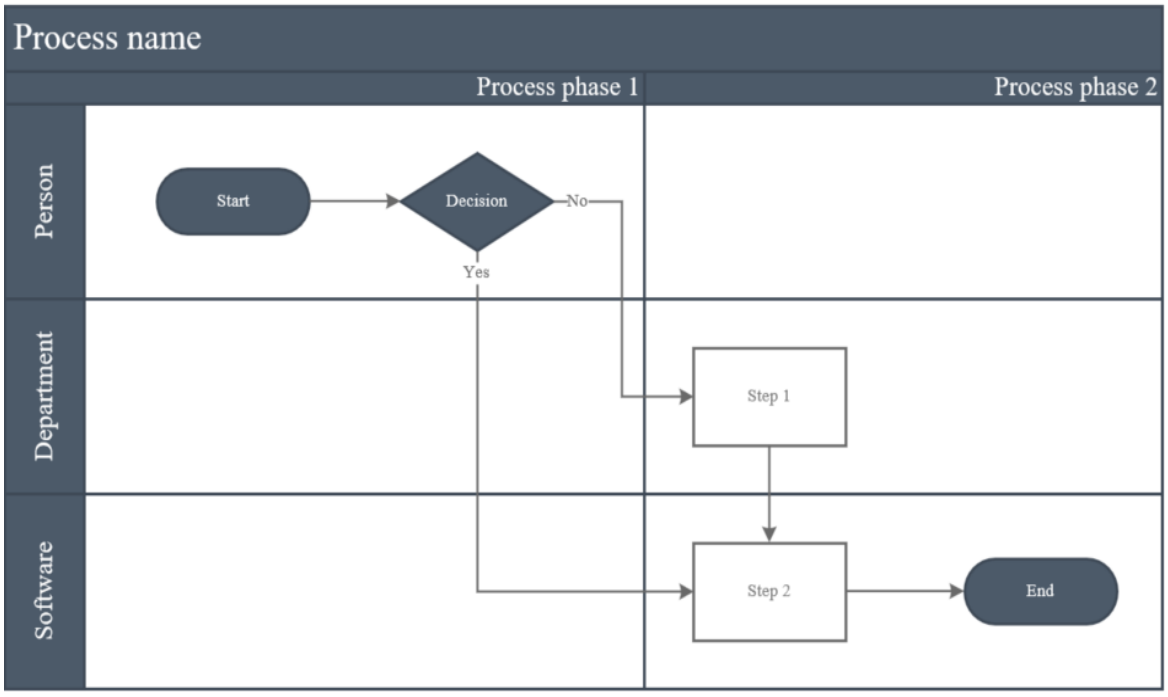

Figure 7. Simple swimlane process map.

In the swimlane process map the swimlanes can be vertical or horizontal and the phases are divided other way around. In Figure 7, the swimlanes are horizontal and the phases are divided vertically. The process flows from the upper left corner down and right, though in the large scale processes this cannot be fully executed. (Sandahl, 2015)

6.2 Workshops

Process mapping is a hot topic and it can be done by different ways. One major thing to remember is that no one can do the mapping process alone. In order to the map being thorough enough, people from different work expertise should be involved. This can be done by interviewing, analysing or by workshops. (Cousins, 2018)

In order to conduct a successful mapping workshop, a few important things need to be checked first. The workshop needs to have a leader who conducts and manages the workshops. Someone should also be an expert in the mapping process and help to avoid the

‘pitfalls’ and know how the processes are mapped. The leader and the mapper can be the same person if they are competent. (Gaillard, 2015)

There is work to be done before starting the workshops. The agenda and the process to be mapped must be clear. It is crucial that the team knows where the beginning of the process is and towards what goal the process is built. The team also needs to know the timetables and the targets of the mapping process. The efficient team should also be formed. Workshop team should not have more than ten attendees. (Triaster, 2018) The formation and decision making within the workshop is presented below in Figure 8.

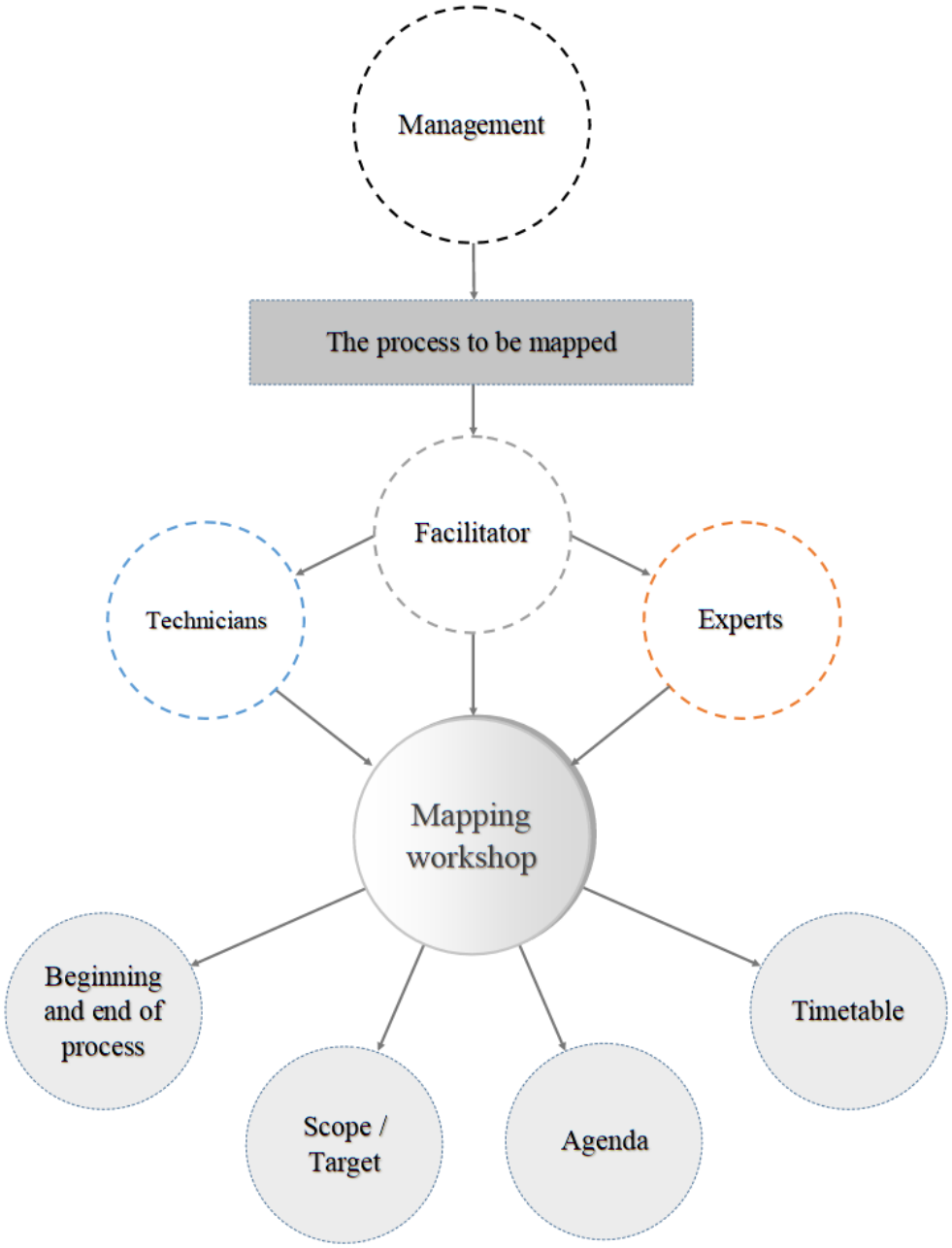

Figure 8. The workshop should consist of a skilled team. The team must accept on few things before creating the process maps.

After the preparations, the workshop can be conducted. Even though the maps are eventually recorded with the chosen software, the best results in teamwork can be achieved with paper and a marker pen. For the base of the process map large empty wall and roll of paper is needed. The paper and Post-It-Notes gives a team more involved feeling and everybody can

participate easier (Kalman, 2002). There should be a set of differently coloured or shaped notes in order to indicate different tasks or process parts as presented in Figure 6. (Triaster, 2018)

Important parts of the process mapping are to create a name for the process, make sure, that everyone in the team sees what is happening in the map and leading the team towards the right targets for example by asking questions. (Dudenhoefer, 2018)

When the map has been finished, there are still few steps to be done. The process map can now be imported to the chosen mapping software. After doing this, the data needs to be checked. The labels should be written correctly and the map should be simple and readable. After this with the help of different team, the bottlenecks and risks can be identified and the project continued. (Triaster, 2018)

EXPERIMENTAL PART

7 INHALATION ANALYTICS LABORATORY

The experimental part of the thesis starts here. This part of the thesis follows the process mapping method presented in Chapter 5.1. The structure of the experimental part is presented in Figure 9 below.

Figure 9. Structure of the experimental part of the thesis and the actual process of finding and managing the data integrity risks within any process.

In Figure 9, the start is mapping the macro process of the inhalation analytics laboratory and the process goes on clockwise till the map expires. This process can be used in any process and can be used to find any risks or bottlenecks in the processes.

Medfiles' inhalation analytics laboratory is located in Kuopio and has GMP certificate. The laboratory handles mainly inhalation drug products which are dry powders and the processes are specified in them. This means for example that the laboratory technicians are trained to inhalation analytics. Most of the instrumentation is also designated for inhalation analytics. All the samples go through the same procedures even if they are not inhalation drug products.

There are two main analytical techniques in the inhalation analytics laboratory, which are high- (HPLC) and ultra-high performance liquid chromatographs (UHPLC). In addition to this titration technique is typically used in the analysis. Some instruments that are uncommon in pharmaceutical laboratory are used in sample preparation in inhalation analytics laboratory.

The inhalation analytics laboratory was chosen for the mapping process, because its processes are quite similar regardless of the product or the client.

8 SAMPLE FLOW IN LABORATORY

Sample flow in Medfiles' laboratories has generally been described in the standard operating procedure (SOP) KL-114. The process described directs the sample flow in all Medfiles' laboratories. The sample flow according to the SOP KL-114 is presented in Figure 10 below.

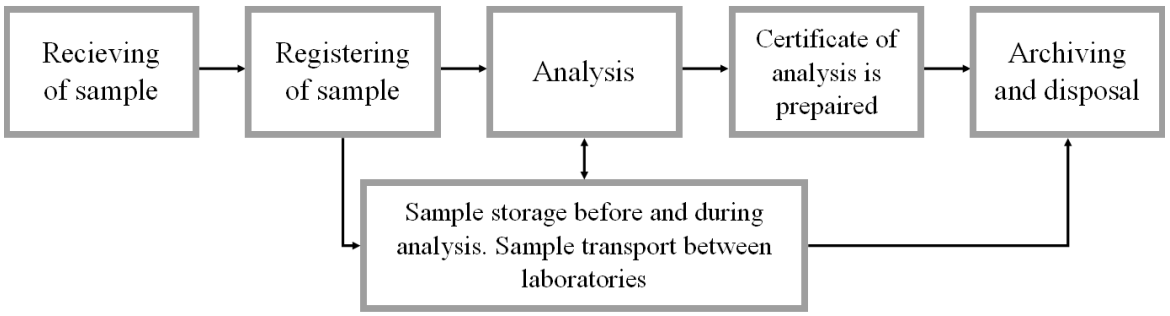

Figure 10. Sample flow in general in Medfiles' analytical laboratory.

When the sample arrives at Medfiles, it is received, and the information about the sample is checked. The information that is needed includes for example the name of the product or the active pharmaceutical ingredient (API) and batch number in order to identify each sample. Before the registering of the sample into Medfiles' system, the sample should have a request of analysis or an analysis plan which informs the needed analyses and specifications if provided. Samples' condition is also checked for there should not be anything abnormal, for example the primary packaging has to be undamaged. (Oy Medfiles Ltd., 2019b)

The samples are registered into Medfiles' system after the pre-checking is done. Each sample receives a sample code. The sample is registered to the laboratory information management system (LIMS) and the sample code is generated there. (Oy Medfiles Ltd., 2019b)

The samples need to be stored in right conditions before, during and after analysis. The right conditions are individual for each sample and are stored according to the customers' requests. (Oy Medfiles Ltd., 2019b)

Physical data of sample analysis is recorded in worksheets, laboratory notebooks or other valid work related documents immediately after the specific action has been performed. The results of the analysis are recorded into LIMS. The results are reviewed from the preparing of the sample to the possible analytical methods and their performance. LIMS generates the certificate of analysis, which can be sent to the customer after the approval of the analysis results. (Oy Medfiles Ltd., 2019c), (Oy Medfiles Ltd., 2019d) The data that is generated concerning the sample, for example the request- and the certificate of analysis are archived and the samples can be disposed after agreed time (Oy Medfiles Ltd., 2019b).

8.1 LabVantage LIMS

Information gained from the laboratory contains everything from the clients and samples to the analytical methods and results. Gathering and archiving this information is important, because it is the greatest part of laboratory's capital. There is a risk that the information is lost, if it is recorded only on paper or the laboratory is using 'hybrid' forms of documentation. Hybrid means that the laboratory uses paper and electrical systems together for recording the data. (McDowall, 2017)

LabVantage's LIMS is a system that can be used to gather and maintain all the information a laboratory produces based on the data integrity points: completeness, consistency and accuracy of data. (Software Point, 2016a)

LIMS was taken into full use in Medfiles' inhalation laboratory in the autumn of 2019 in order to overcome data integrity issues and simplify the workload of laboratory technicians and experts. With the help of LIMS for example all the laboratory information can be gathered into one database and the risks of calculation and data entry mistakes are reduced. (Software Point, 2016a)

In Medfiles' inhalation analytics laboratory the data recording system is still in hybrid form, but with LIMS and few new software the system could be fully electronic. LabVantage offers for example electronical worksheets and solutions so other software can be integrated with LIMS. (Software Point, 2016a)

8.2 Empower 3

Chromatographical analysis methods are used widely in pharmaceutical industry because of the good properties to measure amounts or detect them from the pharmaceutical product. The concerns regarding data integrity in the chromatography data systems (CDS) have risen due to the higher regulations. CDS, that fulfils the data integrity regulations produces huge amounts of metadata and the review process can be compromised. (Longden, 2017)

Empower 3 is a CDS, that is used in the HPLC and UHPLC analytics in Medfiles' laboratories and if needed, it can be used to perform all chromatographic applications. The software is provided by Waters and can fulfil the data integrity requirements according to ALCOA+. (Waters, 2019)

9 MACRO PROCESS

The macro process map was created with the help of a workshop. The workshop team was formed so that there were people with different specialities concerning the process. The first workshop was conducted as a live meeting and there were two laboratory technicians, two experts from the inhalation analytics laboratory and expert from the LIMS team. The group was extremely experienced in their fields and the first draft of the process map was created in less than two hours.

There was no prior process map specifically illustrating the processes in the inhalation analytics laboratory. The only process map generally illustrating Medfiles' laboratory processes is presented in Figure 10. The process map created in the process mapping workshop is presented in Figure 11 below.

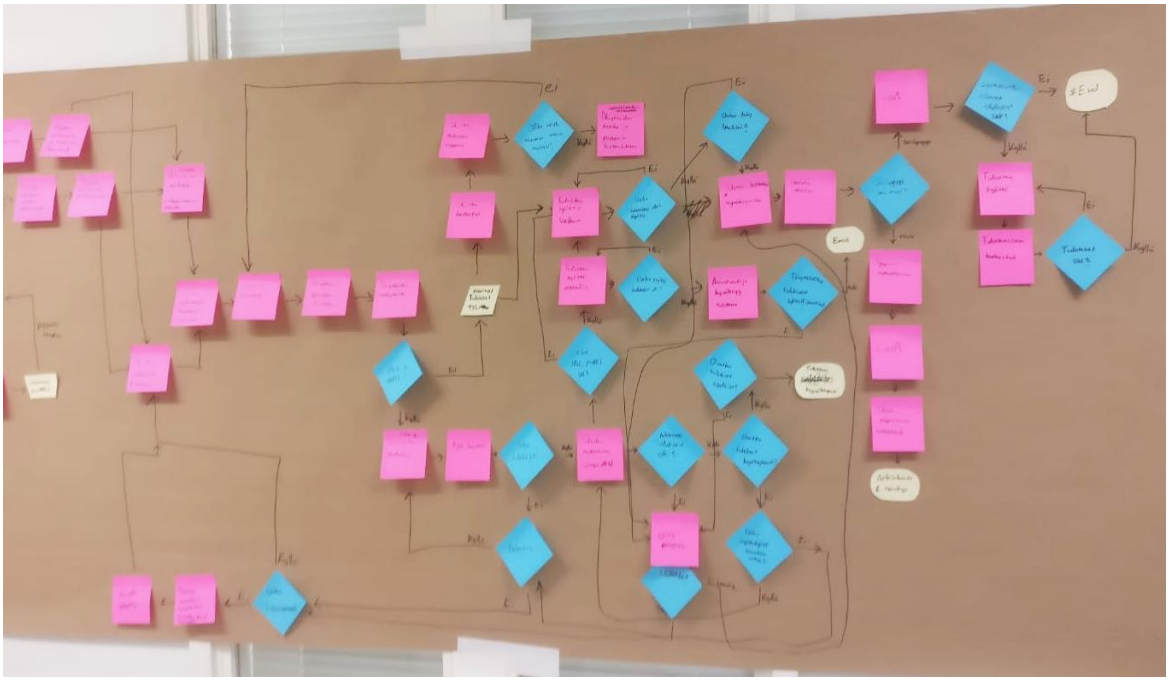

Figure 11. The map created in the process mapping workshop 14.01.2020.

The process map workshop began with a short presentation about the topic of the thesis and basics of process mapping, because the participants were not very familiar with the mapping process. Some of them had watched a short video about process mapping that was sent to them earlier. This preparation made the workshop efficient. The actual mapping process was started with defining the starting point and the end point which were *Sample arrival* and *Archiving and disposal*. The shapes used in the workshop are presented below in Figure 12. There was no need for too many different shapes, because the shaping could be done more efficiently in digital form.

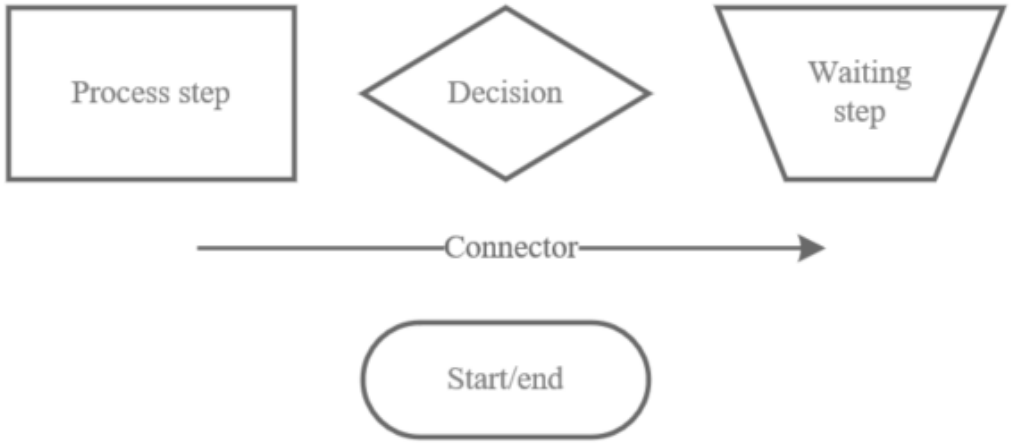

Figure 12. The used shapes in the first process mapping workshop.

The mapping workshop was a success and everyone in the team participated. The team was able to figure out many tricky mapping problems, like open process loops that were not leading anywhere or towards the wrong ending. The process map that was drawn in the workshop had four different endings and one of them was at first defined *Archiving and disposal*.

One of the most important things to remember in both paper- and software mapping is that the boxes are made and put in their right place before the connectors are drawn. This makes the moving of the boxes possible.

After the process map of the inhalation analytics laboratory was finished with the mapping team, the map could be transferred to a digital form. This was done with the Visio Professional provided by Microsoft. The digital map was drawn into a swimlane form, where the process flows from lane to other. The swimlanes from top to bottom are *Sample registering*, *LIMS*, *Document actuary*, *Expert*, *Laboratory technician* and *Empower 3*.

The process was also divided into four different phases according to the general laboratory process of Medfiles' presented in Figure 10. The phases chosen were *Sample arrival and registering*, *Analysis, reporting and review of results* and *Archiving*. This is the actual general process 'map' of Medfiles' laboratories and can be used as the mapping frame of other processes. Below in Figure 13 the beginning of the process can be seen on a digitalized map.

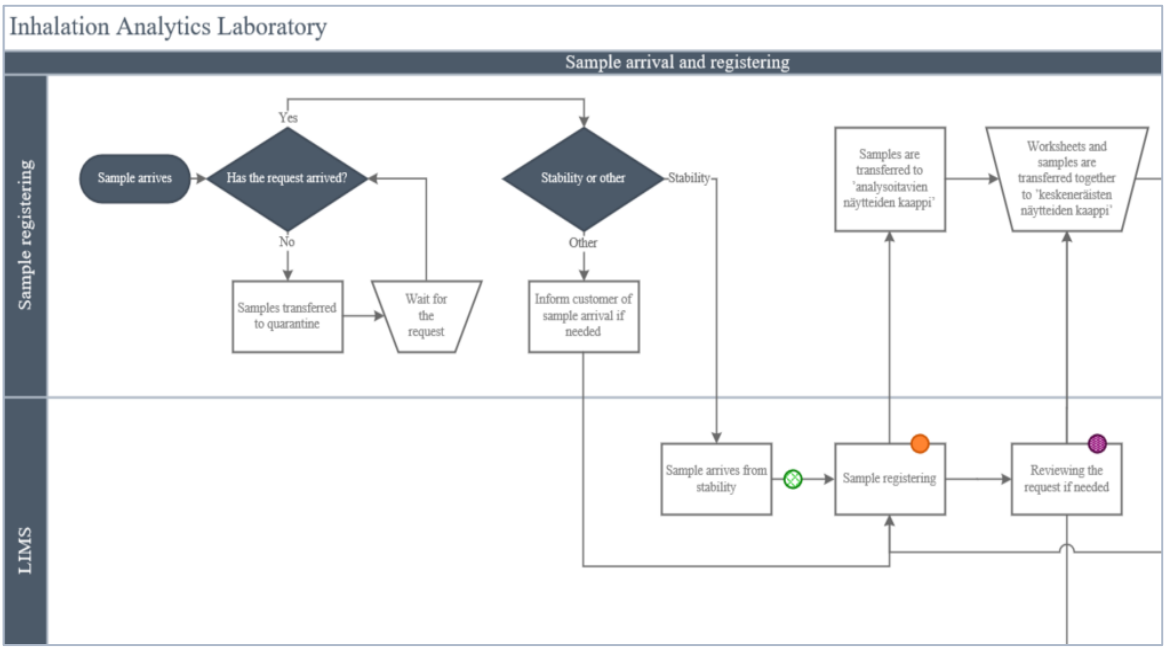

Figure 13. The beginning of the first process map of Medfiles' inhalation analytics laboratory. The first phase, *Sample arrival and registering* can be seen and the two first swimlanes, *Sample registering* and *LIMS*.

The mapping on Visio was done by one person, the facilitator of the workshop. The mapper should be someone who was part of the process mapping workshop, has knowledge of the mapping and knows how to use the process mapping software. (Harris, 2019)

The important things to remember in the mapping with software are that the outcome should resemble the map created in the mapping workshop and should be clear and understandable by other employees. (Kalman, 2002)

The mapping with Visio was done in two days. The final map contained one start, five endings, 20 decisions and 31 process steps from which two were 'waiting' steps where the process could not go on until something happened for example *Wait for the request* in Figure 13. There is a step where the process cannot go on until the analysis request has arrived. The final digitalized map was presented to the laboratory manager, laboratory director, QA (Quality Assurance) and inhalation analytics' team leader in order to find out were there any inconveniences or was the map understandable by the people who will eventually use it. There were a few notes about the missing arrows, steps in wrong swimlanes and a few term related problems, but all were easily solved.

One of the major changes to the map was the addition of indicators to the lanes *LIMS* and *Empower 3*. The problem was that in these swimlanes the role of the performer of the process step was not clear. Four different indicators were added to the map and these can be seen below in Figure 14 and in the map in Figure 15.

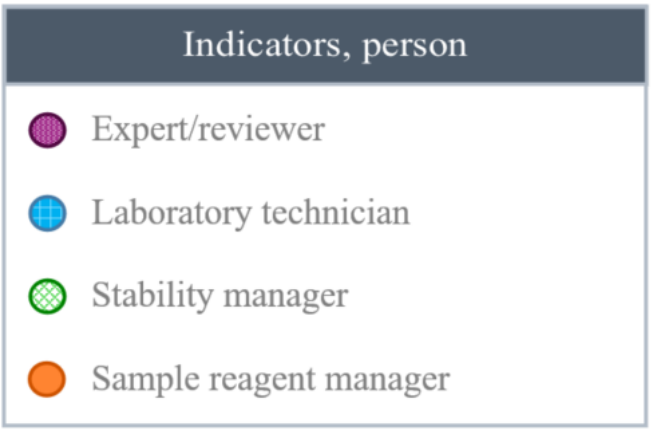

Figure 14. The indicators of different roles in the process map.

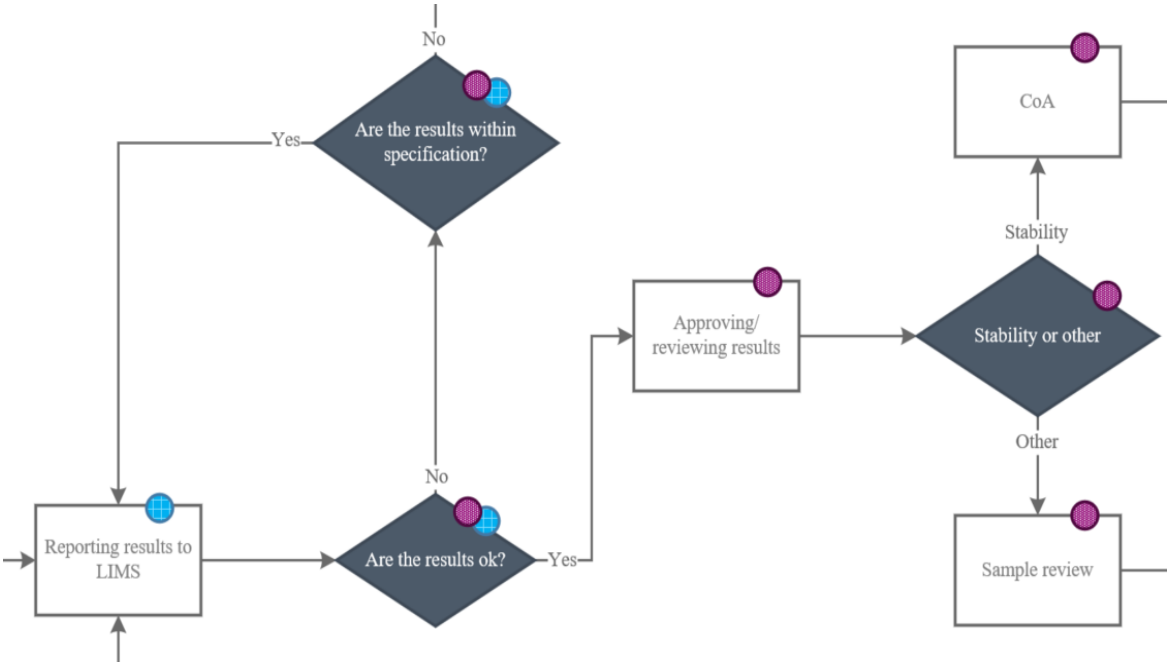

Figure 15. The example of the use of the laboratory technician (blue) and expert/reviewer (purple) indicators in result reporting and reviewing in the LIMS swimlane.

The roles presented in Figure 14 are the actual roles used in the Empower 3 and LIMS software. The stability manager and sample reagent manager are only part of the sample registering part of the process and these indicators can be seen in process in Figure 15.

10 BOTTLENECKS AND RISKS

The bottlenecks and risks found in the macro process map can vary from useless process steps to critical lacks in the process compromising the data integrity. It is crucial to find the most critical risks, but finding the bottlenecks makes the process more efficient. The minor risks should also be found in order to manage them even though the risks in question would be left as are. When even the small risks are acknowledged, their chance of creating a greater risk is decreased. This can happen for example when getting rid of a greater risk in the same process step. In this study, all the located risks were data integrity risks. (Rath, 2008)

Before starting the risk management process, the risks in the macro map needed to be found. This is the second part of the process. The risks were located from the map by a new workshop group. The found risks are presented in Appendix II. This workshop's team was formed as follows: chemist who was the same person as in the process mapping workshop, laboratory manager, QA, development engineer and me. The risk location was done in two sessions. The group was familiar with the ALOCA+ and this made the process more efficient.

The working method was following. The data integrity definition, ALCOA + contains total of nine terms, which are explained Chapter 3. It was decided that the easiest way to locate all the risks was to go through the whole process nine times changing the approach to each process step according to different ALCOA + term. This way of working was extremely efficient and the workgroup was able to complete this part in four hours.

In Figure 16 below the three most compromised process steps concerning the data integrity are presented. There are many data integrity risks in this part of the process because these

steps are done by humans and all the actions are written to worksheets with a pen. These three steps also contain a lot of different subprocesses and tasks.

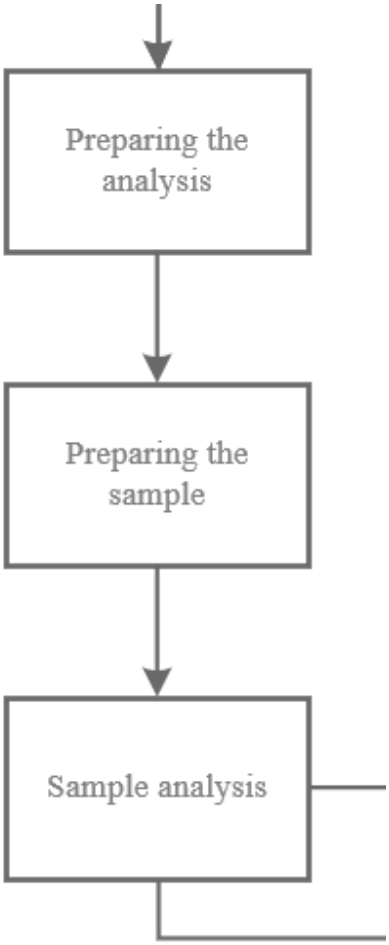

Figure 16. The analysis phase starts with these three process steps.

The part presented in Figure 16 was the most problematic part of the inhalation analytics laboratory's macro map. Almost every aspect of data integrity was somehow compromised here though many of the located risks were small. It was clear immediately that these steps needed visualised subprocesses.

During the first risk locating session the whole ALCOA was gone through the map and there were only the additional four terms from ALCOA + left for the second meeting. It was noted that the process steps done in the LIMS and Empower 3 swimlanes had just a few

compromised data integrity aspects because the software have been validated and checked to fulfil the data integrity requirements.

It should be noted that the following chapters discuss different risks and many of them can happen due to an intentional forgery of the analyses and results etc. The companies should have clear policies concerning forgery and that lying is not tolerated in any form. Medfiles follows policies from Pharma Industry Finland (PIF) – *Code of Ethics* and trust between the employees and employers is established (PIF, 2019).

Attributable

The first ALCOA term is attributable. It means that the creator of the data can be tracked, and the date and/or time of creation and each action is available. If the data has been modified, the modification date and name of the modifier are also available. (McDowall, 2019a)

Many of the process steps in Medfiles' inhalation analytics laboratory contain risks compromising the attributability of data. This is mostly because of the hybrid system. Software that are used in the laboratory save the dates and names of each action but if the action is recorded on the paper, the name and date can be forgotten or dated or named wrong. In the paper system, it is impossible to remove the risk because of the possibility of human error.

There is also a risk in sample request that cannot be removed by Medfiles. The request arriving from the customer is not always in regular form and may be missing some important information, for example the batch number. Sometimes there might be some handwritten information or calculations that can be unclear or wrong. The risk can be minimized by going through the request thoroughly and recalculating the results if possible. If there is unclear information, the customer should be contacted.

Legible

Legible data is readable and it can be understood only in one way. The action that the data describes can be redone without any misunderstandings (McDowall, 2019a). It should be noted that the language has to be understandable to authorities (Price, 2017).

Legibility risks occur in the parts where the work is recorded by a human and by handwriting, as is the case in many of the data integrity risks. The data integrity is compromised, when it is recorded with poor handwriting that is not clear, or the words or sentences used are not unambiguous.

The data integrity risks concerning legibility were found in the analysis part and in worksheet review. These risks cannot be fully removed when there is something done by human and by hand. Nothing seemingly critical came up in this part.

Contemporaneous

Contemporaneous means that the data, for example the result of analysis is recorded at the same time as the action is performed. The actions should have time and date stamps. Back dating should not be done. (McDowall, 2019a) It is not always possible to record the data at the time of the action, but the delay between the action and recording should be as short as possible and well justified (Price, 2017).

Again the data integrity requirements concerning contemporaneousness are hard to fulfil in the parts where the actions are made by human and by hand. The written action, date and time can be what ever and there is no way to prove when the action has actually been done. The only way is to believe the written date and name to be true. Usually, the exact time is not recorded in worksheets excluding a few exceptions in Medfiles' laboratories.

In addition to the analysis part of the process, a few risks exist in the result calculation approval. These risks have been removed faster than were anticipated due to the calculation transferred to the LIMS system.

Original

Original data can be verified to be original. Data is original, when it cannot be modified or deleted without a trace and the original data is left to be reviewed in its original form. The original version should be easily available. (McDowall, 2019a)

The most significant risk was found to be an originality risk of a print. In the laboratory in order to perform different analyses, some samples need to be weighed. The balances in the inhalation analytics laboratory are not connected with any software that records actions. This means that the weighing print cannot be assured to be original. The printer connected with the balance prints the date and time of the printing event, but the results are not recorded anywhere else. There is a requirement that the print has to be taped into the worksheet or if it is part of the calibration event, the print has to be taped into the instrument notebook. This means that it is possible to throw the original print away and do the weighing again.

Originality in worksheets can be assured with the verification of them and all the entries need to be done with the ballpoint pen that cannot be removed. It is not allowed to copy the worksheets and this can be managed by the worksheet approval process. The adding of information on the other hand is extremely hard to control. For example, if the date is in form 1.1.2020 there is a possibility to add zero to alter the number 1 to 10. Text can also be added and corrected without noticing if there is space. The year 2020 is also tricky. If the date is written in form 1.1.20, the year is really easy to change into 2019 or other.

Accurate

Accurate is the last term of the basic ALCOA. Accurate data is error free and should not be modified. In the case, the data needs modification, the original data, date and name should be available. The reason for modification is also important to be known. (McDowall, 2019a)

Again, accuracy problems and risks were located in the process parts where the action is performed by a human. The accuracy of raw data is at risk when there is a possibility to forget something, for example recording the dilution of the sample, which is very crucial information and cannot be added later with certainty, that the added value is right.

The data accuracy risks were also found in the *Sample registering* and *Archiving and disposal*. The sample reagent manager registers the sample to LIMS but there is a possibility

that there is some information missing, for example the batch code. In the worst case scenario, the missing of information is not noticed and the analysis can be done from wrong samples. This is still very unlikely to happen.

In archiving the data and analysis results, there is a possibility that the sample code is miswritten, and wrong papers are linked to the wrong sample. If the linking of sample code and its analysis papers is wrong, it could cause a problem in finding the correct raw data also with other samples. This is because there may be several samples in the same injection sequence of the HPLC or UHPLC run and the raw data can be located only with the one sample. If this sample has a wrong data archived with it, it is almost impossible to locate the raw data.

Complete

Complete is the first of the four additional terms to ALCOA (ALCOA +). Complete data is all the data that has been acquired before, during and after the analysis. In hybrid systems, like Medfiles', the data should be linked together. (McDowall, 2019a)

The completeness of data is at risk every time the data is transferred from one system to another or from paper to a system or the other way around. In Medfiles' laboratory this creates many different problematic process steps because the data has to be transferred quite many times. Completeness of the data may be the biggest lack in Medfiles' laboratory. The data transfer can be easily indicated from the process maps. Usually the transfer is done when the process flows from one swimlane to another. This is presented in Figure 17 below.

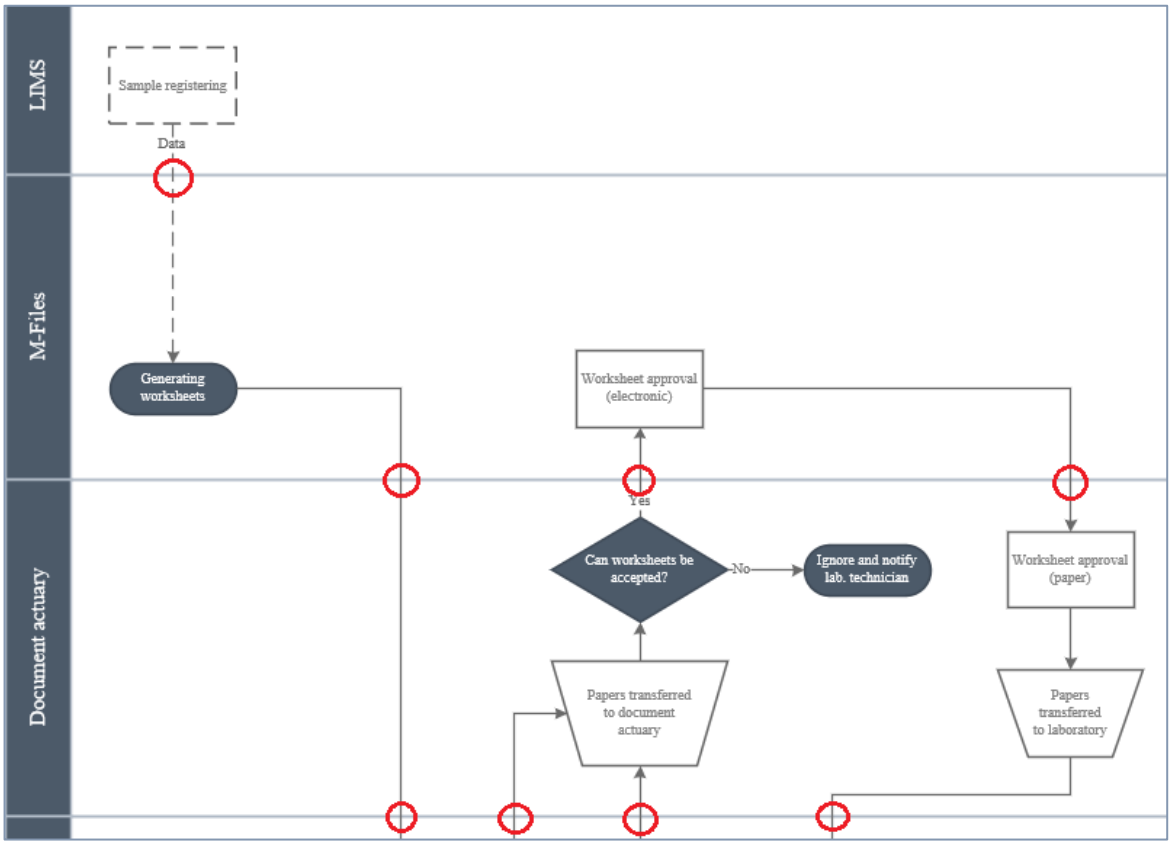

Figure 17. Data completeness may be compromised each time the data is transferred from one software to another, from paper to software or other way around.

Over all the data is first transferred to Medfiles' LIMS software. This data is the information about the needed analyses. The data of the sample preparation is transferred to Empower 3 for result calculation and then the results are transferred to further calculation, that cannot be completed with Empower 3. The data is also transferred to LIMS. The data can be transferred safely concerning the data integrity, if the electronic systems are integrated together and the data transfer is verified to work accordingly in the validation process. This is possible for example between LIMS and Empower 3 (Soares, 2019).

In the laboratory there is a problem with transferring the information from the expert to the laboratory technician. When the sample arrives, there is sometimes no place where the laboratory technician can check how many parallel samples need to be analysed. There usually is also some additional information or guidance about the analysis method that only the responsible expert knows. The information can be found from the Medfiles' SOPs and working instructions or if there is specific or varied information it should be in the request of analysis, but it can also be obtained from the client via e-mail or Skype message. This is

problematic. There should be a verified list somewhere containing all this information and it should be available to laboratory technicians easily.

The integration between LIMS and Medfiles' document management system is needed in order to generate the laboratory worksheets, manage archiving and write OOX- (Out of Specification, Expectation, Trend) and deviation documents. The document management system manages the worksheets and the creation and approval is done through the system. The data is entered to LIMS and the document management system gets the information needed from the LIMS database. The refreshing gap is the problem that creates the risk. The refreshing gap means the time between the automatic data transfers from M-Files to LIMS and other way around. If something is changed in LIMS, the new information updates to the worksheets only once a day and if the worksheet is printed before the update, the information on the worksheet differs from the information in LIMS.

Consistent

In consistent data, the events are in order. This means that dates and times of actions are recorded in order to arrange the events (McDowall, 2019a). The primary instance of data is important to be defined if there is more than one source (Price, 2017).

Consistency of data seemed not to be a great risk in the inhalation analytics laboratory. This risk showed again in the process steps where the data is recorded by hand. In an electronic system, all the actions are recorded automatically with the time and name stamp and can be put in order. In the laboratory, the work is recorded to the worksheets, that lead the work really specifically. The problem is that the worksheets are so specific, that they actually work well only for one specific inhalation analysis client. For other clients, the specific worksheets won't necessarily work. In this case the general worksheets, that contain only lines to write in, are used. There is also a possibility to generate specific worksheets to other products and clients, but even though it is rather easy, it is time consuming and the design and content require efforts.

Enduring

Enduring data is recorded into official forms. These official forms include for example notebooks or worksheets that are verified and can be located (McDowall, 2019a). Data on paper or electronic should be stored so it is safe from loss and degradation (Price, 2017).

In Medfiles' laboratory the data can be considered being enduring. All used documents have been approved and can be located. The markings are done with archival quality pens. Only problematic part is the archiving and specifically the electronic archiving. Samples' raw data on the inhalation analytics laboratory is created and stored in the Empower 3 system. This system is validated and the data is backed up and checked regularly in case of corruption. In LIMS the certificate of analysis is stored in PDF form that is in archivable form.

Available

Available data can be reviewed and is easily accessed. The data should be available for its defined lifetime in every phase (McDowall, 2019a). This part of the ALCOA+ concerns also the third-party suppliers and/or service providers (Price, 2017).

Availability of data in Medfiles' inhalation analytics laboratory is only really a concern in archiving and disposal part of the process. Because of the hybrid system that is used in the laboratory the archivable documents are both on paper and in digital form.

The paper documents are stored in the Medfiles' own premises for a certain period and transferred later to subcontractor's facilities. After this, the documents are not so fast to retrieve physically but there is a possibility to ask a PDFs from the documents the same day. There has been discussions that the paper data would be copied and after that could be seen as PDFs from document management system.

Availability may be challenging for data older than 10 or 15 years when there is a need for retrieving older electronic data. Employees may notice that the data is hard to open due to the data format. One challenge is also the file path which may cause difficulties finding the data. Validated electronic archiving software would help in finding the specific data and ensure the data to be in readable form for its whole lifecycle.

11 SUBPROCESSES

It was noted by the risk workgroup, that there were two parts of the process that needed further mapping. These were *Worksheet approval* and *Preparing of the analysis and sample and sample analysis*. The subprocesses were mapped to the macro process map in order to find the causes of the risks. The subprocesses were mapped with the help of two different workgroups. The first workgroup mapping the *Worksheet approval* consisted of the ICT Manager and two LIMS experts. The other workgroup mapping the *Preparing of the analysis and sample and sample analysis* was the same as in the macro map workshop. The facilitator was in both groups the same as in the macro process map workshop.

11.1 OOS process

The OOS means Out of Specification result and is defined by ECA Academy (2014) followingly: '*A result that falls outside established acceptance criteria which have been established in official compendia and/or by company documentation*'. The OOS result is simply a reviewed result which does not conform the acceptance criteria. (Oy Medfiles Ltd., 2018)

The OOS, or OOX process in Medfiles' laboratories is a process that has been specifically defined in Medfiles' SOP and is quite complicated. In the risk workgroup it was decided that this process would be left out of this thesis because it is specific and a kind of standalone process that can be assessed on its own. There were a few risks concerning the OOS process but these risks have already been assessed earlier in other cases.

11.2 Worksheet approval

The first process map part to be mapped further was the *Worksheet approval*. This process step was chosen, because the role of document management system was important to be seen in the process. The process was chosen to visualise the whole lifecycle of the worksheet from generation to archiving. After the chosen archiving period, the worksheets would be disposed accordingly. In the left top corner of the map is a different start for the process than normally. There is a LIMS swimlane where the data is generated in the *Sample registering* process step. Because this is not the actual start of the process for the worksheet lifecycle it is highlighted with dotted line and the connector of data is also dotted. This part visualises that M-Files (Medfiles' document management system) gets the data for the worksheets from the LIMS database. This process step is presented in Figure 18 below.

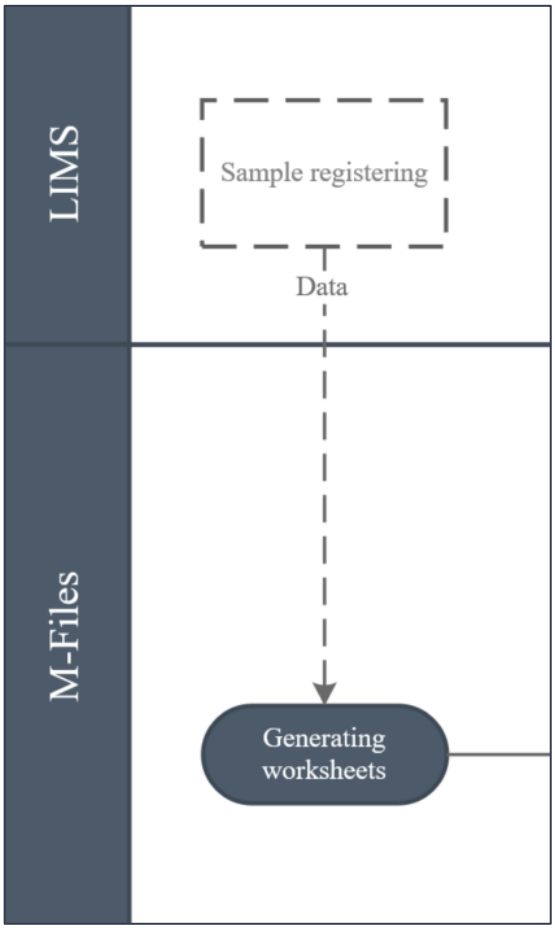

Figure 18. *Worksheet approval* dataflow between M-Files and LIMS. M-Files gets the information needed for generating the worksheets from LIMS database.

The whole worksheet approval process that was mapped in the workshop is divided into four swimlanes: *LIMS, M-Files, Document actuary, Laboratory technician* and *Expert*. The process contains three phases: Sample arrival and registering which is presented in Figure 19 and *Analysis and reporting and review of result*s and *Archiving*. Two latter phases are presented in Figure 20.

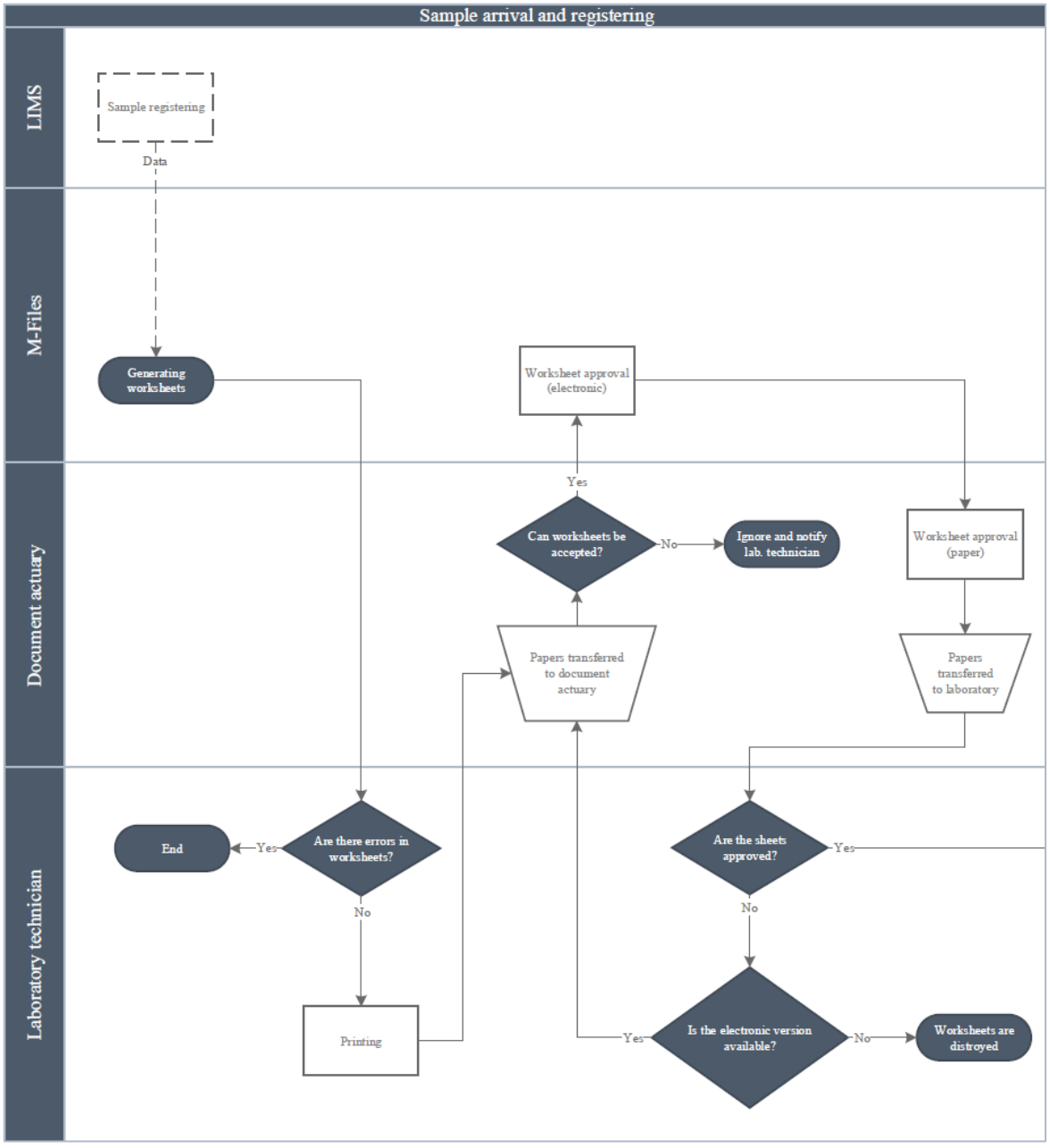

Figure 19. *Sample arrival and registering* part of the worksheet approval process.

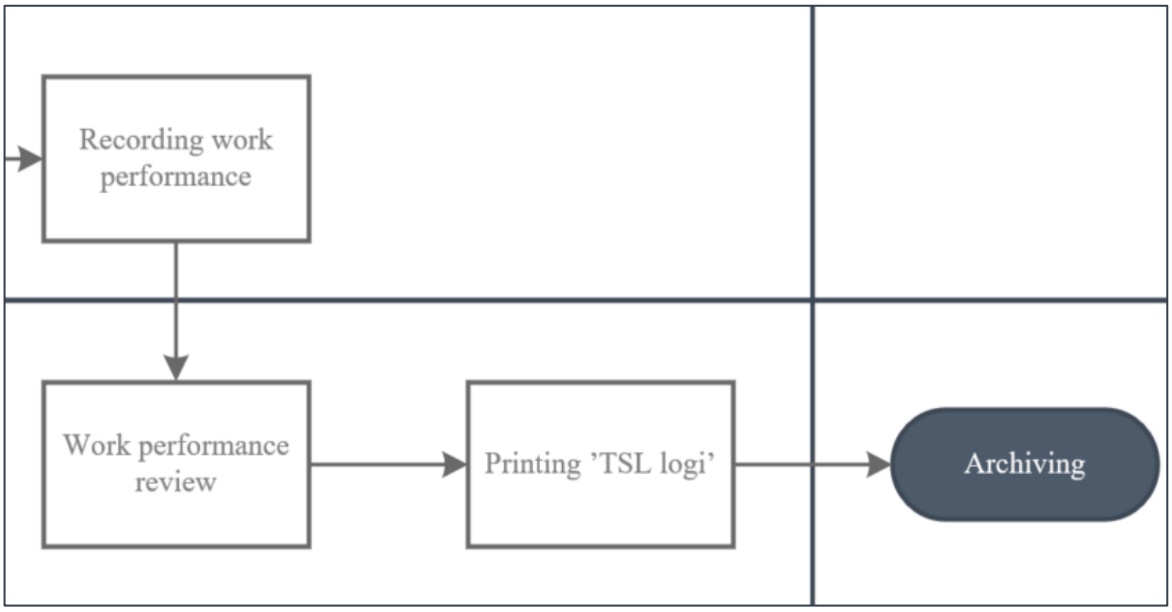

Figure 20. *Analysis and reporting and review of the results* and *Archiving* parts of the worksheet approval process. The two swimlanes are *Laboratory technician* (above) and *Expert* (below).

The worksheet approval process was assessed with the help of the workgroup working on mapping it. The main problem areas can be found in the parts where there is a possibility for a human error. The archiving process contains also a few different risks already discussed in Chapter 10.

The most alarming risk was found in the laboratory technician responsibilities. This risk was discussed because it actually had happened a few weeks earlier prior to the discussion. The risk was that if the laboratory technician forgets to check the approval of the worksheet and it is not approved, data is recorded on the paper that is not trackable or valid. The data integrity is compromised because data is not enduring any longer. If the worksheet containing the original data is declined after the analysis because it is not valid, the originality of data is compromised. In this particular case, the worksheet was not any longer available in Medfiles' document management system, because if the worksheets are not approved electronically within 30 days, the digital worksheet is deleted by the system and cannot be retrieved. This risk was reduced by corrective and preventive actions (CAPA) regarding the deviation generated during the happened accident.

Other notable risk is the actual approving process. The document actuary has to check, if the document can be approved and then approve it first in the document management system. After the system generates the actual date and time stamp for the approval, they are

transferred by hand to the worksheets. It is important that this process is done in this order, because there is a possibility (and actually this has happened) that the electronical approving is forgotten and then there is no linking between the paper and electronic data. This creates problems again in the archiving of the data and endurability of data may be compromised.

There is also a problem with the state transition of the document in the document management system. There is a possibility to transfer the state of the document to 'Unnecessary' by accident. This has happened a few times and the error is noticed only when the worksheet log is printed at the end of the process. The risk of the data being complete is in this case compromised, because there is now an error in hybrid system linking. Also the endurability is not achieved, because the accepted, but (accidentally marked as) unnecessary worksheets are used. There is no way that the laboratory technician could notice this without going to M-Files and checking, if the document is in the right state.

The actual benefit of mapping the subprocesses can be seen very clearly here. When assessing the risks with the risk workgroup, the group was able to name risks concerning completeness, consistency and endurability of data. After mapping the subprocess of the worksheet approval, the workgroup could find risks also concerning contemporaneousness, originality and accuracy of data. These risks could have been missed if the accurate process map had not been created.

11.3 Analysis

The *Analysis* part of the process was further mapped because the risk workshop noted that there were too many different risks concerning this part of the process and it was hard to pinpoint the root causes of each one. There was also a concern that all the risks were not noticed in the risk workshop. There is a problem with the mapping of this part of the process, because there are so many different analyses and products flowing through this same process. The mapping workshop group chose to map the general process and not to go into too much detail. The map can be linked to the macro map from the beginning and the end of the subprocess map.

It was important to see the weighing in the subprocess because it was already known that there is a major risk with the weighing print. Because there is no software on balances that records the weighing events or the person doing the weighing, the data can easily be compromised. There is no possibility to prevent, that the weighing is done multiple times in order to make the weight of the sample falsely acceptable to make the sample result better. There have also been times when the print is unreadable because the printer ink did not work properly.

The process was divided into two phases which were *Preparing the analysis* and *Preparing the sample and sample analysis*. The process flows in two swimlanes: *Laboratory technician* and *Empower 3*. There is only one process step done in *Empower 3* swimlane. The first phase is presented in Figure 21 and the second phase is presented in Figure 22.

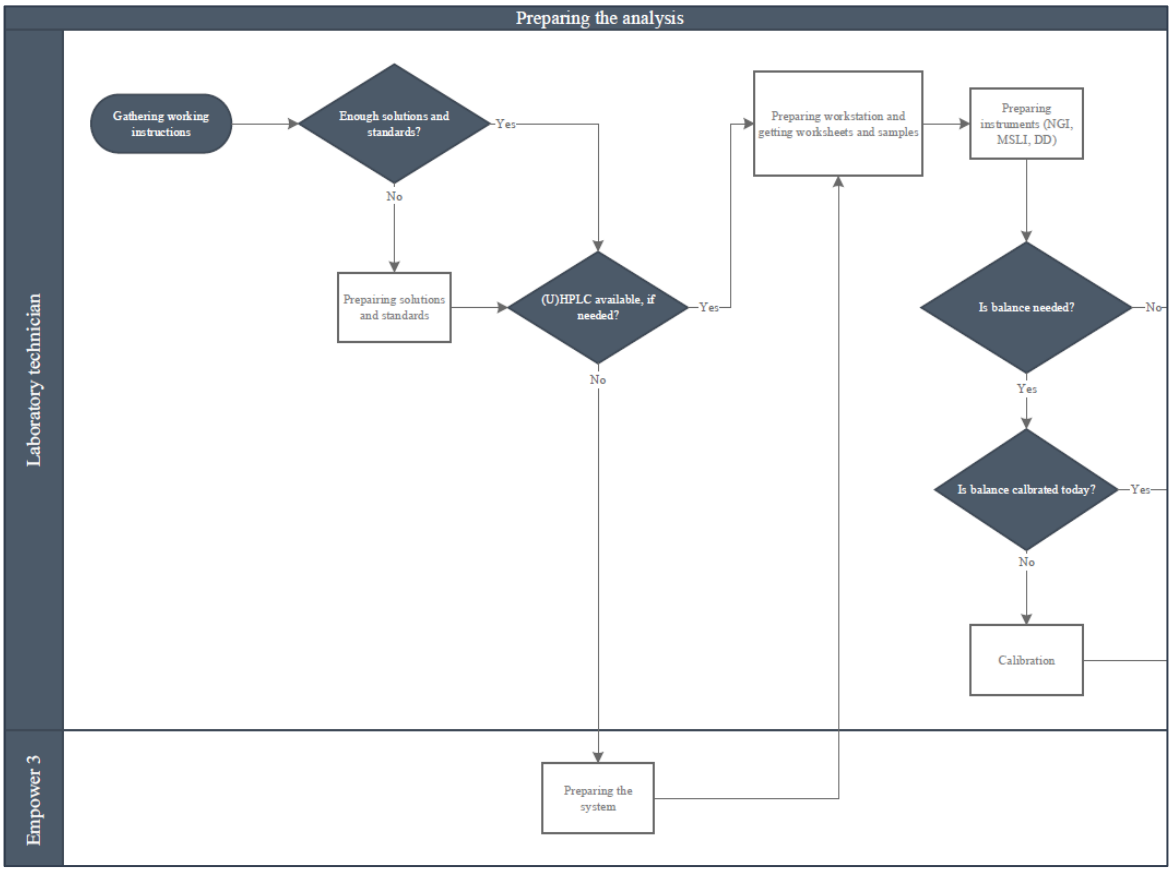

Figure 21. *Preparing the analysis* part of the analysis process.

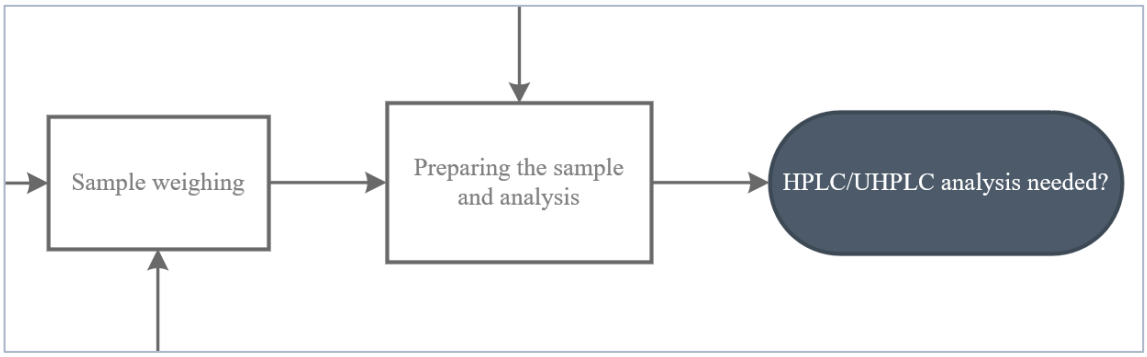

Figure 22. *Preparing the sample and sample analysis* part of the analysis process. These process steps are done in the *Laboratory technician* swimlane.

The risks of this part of the process were possible to locate into a specific process step and the same risks were found this time as in the first mapping workshop. The most problematic thing about the analysis is that is entirely done by a human and where there is a possibility for a human error there is also a risk that someone forges the results. These kind of risks cannot completely be removed. It is extremely important to have a trustworthy personnel in the company and to have an open environment in order to encourage employees to tell if something went wrong.

12 RISK MANAGEMENT

Risk management consists of three phases explained in Chapter 4. These phases are risk assessment, control and review and the process is ongoing until the end of the process map's lifecycle. In this thesis, the first phase, risk assessment also contains the risk prioritizing.

The risk management process was conducted after the creation of the macro process map, where the problematic parts and bottlenecks were identified. The risk identification is part of the risk assessment process, so the risk management was actually started already in Chapter 10. The purpose of the risk management was to define what changes were important to be done in order to remove or reduce alarming risks.

12.1 Risk prioritizing and assessment

The risk assessment was the first part of the risk management process. The risk prioritizing was done as a part of this phase. The risk assessment was done by the same group as in which the risk identification was done. The risk prioritizing and assessment were conducted with the help of FMEA. These methods are presented in Chapter 4. The first step was to generate the wanted model for the process. The used model is presented in Appendix I.

After the FMEA model was chosen, the scope and topic of the risk management process were defined and the identified risks in Chapters 10 and 11 were entered into the model. This included writing down each process step containing the risk or risks and identifying *Potential failure mode, Potential effects of the failure* and *Potential causes of the failure.* The defined topic and scope are presented in Table IV below

Table IV. The defined topic and scope of the FMEA.

Risk assessment subject	Inhalation analytics laboratory
Risk question	Is the integrity of data maintained during the process in inhalation analytics laboratory?
General assumptions	Instruments are qualified, software are validated and fulfil data integrity requirements.

Next the RPN values for each risk were defined. This was done by defining the *severity, probability* and *detectability* values and calculating the RPN by methods presented in Chapter 4.2.

One mistake that was made, was that there was no expert in the workgroup that would have had specific information about the inhalation analytics laboratory process. Due to this mistake, some of the process steps had to be discussed independently with chemists. After the RPN values were calculated the distribution presented in Table V was formed.

Table V. Risk priorities of macroprocess, *Worksheet approval* subprocess and *Analysis* subprocess.

RPN	Number of risks	Percentage
≤ 3	27	57 %
4 - 14	17	36 %
≥ 15	3	6 %

From Table V it can be seen that there were many risks with priority less than or equal to 3, and those risks were just noted and accepted. No further actions were taken for these risks. There were only three risks with priority higher or equal to 15. The low amount of high priority risks indicates that in Medfiles' inhalation analytics laboratory, the data integrity is mostly in control.

12.2 Risk control

The risk control was done after the assessment. The control of risks consisted of two phases that were risk reduction and risk acceptance. The risks with priority number less than or equal to 3 skipped the risk reduction step, because the risks with such a low priority number can be accepted straight away. Many of these low risks would be reduced and removed alongside with the reduction and removal of the higher priority risks. For example, the use of electronic laboratory notebooks (ELN) would remove many risks from this table.

The risks with the higher priority number than 3 were processed through the risk reduction. This was also done with the help of FMEA. This was good because then the whole risk management process can be seen from the filled FMEA. The risk reduction phase included the considering whether the risk had some potential reduction methods or solutions. It should be noted that many of the recommended actions require a huge amount of different preparations, testing and investment, so the changes cannot be made straight away and may require many years delay depending on many different aspects.

The highest priority numbers were obtained from process steps *Work review, Archiving and disposal* and *Additional calculation*: *Result approval.* In the *Work review* process step the found risk was that in the laboratory the experts do not use the comments section at the end of each analysis in the worksheets. The comments section is there to be used as an 'audit trail' to see what corrections the expert or the reviewer has requested and if there is something to be noted in the work review. It was noted by chemists, that the comments section is not efficient, consumes time and is hard to use. For example, if there is just one needed correction. It is easier to just walk to the laboratory and ask the analyst to correct the error. The recommended action for this risk is to make the comments section better and to make the needed changes to the section after thorough enough research.

In *Archiving and disposal* there were two identified high risks. First was that the linking between the paper and the electronic data is not error free. If the linked sample number is wrong, there is no way to find the electronic data after the sample and its data have been archived. The probability of this happening is great, because just one misspelling or unclear handwriting can make the linking go wrong.

The process step *Result approval* of additional calculations included a high priority risk. This risk has been dealt with during the spring 2020 and removed by adding new LIMS calculations.

12.3 Risk review

The risk management process was completed with the risk review. After the FMEA model was filled it was sent for review to employees that took part in the risk management process. There are a few different questions that can be asked in order to make the risk review efficient and thorough enough.

While doing the risk prioritization, assessment and control it was noticed that the risk analysis group should have had a few other experts to take part into the meetings. Because of this, some of the risks had to be discussed separately with two different chemists. The use of FMEA was efficient in excel sheet. This way the risks and process steps can be sorted.

The FMEA should be already used in the risk identification part in order to make extensive notes about each process step.

The new risks were found mainly in the sub mapping in addition to the initial risk identification meetings. Few risks were defined somewhat differently in the risk management process but no new risks raised from the processes during the risk assessment or control.

There were a few risks that were considered high risks, but it was realised in the probability and impact assessment, that the risks were actually low. In the case of the *Work review* process step risk, it was not realised in the identification that the risk would be so high and during the assessment the risk was found to be one of the highest risks in the inhalation analytics laboratory.

The three highest risks could be reduced and probably removed completely by purchasing or updating software. In fact, most of the risks identified could be reduced by this solution. The comments section in the worksheets was taken into immediate processing and the risk in *Result approval* was removed by modifications in LIMS. All the other risks that were to be reduced or removed if possible, should be transferred into the CAPA process.

The updating and management of the FMEA and the process maps will be transferred after this work to the chosen process owner who will have the responsibilities mentioned in the process owner Chapter 5.3.

13 REMAPPING AND FINAL PROCESS MAP

The remapping was done after the subprocesses were mapped and the risk management process was finalized. The purpose of the remapping was to visualise the map again with the new ideas formed during this process. Different improvement ideas were noted from experts, management and the process mapping workgroups. The changes were made in order to make the maps more informative, accurate and easier to read for all employees using it.

The first improvement was to make the links to the macro map for the process steps that included sub maps. It was important for the maps' user to know what part of the process the sub maps visualised. The linking was done with new indicators that mean in general the linking between maps. The chosen indicators are presented below in Figure 23.

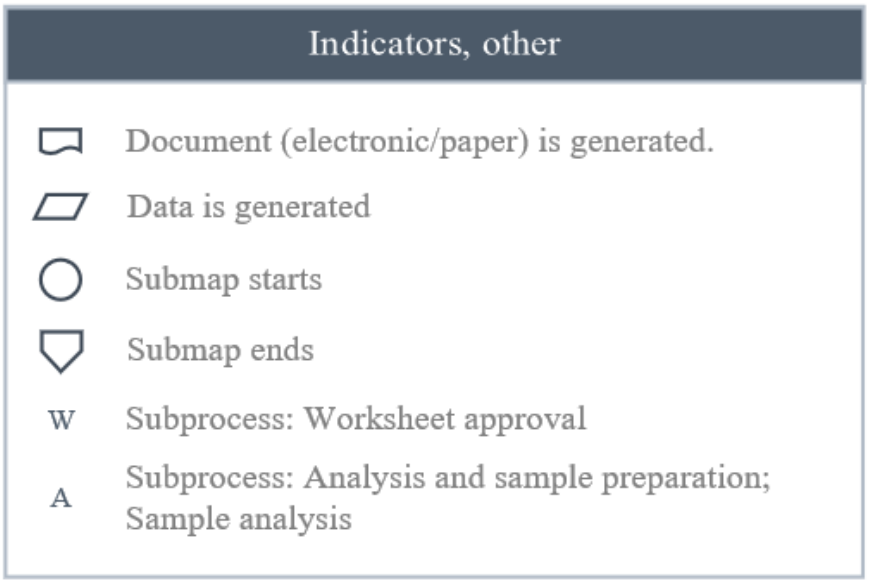

Figure 23. Added indicators to the macro map of the inhalation analytics laboratory process.

The chosen indicators were added in order to make the data integrity issues more visible. Also the sub map indicators were chosen to demonstrate where the sub maps take place. The document generation and data generation indicators are used also in the sub maps. In the process steps, where the data or documents are generated should be sorted out in specific detail. This is done in order to know that the generated data stays integrated. The example of the macro process map the use of indicators is presented in Figure 24 below.

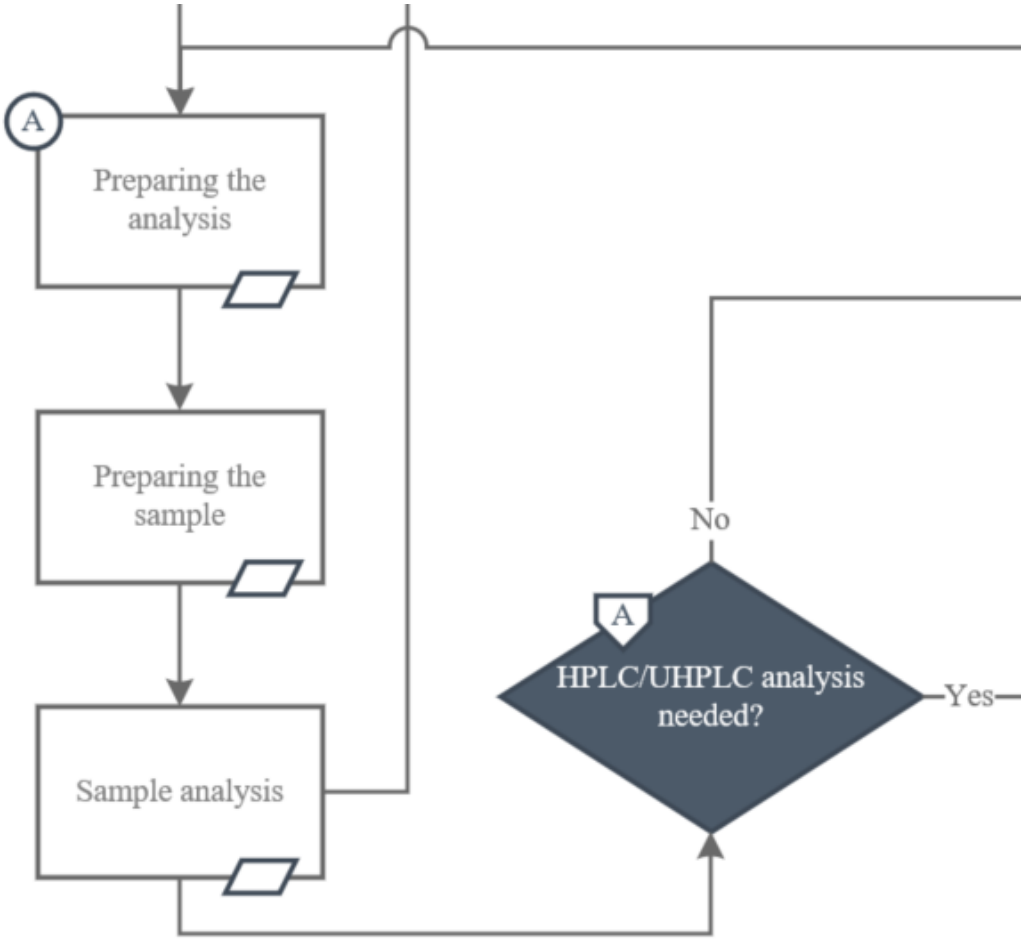

Figure 24. Generation of data and start and end points of subprocess: *Analysis and sample preparation*; *Sample analysis* presented in the macro process map.

The other small detail changed to the maps was the colouring. The swimlane and phase separator colour was changed to be slightly lighter in order to make the separators and connectors differ from each other to make the maps more readable.

The maps should be available for everybody within the organization and the results of the process should be presented to the laboratory personnel. (Harris, 2016) It was decided that the maps would be shared to the personnel so that anyone, who needs the maps, could find them. It was also decided that there would be a meeting or some kind of presentation of the maps after the thesis was done.

14 CONSEQUENCES AND NEEDED CHANGES

After the mapping processes were completed, there was the last part left to do. In order to the mapping process to have any purpose, the found risks, problems and bottlenecks needed to be processed. In the following Chapters, the most urgent corrective actions, recommendations and schedules are presented.

14.1 Worksheets comments section

The comment section was added to the worksheets in 2019 while I was also working as an expert in the inhalation analytics laboratory. The section was added there in order to generate a ‘physical audit trail’ to each analysis and a place where the experts could comment on the work performance. The problem was that the process of adding the comments section was too fast and did not consider all aspects. That is why the comments section is not useful the way it is. The comments section needed refreshment to make it more efficient and more user friendly. The workgroup for the open discussion was assembled and each expert from the inhalation analytics laboratory was invited to take part.

At the moment, the comment section is at the end of the longer analyses but there is no place for commenting at the end of the so called ‘small analyses’ for example odour or appearance. The problem was discussed in the group. It was clear that the problem was the time that is consumed in the comment writing. The analyses are recorded on worksheets and for example delivered dose analysis is six pages long. This means that if there are corrections for example on the first page, the expert has to go to the comment section to the end of the analysis and then write it down. This makes the correction process extremely time consuming and difficult.

After discussions it was concluded that there should be a comment section at the bottom of each page of each analysis, even the ‘small analyses’. In the worksheets all the boxes and

lines have to be filled or crossed out with a name and a date. The new comment section is presented in Figure 25 below.

Reviewers comments ☐ Yes ☐ No

Figure 25. Reviewers comment section at the end of each page of the worksheets.

In Figure 25 after the *Reviewers comment* there are two boxes to be filled. This is to make the box more user friendly. When the box is ticked 'No', there is no need for crossing the box out and writing the name and date. The comment section was tested with delivered dose analysis worksheets. There are over ten different analysis worksheets and it would have taken a huge amount of time to modify each of them just to notice that the new comment section was not working. The new delivered dose worksheets will be tested for a few weeks and after this the feedback will be gathered. The major concern is that is there enough space for the comments. The old comment section was also preserved at the end of the worksheets if there ever is a need for longer explanations. The new comment section will be modified according to the feedback and then taken into full use for all analyses. If the section is not working, new solutions will be discussed.

14.2 New software

Even though the new software is not always the answer, is should be noted that in order to make the laboratory paper free and get rid of the hybrid system, some updates for the existing software are needed. LabX is software that could connect the laboratory balances and titrator used in Medfiles' inhalation analytics laboratory. The program could remove the risk of balance and weighing data loss and there would always be an electronic signature, time and date stamp of the weighing event.

Electronic laboratory notebook, or shorter ELN is the alternative for the paper worksheets. Other good solution would be the laboratory execution system (LES). ELN and LES are add-ons to the LIMS which is already in use in Medfiles' laboratories. The use of these could be a huge step away from the hybrid system, which Medfiles' laboratories struggle with. With the help of the add-on, the laboratory could get rid of the worksheets and all the important information could be found via the software. The ELN or LES could be built for our use so that the software would calculate the needed calculations and check some work performance automatically. For example, has the used instrumentation been right for the specific analysis? The new add-on would also eliminate the possibility of forgetting to make some entries for example the signatures and dates and everything that is defined to be mandatory information, because the software generates them automatically. (Software Point, 2016b), (Software Point (2020)

The difference between ELN and LES is that the reporting to ELN is not so strict and in LES the reporting of the work performance would be more similar to the paper worksheets that are used in Medfiles' laboratories. LES is designed to record the work performance that is predesigned and has to be done according to the precise instruction. This is why the LES may be better solution compared to ELN. (Software Point, 2020) It should be noted that these are not the only solutions and there is a possibility to add an electronic laboratory notebook to Medfiles' system from other service providers. ELN and LES are the solutions from Medfiles' LIMS software provider.

There is a possibility to connect all the important instruments together in the laboratory. Important instruments are those that affect the analysis and produce data. The software called LabX is used in one of Medfiles' laboratories. With the help of LabX Medfiles' inhalation analytics laboratory could get rid of most of the data integrity risks concerning balances and weighing. The software could store data from balances and titrator and have authorised user access control. (Mettler Toledo, 2016) With the help of LabX, the laboratory's data integrity could be enhanced. (Adam, 2018)

The archiving in Medfiles' laboratories needs a fresh perspective and should be updated together with the new electronic archiving system. The archiving system should be acquired to Medfiles' laboratories as fast as possible. It may be the most important software purchase in order to increase the data integrity and get rid of the greatest data integrity risks within

Medfiles' inhalation analytics laboratory. There has already been different software under review but no decisions have been made.

14.3 Integration of software

Many of the risks concerning data integrity in the inhalation analytics laboratory could be removed by getting new software and having an integration between them. This is because when the data is transferred from software to another or from paper to the software, there is a possibility of data changing due to the entering of wrong values or in wrong form, for example changing masses from mg to g. If validated integration between the software existed, the data could be transported via software without a need for checking the data transfer.

Software developers provide different integration or connection add-ons to their software. With the help of these add-ons for example LabX can be integrated into LIMS and data can be automatically transferred via the connection. (Mettler Toledo, 2007)

There is also a possibility for integrating LIMS to Empower 3 so the CDS could get the sample masses straight from the balances (Mettler Toledo, 2007). With the connection, all the chromatographic raw data could be transported automatically to the LIMS for result calculation and reporting. (Software Point, 2014)

There are many different benefits in addition to already mentioned time saving and error free data. Integration of software improves the laboratory productivity and efficiency. There is no more time consuming data transfer checking which can also be very challenging due to a huge amount of number data. When there are no errors, there is nothing to be corrected and time is saved also for laboratory technicians. There have been discussions about the integration of the Medfiles' software but no decisions have been made. The integration of wanted software should be done by a defined schedule within a few years in order to increase laboratory's data integrity. (Curtis, 2019)

15 OTHER USES FOR PROCESS MAP AND FUTURE

The process map constructed in this thesis has also other possible uses that have come up during this project. As said at the beginning of this thesis, there is a wide range of reasons why different organizations want to do process mapping. The reasons vary from process intensification to buying a new instrument to the process or in order to identify risks within the process. The main reason for all this is to use process mapping as a tool for visualising the process. Without process maps, the process could be extremely abstract. The process map also helps to see who is responsible for what action. (Cooper, 2017)

The map created during this thesis is modified to visualise the process, so the data integrity risks are easier to find. The purposes of the map are infinite. Medfiles has three laboratories in Kuopio with different specializations. Even though the laboratories aim to have the same operating procedures, they vary quite a bit. Each laboratory has their own personnel and are located in different addresses. This easily generates differences between the operation procedures. With the help of the maps that could be generated to each laboratory, the processes could be more unified.

Now that the process map for the inhalation analytics laboratory has been drawn, there are uses for it in the form as it is. For example, the training and introduction of new staff can be done with the help of the maps. If a new product is accepted to the laboratory, the adaptation is easier with the help of the map and new instruments and process steps can be developed or added easier to the process. (Cooper, 2017)

There were some discussions about the future of the process mapping in order to find data integrity risks within Medfiles' processes in the Kuopio laboratories but also in other departments before starting this thesis. As the process mapping was done it was noticed that the maps have many different other efficient uses and possibilities as listed above. For example there have been two SOP updates where the process mapping was used to visualise the validation of computerised systems.

The maps generated for this laboratory and for the SOPs have gathered positive feedback, because the complicated and long processes are quite hard to understand without visual help. The mapping of the processes should be done at least for each laboratory in Kuopio in a

similar way as was done for the inhalation analytics laboratory. The maps could be done also to visualise different laboratory maintenance processes for example for instruments.

The list for good use purposes goes on but it is important to remember that the maps need to be managed and their purpose has to be fulfilled. For example if the map's purpose is to identify risks, the risk management has to be ongoing till the end of the process' lifecycle. This means that there should be employees, who manage the lifecycle of the processes and the maps. There should be an appointed process owner for each process. The process owner is responsible for mentioned tasks.

16 CONCLUSIONS

This thesis was done in order to study the data integrity risk identification with the help of the process mapping method. The experimental part of the thesis was conducted in Medfiles' inhalation analytics laboratory in Kuopio.

Process mapping method was studied thoroughly in the theoretical part of the thesis and different aspects concerning the GMP environment and regulations were discussed. The data integrity has been a major concern in the pharmaceutical industry due to various fraud cases and violations. The data integrity is so wide topic that there was an urgent need for an efficient risk management system including the risk identification in the Medfiles' laboratories and processes.

The process mapping has been a successful method for visualising the organizations' processes for decades and is still getting more and more popular. The problematic thing can be the distorted views about the topic, which can cause huge losses to the company when applied wrong. These can be for example the understanding that only the automatized system could remove the bottlenecks and that human errors are not tolerable. This is usually not the case. The process mapping method has to be efficiently studied and implemented to each organizations needs. The tools useful for one company can be time consuming for other.

When the process mapping and the data integrity risk management systems are unified, the outcome may be an extremely efficient management of the processes and the risks. When the maps are updated and new process steps are added or some removed it is a great moment to revaluate the data integrity in that process. Other uses for the maps can also be added and integrated to the map's maintenance process. Other possible uses could for example be identification of costly and inefficient process steps or identification of risks concerning employee safety.

The low amount of greater risks and amount of risks overall indicate that in Medfiles' inhalation analytics laboratory the data integrity issues are under control. The three most hazardous risks have been noted. One of these risks has been taken under reduction by worksheet modification and discussion but these actions have not yet produced any results. One of these three risks was greatly reduced and will be probably removed within a few

months because of new calculation procedures within the process. The calculation procedures were not changed due to this study but due to the LIMS development.

Most of the risks identified within the laboratory's procedures could be directly or indirectly reduced and even removed with some new software and integrations between existing and possible new software, though this would require investments. It should be noted that all the problems cannot and should not be solved by new software but with the help of these updates, the laboratory system could become fully electronic instead of hybrid, which it is now.

The process mapping method is a great and efficient way to visualise various processes and identify not only data integrity risks but other risks as well and find bottlenecks that slow down the processes and can be costly. The mapping of the processes has infinite range of applications and only the imagination and creativity is the limit.

REFERENCES

Abubakker, H.T., 2012, *Practical Guide to Creating Better Looking Process Maps*, Six Sigma Tools & Templates, accessed: 02.12.2019, available at: www.isixsigma.com

Adam, E., 2018, *Does Your Lab Lack the X Factor?* SelectScience. accessed: 27.02.2020, available at: www.selectscience.net

Cooper, C., 2017, *The Top 10 Benefits of Process Mapping*, Triaster Blog, accessed: 24.02.2020, available at: www.blog.triaster.co.uk

Cousins, M., 2018, *Process Mapping Techniques: 6 Methods to Create Process Maps*. Triaster Blog, accessed: 03.12.2019, available at: www.blog.triaster.co.uk

Curtis, D., 2019, *Benefits and Best Practices for Integrating Laboratory Instruments with LIMS*, Astrix Technology Blog, accessed: 25.02.2020, available at: www.astrixinc.com

Dudenhoefer, L., 2018, *Process Mapping Checklist: How to Make an Accurate Process Map*, Triaster Blog, accessed: 04.12.2019, available at: www.blog.triaster.co.uk

EMA, 2015, *ICH guideline Q9 on quality risk management*, European Medicines Agency

EMA, 2013, *EudraLex - Volume 4 Good manufacturing practice (GMP) guidelines - European Commission*, European Medicines Agency

FDA, 2018, *Data Integrity and Compliance With Drug CGMP Questions and Answers Guidance for Industry*, U.S. Food and Drugs Administration

Fleischmann, A., Schmidt, W., Stary, C., Obermeier, S., Börger, E., 2012, *Subject-Oriented Business Process Management*, Springer, Berlin, Heidelberg

Gaillard, C., 2015, *How to Facilitate Successful Process Mapping Sessions*, [WWW Document], BAtimes. accessed: 03.12.2019, available at: www.batimes.com

Harris, E., 2019, *3 Common Business Process Mapping Mistakes (and How to Avoid Them)*, Triaster Blog, accessed: 22.01.2020, available at: www.blog.triaster.co.uk

Harris, E., 2016, *Six Major Benefits of Process Mapping*, Triaster Blog, accessed: 24.02.2020, available at: www.blog.triaster.co.uk

Helmers, S., 2018, *Four Components of an Effective Process Map*, accessed: 02.12.2019, available at: www.netmind.net

ISO, 2018, *Riskienhallinta. Ohjeet*., ISO 31000:2018, International Organization for Standardization

Kalman, H.K., 2002, *Process Mapping: Tools, Techniques, & Critical Success Factors*, Performance Improvement Quarterly, Vol. 15, No. 4, pp. 57–73

Kiran, D.R., 2017, Chapter 26 - Failure Modes and Effects Analysis, in: Kiran, D.R. (Ed.), *Total Quality Management*, Butterworth-Heinemann, pp. 373–389

Liu, H.-C., 2016, *FMEA Using Uncertainty Theories and MCDM Methods*, Springer, Singapore

Longden, H., 2017, *Data Integrity*, LC-GC North America

Madison, D., 2005, *Process Mapping, Process Improvement, and Process Management: A Practical Guide for Enhancing Work and Information Flow*, Paton Professional

McDowall, R.D., 2019a, *Data Integrity and Data Governance: Practical Implementation in Regulated Laboratories*, The Royal Society of Chemistry, U.K., Croydon

McDowall, R.D., 2019b, *Data Integrity Focus, Part 2: Using Data Process Mapping to Identify Integrity Gaps*. LC-GC North America, Vol. 37, No. 2, pp. 118–123

McDowall, R.D., 2017, *Validation of Chromatography Data Systems: Ensuring Data Integrity, Meeting Business and Regulatory Requirements*, The Royal Society of Chemistry, U.K., Croydon

Mettler Toledo, 2016, *Laboratory Data Integrity with LabX*, [WWW Document], accessed: 27.02.2020, available at: www.mt.com

Mettler Toledo, 2007, *LabX Software, Complete Overview, Total Control, Smart Laboratory Instrument Software*

MHRA, 2018, *"GxP" Data Integrity Guidance and Definitions*, Medicines & Healthcare products Regulatory Agency

Nichols, J.H., 2011, *Laboratory Quality Control Based on Risk Management*, Annals of Saudi Medicine, Vol. 31, No. 3, pp. 223-228

Oy Medfiles Ltd., 2020, *Computerised systems and responsibilities, ICT-16-01*, Standard Operating Procedure

Oy Medfiles Ltd., 2019a, *Quality manual - Version 3*

Oy Medfiles Ltd., 2019b, *Näytteiden kulku analytiikka- ja tuotekehitysyksikössä, KL-114-07*, Standard Operating Procedure

Oy Medfiles Ltd., 2019c, *Laboratory notes, KL-101-08*, Standard Operating Procedure

Oy Medfiles Ltd., 2019d, *LIMS Vellamo, Asiantuntijoiden ohje*, WPI

Oy Medfiles Ltd., 2018, *Investigation of out of specification (OOS) result, KL-605-04*, Standard Operating Procedure

PIC/S, 2018, *Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments*, Pharmaceutical Inspection Co-operation Scheme

PIF, 2019, *Code of Ethics*, Pharma Industry Finland

Price, D., 2017, *Overview of GxP Data Integrity Guidelines*, GMP Platform, accessed: 29.01.2020, available at: www.gmp-platform.com

Rath, F., 2008, *Tools for Developing a Quality Management Program: Proactive Tools (Process Mapping, Value Stream Mapping, Fault Tree Analysis, and Failure Mode and Effects Analysis)*, International Journal of Radiation Oncology Biology Physics, Vol. 71, No. 1, pp. S187–S190

Rattan, A.K., 2018, *Data Integrity: History, Issues, and Remediation of Issues*, PDA Journal of Pharmaceutical Science, Vol. 72, No. 2, pp. 105

Rosemann, M., 2006, *Potential pitfalls of process modelling: part A*, Business Process Management Journal, Vol. 12, No. 2, pp. 249–254

Sandahl, K., 2015, *How to create a good swimlane diagram* [WWW Document], accessed: 21.01.2020, available at: www.gluu.biz

Sharp, A., McDermott, P., 2009, *Workflow modelling: Tools for process improvement and application development*, 2nd Ed., Artech House, INC.: Norwood

Soares, M., 2019, *Best Practices for Integrating LIMS with Empower*, LIMS Implementation, accessed: 29.0.2020, available at: www.astrixinc.com

Software Point, 2020, *LabVantage ELN – Dokumentaation ja prosessin hallinnan tuki*, Webinar

Software Point, 2016a, *Mikä ihmeen LIMS?* [WWW Document], accessed: 6.11.2019, available at: www.softwarepoint.com

Software Point, 2016b, *LabVantage ELN & LES*, [WWW Document], accessed: 21.02.2020, available at: www.softwarepoint.com

Software Point, 2014, *Labvantage Connect*, [WWW Document], accessed: 25.02.2020, available at: www.softwarepoint.com

Stankiewicz, A., Moulijn, J.A., 2005, *Re-engineering the chemical processing plant: Process intensification*, Taylor & Francis Group

Triaster, 2018, *The Ultimate Guide to Business Process Mapping* [WWW Document], accessed: 12.11.2019, available at: www.triaster.co.uk

Visual Paradigm Online, 2020, *What is Cross-Functional Flowchart?* [WWW Document], accessed: 21.01.2020, available at: www.online.visual-paradigm.com

Waters, 2019, *Empower 3 Chromatography Data Software: Waters*, [WWW Document], accessed: 20.11.2019, available at: www.waters.com

WHO, 2016, *Guidance on Good Data and Record Management Practices*, World Health Organization

APPENDICES

Appendix I	FMEA, risk assessment model, 1 page
Appendix II	Data integrity risks, 2 pages

APPENDIX I – FMEA, risk assessment model

FMEA – Risk assessment model

dd.mm.yyyy

Risk assessment subject:

Risk question:

General assumptions:

Process step	Potential failure mode	Potential effect(s) of failure	Potential cause(s) of failure	S	P	D	RPN	Recommended actions	Schedule	Comments

S = Severity
P = Probability
D = Detectability
RPN = Risk Probability Number (= S · P · D)

APPENDIX II – Data integrity risks

The found risks are presented in table below. The x means, that there is recognized data integrity risk. Colour indicates the fatality of the worst risk found in each process step followingly: red = high risk, yellow = medium risk, green = low risk. Process steps written in grey and with grey background are mapped further and processed separately. See the sub maps below and about the OOS process in Chapter 11.1.

Process step	Attributable	Legible	Contemporaneous	Original	Accurate	Complete	Consistent	Enduring	Available
Has the request arrived	x	x							
Sample registering	x				x	x			
Worksheet approval						x	x	x	
Preparing the analysis	x	x	x	x	x	x	x		
Preparing the sample	x	x	x		x	x	x		
Sample analysis	x	x	x		x	x	x		
Review of injection sequence	x						x		
Reporting (Sign off 1)	x				x				
Work review	x	x		x					
OOS process	x	x	x						
Are the results acceptable (Sign off 2)						x			
Result calculation	x					x			
Result approval	x		x						
Reporting to LIMS						x			

Appendix II - Data integrity risks

Process step	Attributable	Legible	Contemporaneous	Original	Accurate	Complete	Consistent	Enduring	Available
Additional reporting						x			
Archiving and disposal							x	x	x
SUBPROCESS, Worksheet approval									
Generating worksheets				x	x	x			
Worksheet approval (electronic and paper)						x		x	
Are the sheets approved?								x	
Printing 'TSL logi'						x		x	
SUBPROCESS, Sample preparing and analysis									
Preparing solutions and standards	x	x	x		x	x			
External test	x	x		x					
Sample weighing	x	x		x					
Preparing the sample and sample analysis	x	x	x		x	x	x		