A spark of foresight – Configurations of Toyota Production System practice bundles in healthcare operations through a subunit lens.

A fuzzy set Qualitative Comparative Analysis

by

Martin Roeger

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy.

Warwick Business School

September 2020

1 CONTENTS

2 TABLE OF FIGURES

3 TABLE OF TABLES

Acknowledgements

Thank you to my readers, you are the reason and reward for this thesis. Thanks also to my supervisors, Dr. Nicola Burgess, Prof. Dr. Mark Johnson and Prof. Dr. Leroy White, who guided me on this path. Thank you! In addition, I would like to acknowledge all the support I received from my family and friends.

This project was financially supported by Heart of England NHS Foundation Trust and the Collaboration for Leadership in Applied Health Research.

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

Martin Roeger, 22/09/2020

'Ohno's approach was never to implement a particular tool, but to build appropriate social and technical capabilities to fit the circumstances' (Lander and Liker, 2007, p. 3684).

4 INTRODUCTION

4.1 RESEARCH BACKGROUND

Every day, healthcare professionals are faced with carrying out an unpredictable workload while running the gamut of human characteristics. Their training to become flexible and adaptive champions of their field enhances their abilities to improve the quality of life for all of us. Unfortunately, 12 hour waiting times, crowded A&E floors, and burned out doctors call for drastic measures or enforced improvements if healthcare systems are not to collapse. The Lean Production System has established an enviable track record for achieving improvements in manufacturing and service organisations. It provides a multitude of tools, techniques, and philosophies that may be adopted to improve the vast and complex realm of modern healthcare.

Successfully implementing the Lean Production System in healthcare organisations can effect improvements that go well beyond expectations; however, unsuccessful implementations can amount to little more than resource-burning changes. While numerous, complex, and industry specific barriers hamper every attempt to implement Lean systems in the healthcare sector (Leite et al., 2019), an in-depth understanding of intertwined Lean methods enables breath-taking rises across performance indicators.

For these reasons, Lean practitioners are constantly seeking new insights into links between Lean methods that can cope with the complexity of the healthcare system and impact on performance indicators. Faced with the vastness and knottiness of the issues, many practitioners have settled for shoehorning generic 'best practice' Lean methods (Zirar et al., 2020) onto their own specific problems. They thereby risk burning through scarce because they have overlooked the essential characteristics of their setting. Overcoming these difficulties can only be achieved by improving our understanding about when simplicity rather than specialisation, or vice versa, will lead to superior performance results.

4.2 Research objectives

A lack of research surrounds the definition of Lean practices. This will be addressed in the literature review. Instead this study will use the following definition of management practices, developed based on Galbraith (1977)'s work: Management practices are used by managers to improve the efficiency and/or effectiveness of operations through working methods, approaches, and innovations.

This thesis will follow this line of argument and study the relations between various Lean practices to identify how combining these in different ways impacts on performance indicators. By comparing subunits instead of entire organisations, this research will also provide deeper insights into how Lean implementations impact on complex healthcare organisations. These aims will be addressed through the following research questions:

- How do combinations of Lean practices lead to superior performance improvement in healthcare subunits?
- How is the `choice' of the practices dependent on the organisational setting?
- Which practices are required and/or sufficient for a higher impact on bundle performance?

Given the complexity of healthcare operations and organisations, the novel method fuzzy set Qualitative Comparative Analysis will be applied.

While some research advances have been made, there remains significant uncertainty about the effects of combinations of Lean practices on performance indicators such as quality, efficiency, and accessibility (Alkhaldi & Abdallah, 2019). This thesis will address that gap by identifying configurations of Lean practices that lead to superior performance improvement, while also investigating practices that are necessary or sufficient for this outcome. This will allow future studies to compare different combinations of Lean practices across other settings and industries. The identification of the differences and similarities across various Lean practices will support further developments and optimisations of the Lean approach – ultimately, strengthening health services. In particular, knowing which combinations of Lean practice are required/sufficient to effect specific performance outcomes will enable Lean managers to focus resources where they are most needed.

4.3 Research structure

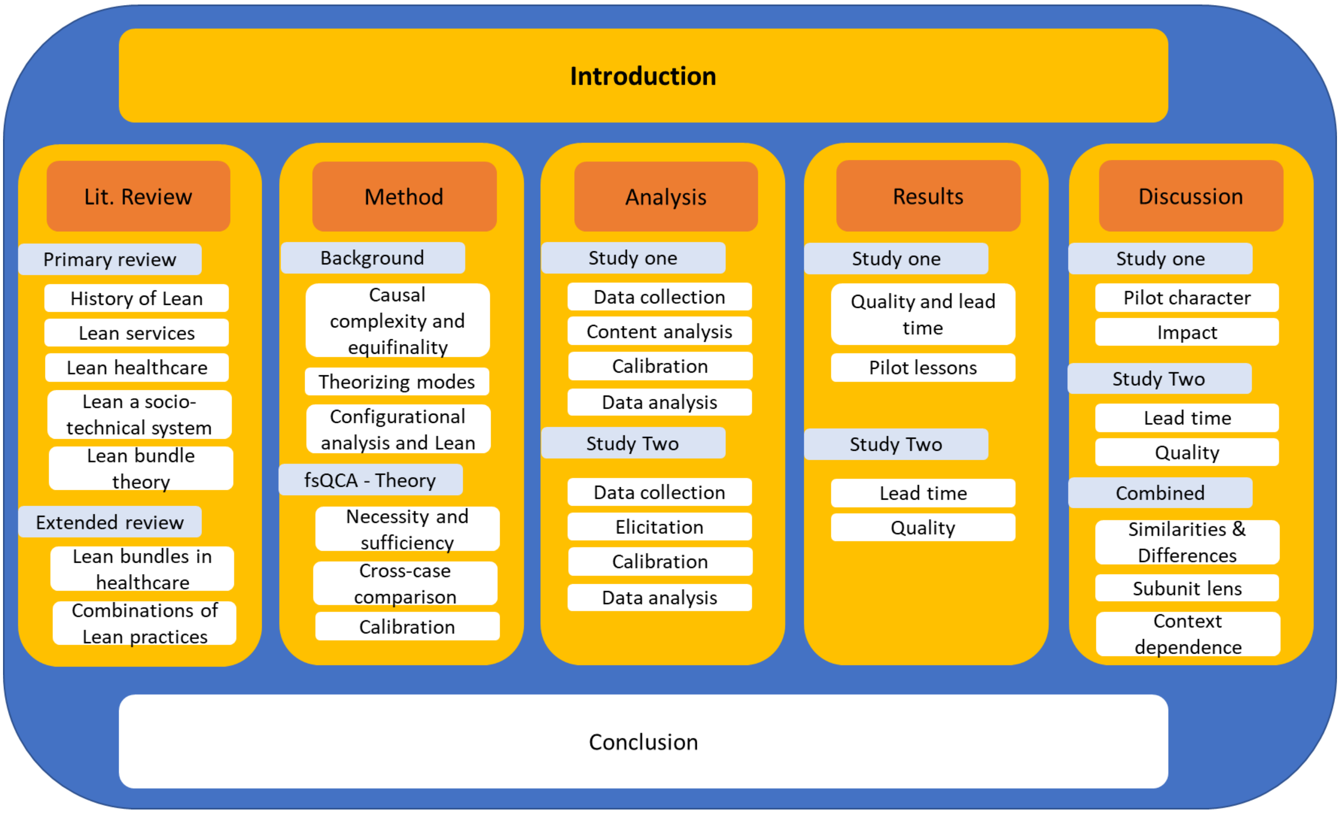

Figure 1: Structure overview

Figure 1 presents a summary of how the thesis is structured. The following paragraphs will discuss in greater detail how each individual chapter contributes to the overall framework of the thesis.

4.3.1 Chapter 2 – Literature Review

In the second chapter, two literature reviews will be presented: a primary review and an extended review. The primary review will focus on the background of the Lean Production System (referred to henceforth as 'Lean'). A general overview of the history of Lean will be given. Furthermore, the chapter will introduce and discuss Lean in services and Lean in manufacturing. The differences between the two approaches will highlight the need for a Lean healthcare approach. The subsequent examination of the Lean approach from a more abstract perspective will include a discussion of the socio-technical nature of Lean as well as Lean bundle theory. This thesis will argue that Lean bundle theory can be successfully adopted by healthcare organisations.

The second literature review will present an extended search for combinations of Lean practices in healthcare. Thus, this thesis will provide an in-depth understanding of which combinations of practices frequently occur and how many practices they contain.

4.3.2 Chapter 3 – Methodological Argument

The aim of the third chapter is to offer a more detailed perspective on the contextual setting in which the empirical work will take place. Furthermore, it will introduce the still novel method of fuzzy set Qualitative Comparative Analysis (fsCQA).

Regarding the contextual setting, this chapter will start by introducing and defining notions such as equifinality, asymmetry, causally complex, and causal conditions. This thesis will argue that any Lean implementation in healthcare faces complex, multidisciplinary, heterogenic organisations, and processes. In such an intricate environment, traditional linear models struggle to explain phenomena. Therefore, the chapter will discuss the differences and suitability of a number of theorising modes. Based on this discussion, the thesis will argue in support of a neo-configurational lens, and will explain in detail the advantages this offers to configurations and typologies. The adoption of a configurational lens provides the foundation for the application of fsQCA.

The second part of the chapter will introduce fsQCA and elaborate on recent developments in its development and ongoing discussions about its implementation. Additionally, the chapter will explain the major elements of fsQCA, namely calibration, cross-case comparison, necessity, and sufficiency. At the end of the chapter, a discussion about the uniqueness of the setting and the contextual suitability of the method will make the way for the analysis chapter.

4.3.3 Chapter 4 – Analysis

The fourth chapter will present the empirical procedures of the analysis. As the complexity and novelty of the method creates a degree of uncertainty about the actual data collection and analysis, the thesis is split into two studies, with the first acting as a pilot. Because of this split, the chapter will discuss both studies individually. Hence, the analysis of the first study will present the data collection process, the content analysis, and the calibration, before going on to discuss the data analysis proper. The same elements will be discussed for the second study but with the addition of an elicitation element. Specifically, while the data analysis in the first study is presented via the traditional fsQCA approach of truth table analysis, the analysis in the second study will highlight the advantages of superset/subset analysis. Additionally, both studies will argue in support of the necessity and sufficiency analysis.

4.3.4 Chapter 5 – Results

Following the analysis, the fifth chapter will present the results. Like the analysis chapter, the results chapter will distinguish between the first and second studies to better identify the individual results. The results of the first study will detail the impact on performance and the identified configurations of Lean practices. The results of the second study will present configurations that lead to strong lead time and/or quality improvement. Both studies' results will explain the sufficient and necessary practices for each performance indicator.

4.3.5 Chapter 6 – Discussion

The sixth chapter will consist of a discussion about the individual studies, as well as their combined contribution. The insights from the pilot study will be discussed, as well as the lessons learnt from it. The chapter will then focus on the performance impact of the identified Lean configurations and the differences between the two studies. Furthermore, the thesis will argue that fsQCA is a suitable tool for the investigation of Lean combinations in healthcare in that it provides in-depth understandings about the relationships between practices as well as the importance of individual practices for the overall performance impact. The chapter will close with a detailed discussion of the applied subunit lens, its managerial implications and limitations. Finally, the thesis will argue that its findings suggest that future research adopts a subunit lens rather than an organisational lens when exploring which Lean practices are needed for targeted performance impacts.

4.3.6 Chapter 7 – Conclusion

The seventh and final chapter will summarise and synthesise the effects of the entire thesis. It will explain how the identified Lean configurations enable researchers to compare results across industries and also allow managers to focus their resources on the practices that are most important to their setting. In addition, the thesis will argue that Lean bundles in healthcare should be regarded as context dependent rather than as predefined constructs. This insight will prevent practitioners from falling into the trap of adopting 'best practice' Lean bundles; to gain maximum benefit from a Lean implementation, they should give thorough consideration to the operational and organisational context. Finally, the chapter will close with an acknowledgement of the study's overall limitations, some of which offer suggestions for future research.

5 LITERATURE REVIEW

5.1 CONCEPTUAL FRAMEWORK

5.1.1 Introduction – Literature Review

The literature review will address past and current research in the area of Lean Production Systems. A general review will be followed by an in-depth presentation of the shortfalls in the current state of the scientific discussion.

First, the review will give an extended presentation of the history of Lean. This will take the form of a discussion about the development of various general process improvement methods in pre-war Western factories, followed by a Japan focused perspective on the foundation of Toyoda (the precursor to Toyota) and the Toyota Production System (TPS). The history section will then address the TPS research at the Massachusetts Institute of Technology. Overall, this section will function as a summary of Lean, which will enable the reader to better understand later remarks or references in the thesis.

The literature review will then switch to a critique of the past discussion of Lean. It will present the ongoing struggle to develop a suitable definition for Lean and highlight the main perspectives of the Lean Production System.

The next section will present and discuss (Shah and Ward, 2003)'s definition of Lean by examining it through a socio-technical lens and transferring this argument onto the causal derivation of Lean bundles. At the end of this section, an initial discussion about the limitations of this theory and lens will be presented.

After identifying the limitations in previous studies, the review will discuss (Hadid et al., 2016)'s creation of the Lean service bundles. Therefore, a presentation of different streams of Lean service theory will be presented.

The next section will address the transfer of Lean services into healthcare. The presentation of a wider theoretical framework for Lean service bundles in healthcare will end the general review before the focus switches towards a systematic literature review (SLR). The SLR will analyse studies that have explored individual Lean practices or sets of Lean practices implemented in healthcare to depict the occurrence of a pattern between certain practice combinations. Figure 2 summarises the progression of the literature review chapter.

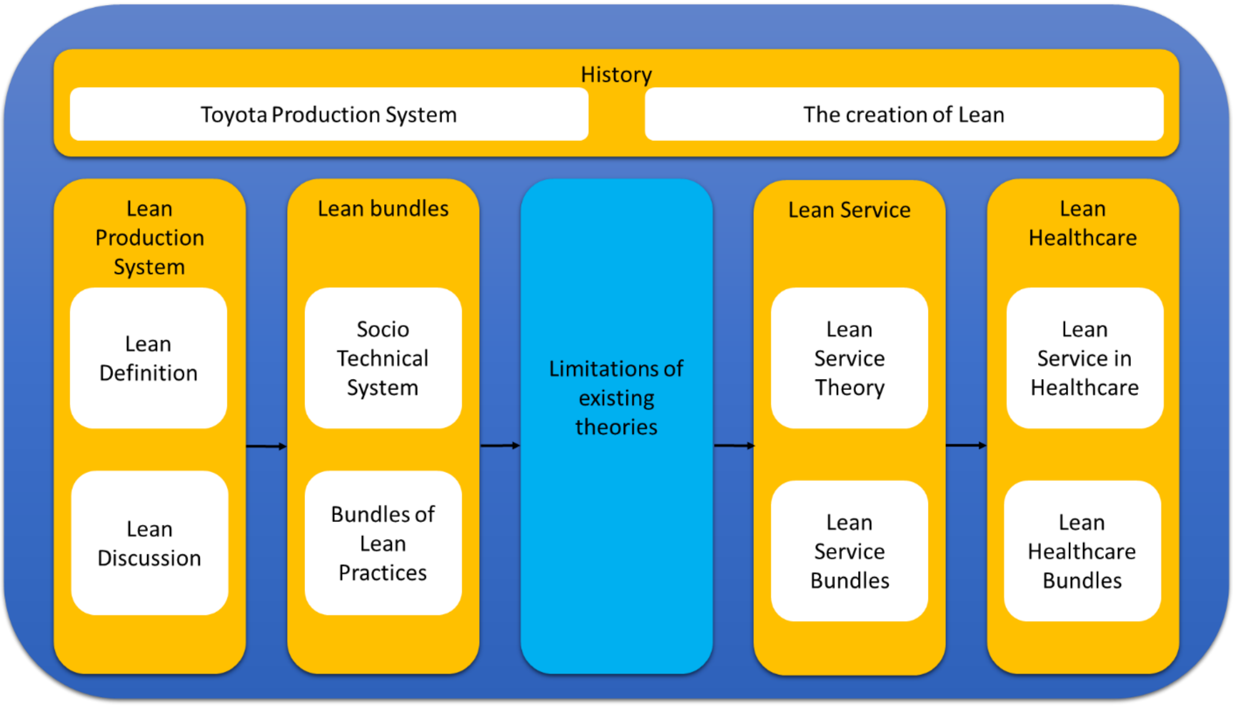

Figure 2: Overview of Literature Review

5.1.2 Literature Review

5.1.2.1 The history of Lean

5.1.2.1.1 Toyota Production System

According to Holweg (2007), Lean is based on several process improvement approaches but it is mostly derived from the Toyota Production System (TPS). However, even the development of TPS has several influences. Following the industrial revolution in the West, several factories/organisations tried to actively develop basic process improvement measures to achieve an advantage in competitive markets (Reinertsen, 2005). While it cannot be categorically claimed that process improvement techniques emerged in the late 19th or early 20th century, the push for better performance certainly began at the turn of the 20th century. Around 1900, the rise of a number of improvement approaches began, which included scientific management (e.g. Taylor, 1919). Components of this approach had however been previously utilised. For example, piecework (Brech and Brech, 1975) and bonus or premium systems and rate fixing (Elbourne et al., 1979) were introduced to the British cotton industry before 1900. Two other well-known examples are the adoption of Taylorism in the European pharmaceutical industries (Homburg, 1978) and the advancements of performance measurement methods after the First World War (Bourne et al., 2003). In 1910 in the USA, Henry Ford introduced a mass production system that utilised a moving production line. The entire production line was dedicated to the Model T, reducing system variations and streamlining the manufacturing processes by introducing rope powered conveyor belts (Wilson and McKinlay, 2010). Womack (1990) saw in Ford's mass production system and synchronized flow the antecedents of the Lean Production System.

All of these early approaches to production are likely to have had an impact on the development and creation of the Toyota Production System. According to (Fujimoto, 1999), Sakichi Toyoda introduced a combination of Taylorism, bonuses, and performance measurements in his spinning and weaving company after the First World War. In 1924, Toyoda developed a loom that stopped when a strand of yarn broke (Ohno, 1988b), thereby decoupling the operator from the machine; this allowed a single worker to supervise several machines instead of just one. This became known as autonomation (Monden, 2011). The combination of autonomation, stop the line, and root-cause analysis formed the early elements of Jidoka ('intelligent automation') (Ohno, 1988a). The patent rights to the automated loom were later sold to the Patt Brothers, the financial gains of which provided funding for the advancements of the Toyota Motor Corporation (Holweg, 2007).

The development of Jidoka was the foundation of the Toyota Production System. During the wars, general resource limitations and American sanctions against Japan provided constraints on the country's economic growth and stimulated the need for more process improvements. In 1937, using the earnings from his previous endeavours, Kiichiro Toyoda founded the Toyota Motor Corporation. Furthermore, he pursued the development of TPS. For example, he integrated Jidoka with his philosophy of Just in Time (JIT), '…the production of the necessary products in the necessary quantities at the necessary time in every process of a factory and also among companies' (Monden, 2011, p. 35). After the war, he visited the Ford factories in the US to gain an understanding of their process efficiencies (Vaghefi et al., 2000). In 1951 the ‘Creative Idea Suggestion System’ was introduced. This actively motivated staff to fully engage with process improvement. In addition, the Toyota company created a knowledge sharing system between itself and its suppliers (Dyer and Singh, 1998). In 1943, Taiichi Ohno joined the company as the supervisor of part of the production line. Ten years later, he became known for identifying the seven categories of waste (Monden, 1983) and further developed TPS by adding a first form of Kaizen and Kanban (Ohno, 1988b). In 1960, the industrial engineer Shigeo Shingo came to Toyota as a consultant. He finalized Kaizen, added Single Minute Exchange of Die (SMED), and creative thinking to TPS. His work strengthened the philosophical aspects of TPS (Monden, 2011).

According to (Ohno, 1988b), TPS diverted the individual engineers’ focus from machinery to process flow. The introduction of TPS allowed Toyota to obtain cost reductions, higher quality standards, and greater variety, as well as generating the firm’s ability to adapt to customer preferences based on a speed-focused throughput time.

The strength of the achievements of TPS became more visible in the next centuries. According to (Swanson, 1999), the Second World War had destroyed entire production lines in Europe and Asia. This created a lack of competition that generated easy rewards for the post-war companies that actively pursued growth. Although there were resource constraints in the early post-war years that forced efficiency improvements, these companies generally had little need to push for performance optimisation. Manufacturing in the USA had not been pulled down by war in the same way as it had in Europe and Asia, but the potential for market growth rendered the need for efficiency improvements less pressing. The Japanese economy however had suffered greatly during the war and resource constraints remained in place in the post-war period. As a result, Japanese companies pushed at this time for major efficiency gains through waste reduction. Even after the millennium, this focus on efficiency remains one of Toyota’s competitive advantages (Fane et al., 2003).

5.1.2.1.2 Lean production

By the end of the 20th century, several automotive production systems had been identified. For example, Shimada and MacDuffie (1986) conducted research on Honda, Nissan, NUMMI, and Mazda, investigating the organisational status of the companies after the implementation of production systems. They used a specific benchmarking framework, classifying the organisations as either fragile, robust, or buffered. Even in this plethora of systems, the Toyota Production System remained omnipresent. It was during this time that Womack's MIT International Motor Vehicle Program (IMVP) research group changed its terminology to Lean, better representing the actual definition, and this was the basis of today's Lean (Holweg, 2007).

Another member of the IMVP programme, (Krafcik, 1988), investigated the relationship between corporate parentage, culture, technology level, and plant performance. He used the term 'lean production' and showed that corporate culture and parentage have a stronger relation with performance than technology level has. (Womack et al., 1990) used the terminology 'lean production' synonymously to provide a contrast between the TPS and the 'mass production system' of the West. They argued that the superiority of Lean compared to mass production was inherent and would eventually become visible in all value creating activities, from the distribution of manufactured products to health care delivery.

During the 20 years that followed (Krafcik, 1988)'s introduction of the term 'lean production' (which at that time was concerned with waste reduction) a number of scholars called for clarity in the definition of the Lean Production System. (Bartezzaghi, 1999) pointed out that the current definitions were unhelpfully vague and could lead to confusion. (Karlsson and Åhlström, 1996)'s extensive study indirectly provided an example of this. They identified 18 major determinants that affect Lean. Several of these tools, sub-tools, techniques, and principles pointed in different directions. In addition, so many different determinants had created a practical challenge for implementation. For example, (Oliver et al., 1996) showed that identifying all the determinants required to complete the Lean questionnaire would take around 5 days of managerial time. The complexity of the Lean

Production System had been identified early on but scholars suggested various pathways for how the theory should be adapted. Lewis (2000) suggested viewing each Lean implementation as individually dependent upon business performance, strategic objectives, and managerial experience. In contrast, (Muffatto, 1999) and (Hines et al., 2004) indicated that the continuous improvements characteristic of Lean make it a permanently evolving system, which is hard to constrain and define. Therefore, any definition could only be a valid illustration of Lean for a certain amount of time. A complete definition of Lean would require a wide and considered abstraction combining theories that generalize all the complex elements, their relationships, and the different organisational impacts, all of which depend on context. Shah and Ward (2007) heeded this call by suggesting the following Lean definition: 'Lean production is an integrated socio-technical system whose main objective is to eliminate waste by concurrently reducing or minimising supplier, customer, and internal variability' (p. 791)[1]. In addition, (Shah and Ward, 2003) characterised the main elements of Lean as bundles, which were Just-in-time (JIT), Total Quality Management (TQM), Total Productive Maintenance (TPM), and Human Resource Management (HRM).

Although this model of Lean, especially the concept of the bundles, attracted initial scepticism from a number of scholars (e.g. Mi Dahlgaard-Park and Pettersen, 2009), it remains one of the most cited Lean definitions.

(Shah, 2002, Shah and Ward, 2007) combined several theoretical constructs in her ground-breaking 2002 publication. First, she considered Lean as a wide-ranging system that consisted of several practices. This enabled her to enter the discussion of Lean without having to support a specific side of the debate. She then used socio-technical systems theory (Shah and Ward, 2007) in combination with configurational theory (Shah, 2002) to explain the relationships between the various Lean practices. Third, she defined the constructs of Lean practice sets as bundles (Shah and Ward, 2003), which will be discussed in more detail in the

[1] A socio-technical system consists of traditional tool/ process focused elements and supportive elements concerned with people and their relationships Trist, E. (1981). The evolution of socio-technical systems. *Occasional paper*, *2*, 1981. . A more detailed explanation follows on the next pages.

additional literature review. Individual elements of Shah's work are significant to this study. Elements such as her Lean discussion, socio-technical system theory, configurational theory (in the method section), and Lean bundle theory will all be discussed in detail.

5.1.2.1.3 The Lean debate

The following section will discuss different perspectives on the Lean production system. In the literature, two main perspectives are discussed equally – 'toolbox lean' (Bicheno, 2004, Nicholas and Soni, 2005) and 'lean thinking' (Liker, 2003, Womack and Jones, 2010). The toolbox lean perspective, sometimes also referred to as 'real Lean' (Nørby et al., 2008, p. 213), considers Lean as a set of tools and practices that share the main purpose of eliminating waste (Shingo and Dillon, 1989, Bicheno, 2004, Monden, 2011). In contrast, lean thinking considers Lean to be a philosophy that is applied to reduce waste and increase customer value (Hines et al., 2004, Holden, 2011). Therefore, it can be summarised that 'toolbox lean' is a practical and project focused approach, while 'lean thinking' considers the philosophical perspective to be an integrated part of Lean. This is not to say that the two perspectives are contradictory. Indeed, (Stentoft Arlbjørn and Vagn Freytag, 2013) pointed out that both perspectives might exist on different levels. Lean thinking would take a wider perspective that is akin to principles and philosophy, while toolbox Lean would be the base of a hierarchical triangle. Nevertheless, the two perspectives have fundamentally different understandings and targets. Targets can be internally orientated (Schonberger, 1982, Ohno, 1988b, Shingo and Dillon, 1989, Bowen and Youngdahl, 1998, Feld, 2000, Liker, 2003, Monden, 2011) or externally orientated (Schonberger, 1982, Womack et al., 1990, McDermott and Stock, 1999, Bicheno, 2004, Womack and Jones, 2010, Dennis, 2016). According to (Mi Dahlgaard-Park and Pettersen, 2009), the differences in the targets are very small. Though externally orientated customer satisfaction improvement varies considerably from internally orientated inventory or cost reduction measures, neither of these two orientations can be considered to be an appropriate description of Lean. (Hines et al., 2004) state that Lean consists of an operational dimension and a strategic dimension. Therefore, the more recent studies do not consider the two perspectives to be conflicting. Nevertheless, studies tend to have a soft but inherent bias towards either 'toolbox lean' (Shah and Ward, 2003, Shah and Ward, 2007) or 'lean thinking' (Joosten et al., 2009, Mazzocato et al., 2010, Papadopoulos et al., 2011). However, as (Hardcopf and Shah, 2014) point out, this tendency might be related to the organisational context, and they use the organisational culture as their example. Given the different organisational units in which Lean practices might be applied, as well as the theoretical relationship between the Lean Production System and its practices, these findings can be abstracted to a practical level.

Therefore, it can be concluded that a single Lean Production System might consist of different groups of practices, each having shared characteristics, targets, and/or intra-relations. One group might be concerned with tools, techniques, and processes, while another one group is focused on interpersonal connections and their impact. This differentiation is the subject of the next section.

5.1.2.2 Socio Technical Systems

5.1.2.2.1 The Development of STS

At several junctures Lean bundle theory is interlinked with the socio-technical nature of operational practices. The theory of socio-technical systems (STS) was originally introduced between the late 40s and early 50s (Trist and Bamforth, 1951). STS considers organisations as constructs that contain groups of social and technical components. Several scholars have investigated the STS in different industries, such as mining (e.g. Emery, 1974, Trist, 1981), textiles (e.g. Rice, 1953, Rice, 2013), automotive (e.g. Melman, 1958), telecommunications (van Beinum, 1963), public transportation (e.g. Carmien et al., 2005), and healthcare (e.g. Qureshi, 2007).

According to (Trist, 1981), the characteristics of STS allow the traditional technical perspectives to be combined with social components. In this way, new theories, substituting older ones, formed a new organisational system perspective that combined the technical and social perspectives. A technical system includes elements like equipment, techniques, tools, and processes; in Lean, this would be Kanban (Shah and Ward, 2003). On the other hand, a social system consists of people and their relationships (Trist, 1981); this might take the form of Lean staff training (Karlsson and Åhlström, 1996). Both sides exist separately but they are clearly interdependent in that improving one requires improvement in the other, and they share the objective of generating an optimal performance result (Trist and Murray, 1990). This concept may also be described as complementary relations. In more practical terms, reinforcing the components of the technical system (e.g., by increasing investment in its practices) at the expense of the social components, or vice versa, will result in a suboptimal performance (Fox, 1995).

5.1.2.2.2 The relationship of STS and Lean

The social and technical sides interact, resulting in improved performance (Dabhilkar and Åhlström, 2013). Since 1990, STS has received increased attention in the Operations Management research area and it is used to improve knowledge about the performance effects of improvement systems, like Lean. For example, (Lander and Liker, 2007) investigated TPS through the lens of STS. Their findings indicated that combining technical components and social components improves organisational performance by supporting the understanding and learning of organisational processes. Specifically, the combination of Jidoka, JIT, and the maximal utilisation of workers' capabilities creates synergies between the social and technical sides of STS. (Manz and Stewart, 1997) used STS to examine TQM's impact on performance. The results indicated that improving either side of STS would be enough to add value. However, when both components are implemented together, their interaction has an impact on performance that is greater than the sum of its parts (Hadid and Afshin Mansouri, 2014).

While many scholars broadly support the theory of STS, different viewpoints have been expressed, and these will be elaborated on now. According to (Adler and Docherty, 1998), most of the 1970s research recommended the social side (Trepo, 1979, Rollier, 1979). However, in the 1980s, there was a shift of interest as younger scholars focused their attention on the technical side (Mumford and MacDonald, 1989). The two sides were melded in the notion of the 'balanced business dimension' that gained traction in the 1990s (Stace, 1996, Mumford, 2000). While all the perspectives crop up in the recent literature, it seems that many scholars (e.g. Bjerknes and Bratteteig, 1995) support a balanced business dimension. The paradigm shift is still ongoing. While the balanced lens seems to have slowed down the discussion, several questions remain unanswered.

In Operations Management research, the theory of STS has been expanded to several other subfields, for instance, operations control (Wilson et al., 2007), quality management (Asif et al., 2009) and Lean. In the case of Lean, (Shah and Ward, 2003) described a possible complementary relationship between Lean practices as well as the Lean sets of practices. (Shah and Ward, 2007) extended this assumption and defined Lean as a system, specially an STS. However, STS was not a novel concept for Lean theorists. In the 90s, (Huber and Brown, 1991) employed an STS lens to investigate an implementation of the cellular lean manufacturing system. In that study, the authors explained how the theoretical framework could be complemented by adding six HR practices, resulting in an enhancement of Lean's effectiveness. In particular, the social practices of job analysis, training, selection, planning, employee relations, and reward structures increased the performance impact of cellular Lean practices through improvements to the quality of work-life and employee attitude. While these findings explained how two sub-sets of Lean work together, they still did not explain the overall entity. (Cua et al., 2001) went some way towards this by linking the level of implementation of the social and technical practices in the sub-sets TQM, JIT, and TPM to manufacturing performance.

These studies share a belief that Lean, and STS are directly related to each other. Indeed, few studies oppose this notion. For example, (Niepcel and Molleman, 1998) highlighted that while the characteristics of Lean and STS show strong similarities, such as multifunctionality, the attention given to team work, and the need for feedback, there are differences in how these are addressed, such as how they control and coordinate work. Abstracting (Shah and Ward, 2007)'s definition of Lean allows the Lean system to be viewed as a socio-technical system without going so far as to require it. While the socio-technical lens provides a powerful explanation of the relationships in a lean system, it might be that the simple assumption of socio-technical classifications for practices might lead the observer to overlook the relevance of other relations. Therefore, while this study accepted and utilised Shah and Ward (2007)'s socio-technical interpretation of Lean, it nevertheless approached the conceptual and empirical work with an open-minded and neutral lens.

5.1.2.2.3 Relationship of STS with performance measures

The following section will discuss how seeing Lean as an STS enabled performance gains. For example, Malmbrandt and Åhlström (2013) identified a variety of Lean service practices and used them to create a tool for measuring the progress of the implementation. Through a systematic literature review, Hadid and Mansouri (2014) extended this list of Lean practices. Thus, they enabled the distinction of practices as either technical or social. Moreover, they delivered evidence to support the assumption that considering Lean as an STS has positive performance impacts. In practice, this means that an implementation that links social and technical practices is more likely to perform well than an implementation that does not. For instance, (Piercy and Rich, 2009)'s study showed that technical practices eliminate non-value-adding activities, free up time, and reduce costs, cycle, and lead time; customer value was thereby increased. In addition, by emphasising the importance of performing tasks correctly the first time, technical practices have enabled organisations to improve their service quality and reduce their failure rates and costs (Swank, 2003, Piercy and Rich, 2009).

Furthermore, (Hadid et al., 2016) provided evidence that technical practices have an independent connection with operational performance, and zero association with financial performance. In contrast, the combination of social and technical practices had a positive impact on both performance categories. The study therefore tested for interactions between the two sides. The authors' findings revealed that distinctive bundles or sets of Lean service practices improve diverse performance dimensions. These findings support the assumption that service organisations should focus on a wider set of Lean service practices to surmount the barriers to great performance (Cua et al., 2001).

While the relevance of both sides of STS to Lean has been established, detail is still lacking about the individual practices and their relationship to performance. Studies (e.g. Hadid and Mansouri, 2014) have either viewed a set of Lean practices as an STS or investigated the influence of individual socio and/or technical individual practices (e.g. Swank, 2003). Nevertheless the amount of studies linking both remains limited, hence forcing this work to explain how combinations of Lean practices influence performance, in greater detail.

5.1.2.3 Lean practice bundles

In the operational environment, practices do not normally occur individually. In most cases, individual practices operate in sets of ‘interrelated and internally consistent practices’ (Shah and Ward, 2003), also called practice bundles.

It is likely that the term Lean practice emerged from the usage of ‘practice’ in the management research field (Galbraith, 1977) and transitioned step by step into other business domains. For example, (MacDuffie, 1995) investigated whether bundles are not so much a group of random individual practices that have a link to performance but rather contain consistent and interrelated HR practices. His results underlined the findings of (Arthur, 1992), which showed that a successful HR bundle must be implemented in addition to a bundle of core business functional practices. (MacDuffie, 1995198, p.198) defined bundles as ‘interrelated and internally consistent HR practices, rather than individual practices’. While (Cua et al., 2001680, p.680) were still using the term ‘set of practices’, (Shah and Ward, 2003) adopted the HR definition, and validated it empirically within a larger manufacturing practice set/group. Moreover, they stated that interrelated practices form a bundle, they presupposed four different manufacturing practice bundles (TQM, TPM, JIT, and HRM), and demonstrated their relationship to operational performance.[2]

Previous to that study, only a few operations management studies had tried to empirically investigate the relationships between JIT and TQM (Flynn et al., 1995),

[2] An overview about the bundled practices of Shah and Ward (2003) can be found in the Appendix A1.

TPM, TQM, and JIT (Cua et al., 2001, McKone et al., 2001) and a complementary relationship between TQM and JIT, where HRM can operate as enhancer of this relationship (Furlan et al., 2011). (Shah and Ward, 2003) were the first scholars who tested all four bundles and their combined impacts on performance. Later, (Womack and Jones, 2010) supported these findings and described the combination of Lean practices as a system, from which benefits amass. While the four practice bundles of (Shah and Ward, 2003) were validated empirically and are commonly used in recent manufacturing studies (e.g. Fullerton et al., 2014, Bortolotti et al., 2015b), their generalisability or transferability to the service industry has only recently gained momentum.

In an older article, (Flynn et al., 1995) highlighted that the combination of certain practices enables the creation of competitive advantages. They thereby provided an argument in support of bundled practices empowering operational benefits.

(Hadid et al., 2016) based their work on this theorem. Their findings indicated the existence of interactions between Lean bundles, especially the effect exerted by one bundle on performance while also increasing the effect on the performance of another bundle. Therefore, they did not simply use the socio-technical system theory to distinguish between social and technical practices but also linked these to specific outcomes. Seemingly, (Bortolotti et al., 2015a) intended to advance the understanding of STS in Lean through the introduction of the notions of hard and soft practices. They interpreted the (Shah and Ward, 2007) definition of Lean as a system of interrelated hard and soft practices. While hard practices are technical and analytical lean tools, soft practices are related to people, their relations, and behaviour. The authors' findings indicated that managers often focus on the implementation of hard practices and seem to overlook the soft practices. In addition, successful manufacturing plants were more likely to have implemented a wide range of soft practices to supplement the hard practices. Unsuccessful plants generally did not do this. While soft lean practices are important for the achievement of high performance (e.g. Samson and Terziovski, 1999, Matsui, 2007) and the maintenance of the longitudinal performance (Hines et al., 2004), hard practices operate as a predictor of high performance (Taylor and Wright, 2006).

However, not all hard practices are associated with strong performance outcomes. For example, the performance impacts of benchmarking cellular work teams as well as advanced manufacturing technologies were considered to be context dependent (Dow et al., 1999). Thereby, the relationship between both practice types is essential for the influence on performance.

Although the findings of both (Hadid et al., 2016) and (Bortolotti et al., 2015a) support the understanding of Lean as an STS, they also indicate that there still is a lot of unknown about the boundaries (e.g. What is part of Lean?) and relationships between different Lean practices (e.g. socio vs technical / soft vs hard). This review identifies this as the first research gap. The following section will discuss another.

5.1.2.4 Limitations of Shah's Lean bundle theory

Mi Dahlgaard-Park and Pettersen (2009), responding to (Shah and Ward, 2003), showed that the characteristics of HRM and supply chain management are not definable parts of Lean. Their findings (scores) for the bundled techniques SCM and HRM were strong enough to conclude a link with Lean but that it was not an essential one. However, the authors did not further elaborate on these findings in relation to Shah and Ward (2003)'s bundles. They therefore question the applicability of these bundles in practice without going so far as to refute Shah and Ward (2003)'s bundle theory.

Several studies have further investigated Lean bundles. For example, (Dal Pont et al., 2008) showed that an HR practice bundle had only an indirect impact on operational performance, which was in fact achieved through the other bundles, in that the HR practices create the foundation for the implementation of other practices. They thereby enhance the performance impact of JIT and TQM. However, these limits of Shah and Ward (2003)'s bundles become only relevant when seen through a homogenous lens that views bundles as somehow static. Considering products and processes as homogenous allows the majority of variables to be kept constant for the whole population (Porter, 2011). This lens gained traction because of the relative homogeneity of the processes in the manufacturing industry.

According to (Bonavia and Marin, 2006), the manufacturing industry can be considered as a homogeneous sector because operational practices are quickly adopted by market participants. In addition, production variables show little variation.

Most early studies investigating Lean bundles assumed this generalisation as a necessary step to investigating bundle relations (e.g. Flynn et al., 1995, Cua et al., 2001, McKone et al., 2001, Shah and Ward, 2003, Furlan et al., 2011). However, the predefined bundles created by such generalisation of practices meant that the impact on performance of individual practices was either disregarded or averaged out in the construct of the Lean bundle. For example, (Flynn et al., 1995) investigated the relations between a set of TQM and JIT practices. The authors' findings indicate that `perceived quality market outcomes' and `percentage of items that passed final inspection without rework' had the significantly highest impact on performance of the TQM practices. In addition, `fast throughput' and `automation for cost reduction' had the highest impact on performance of the JIT practices. However, these individual differences within the bundles were overlooked in the focus on the relations between the bundles. While the work around Lean bundles has contributed to the understanding of the Lean Production System (Dal Pont et al., 2008), it overshadows questions about which relations between the individual practices impact on the overall performance of the bundle. The given four bundles of Shah and Ward (2003) draw the attention towards the constructs of these bundles but do not necessarily provide answers about which contents of these bundles drive the performance impact of the specific bundle. This study will try to address this question:

Which practices are required and/or sufficient for a higher impact on performance of the bundle?

The generalisation from Lean practices to Lean bundles requires the bundles to be somewhat homogenous in order to juxtapose them. Considering Lean bundles as homogenous items may simplify their comparability, but it comes with drawbacks, the first of which is the implication that all bundled practices share similar characteristics. This also neglects the extent of the differences in the alignment of these characteristics. This limitation becomes especially visible when comparing the impact on performance of the bundled practices. For instance, (Cua et al., 2001) investigated the impact of TQM, TPM, and JIT practices on the following performance dimensions: cost efficiency, conformance quality, on-time delivery, volume flexibility, and weighted performance. The findings of the authors supported the general characteristics of TQM, TPM, and JIT but also showed that only the general practice `committed leadership' and the TPM practice `emphasis on technology' showed significant relations to all performance dimensions. Both (Cua et al., 2001) and (Shah and Ward, 2003) generalised the characteristics of the practices for the bundles. The vastness and diversity of the field of Lean practices requires the investigation of bundles and their relationships to be somewhat simplified and the assumption of homogeneity enabled Shah and Ward (2003) to unite many perspectives on Lean manufacturing. Thus, intentionally or otherwise, bundles were viewed as homogenous constructs. However, the price paid for this is a lack of clarity about the individual practices. The impact of this limitation can be assumed to be less in a homogeneous environment, like the manufacturing industry, and more influential in heterogeneous environments, like the service industry. For example, in the manufacturing industry it can be assumed that production lines are generally similar. The inputs are known, the transformation process is known, and the output can be assumed. Therefore, these processes can always be improved by a specific group of practices that will maintain their similarity of relationships in another production line. In the context of this setting, it is logical to assume a quasi-constant relationship between bundles. However, when we turn to healthcare, the pathway of an A&E patient is unknown. The input has limited predictability until the first examination, and even than it remains possible that the assessment was incomplete or even wrong. Once the appropriate assessment has been accomplished, the transformation process is mostly clear, but its output still is not. Lean implementation in service industries is more likely to need tweaking to fit the context than its implementation in similar service process lines. Therefore, it can be argued that Lean practices might need to be more flexibly grouped in the service sector than in the manufacturing industry. The boundaries of existing bundles (e.g. JIT, TQM) might need to be considered permeable rather than rigid. Thereby these bundles might include uncommon practices but adhere the service setting.

(Prajogo, 2005) pointed out that the service sector is generally more diversified, differentiated, and specialized than the manufacturing industry. Several scholars have acknowledged that service industries differ from manufacturing industries, especially in terms of labour intensity, work specialisation, perishability, heterogeneity, simultaneity, intangibility, and the presence of patients/customers during the work process (Bowen and Youngdahl, 1998, Sampson and Froehle, 2006, Büyüközkan et al., 2011).

(Shah and Ward, 2003) believed that their Lean bundle theory could be applied to similar industries but that it might require further research in other more diverse sectors.

5.1.2.5 Lean Service

According to (Hadid et al., 2016), growing interest has emerged in recent studies surrounding the implementations of Lean manufacturing practices in the service sector, introduced as Lean service (Bowen and Youngdahl, 1998). (Papadopoulos et al., 2011) investigated how organisational players in healthcare shift their network allegiance over time and thereby impact on Lean manufacturing implementation outcomes. In addition, several others have discussed the adaptability of Lean practices to a wide range of service industries. For the majority of the studies, this was explored through conceptualisation or case-type studies (Suárez-Barraza et al., 2012). While these studies have extended the literature, several scholars have stressed that there is a lack of large-scale studies (Malmbrandt and Åhlström, 2013, Hadid and Afshin Mansouri, 2014), studies with a wider focus on Lean from a systems perspective (Radnor et al., 2012), and studies that assume interactions between practices and their environment (Bortolotti et al., 2015a, Hadid et al., 2016). There have been some efforts made to remedy these deficiencies. For example, (Alsmadi et al., 2012) investigated the implementation of Lean manufacturing practices into a service industry setting and the dependencies between them, and (Poksinska et al., 2013) examined how leadership behaviours change as a Lean implementation adjusts the organisational environment. While these studies share similarities in structure, their empirical limitations and focus on the impact of isolated Lean practices leads them to ignore potential interactions (e.g. Talib et al., 2013) among each other (Hadid et al., 2016). In conclusion, while a foundation of shared understanding around Lean service has been created, there remain gaps in the scientific discussion.

Several studies indicated the limitations of the conceptualisation of Lean in the healthcare environment (D'Andreamatteo et al., 2015). The primary barriers are the industrial assumptions behind the Lean Production System that hamper transferability (Hadid and Afshin Mansouri, 2014). To avoid this while considering Lean as a set of Lean practices, a finer-grained conceptualisation of Lean bundles in healthcare is required. However, the process of conceptualisation needs to address the discussion surrounding Lean bundle theory, the individual conceptualisation of Lean services, as well as the different research perspectives on the Lean production system.

5.1.2.6 Lean Service in Healthcare

5.1.2.6.1 The transition of Lean manufacturing to healthcare

According to (Metters and Vargas, 2000), the main difference between service and manufacturing organisations is the contact with customers/patients, which leads to highly connected and diverse task-specialist units. Therefore it is questionable if (Shah and Ward, 2003)'s findings can be replicated in the service industry. While several scholars have validated individual elements, namely JIT (e.g. de Souza and Pidd, 2011), TQM (e.g. Mannon, 2014), and HRM (e.g. Lorden et al., 2015), only limited evidence exists to validate the use of all four bundles together in the service industry. This raises the following question:

How is the 'choice' of the practice bundle content dependent on the organisational setting?

This study will argue that the predefinition of bundles is limited; their dependence on the organisational setting seems to be stronger in the service industry than was posited by the mainly manufacturing driven Lean theories. This study does not reject (Shah and Ward, 2003)'s bundle theory, but raises questions about how far it can be applied to the service industry.

5.1.2.6.2 Lean Service in Healthcare

While a considerable amount of research has been conducted that conceptualises Lean bundles in manufacturing industries (e.g. MacDuffie, 1995, Cua et al., 2001, Shah and Ward, 2003), few studies have considered healthcare.

The scarcity of the research means that a practical conceptualisation approach, such as framing or summarising, would be difficult to apply (Remler and Van Ryzin, 2010). However, wider triangulation might support the creation of a concept by assembling surrounding theories (Singleton Jr et al., 1993, Neuman, 2002). In this case, these would be a combination of Lean manufacturing, Lean service, individual practice impacts in healthcare, configurational theory, socio-technical theory, and causal asymmetry.

The conceptualisation of Lean bundles in healthcare is founded in the definition of (MacDuffie, 1995, p.198), who referred to bundles as 'interrelated and internally consistent […] practices, rather than individual practices'. (Shah and Ward, 2003) utilised this definition and conceptualised the four main Lean bundles as TQM, TPM, JIT, and HRM. However, their study was conducted in manufacturing rather than in a service industry and thus this study looks to the six socio-technical Lean service bundles of (Hadid et al., 2016) (namely, process, customer value, error prevention, financial, human, and motivation) for the second foundation of the conceptualisation. While the work of Hadid and Mansouri (2014) included healthcare studies, Hadid et al. (2016)'s did not. Their work found that social and technical Lean service bundles improve each other, and it contributed to the literature by providing a generalised perspective on service bundles. However, this generalisation led to reduced attention being given to healthcare characteristics.

As a result, the third foundation of this study draws on the different characteristics of service industries in general and healthcare services in particular. According to (Morton and Cornwell, 2009), healthcare services differ from other services in their unpredictability, the professional nature of the production class, and their service orientation. Healthcare processes can be considered unpredictable because of their irreducible variability and task ambiguity. The variability of customers in online shops, consultancy companies, or airlines can be standardized to an extent, but the variability of patients cannot. Furthermore, when a patient enters a care process, nobody knows which tasks will be required until he/she has exited the process. In addition, healthcare services differ from other services by the professional nature of the service delivery. If there is a knowledge gap about the health of the human being, it is difficult to establish whether this is because the doctor or nurse is not 'good' at their job. Finally, there is the service orientation of healthcare. While several healthcare services involve direct customer contact, patients are also concerned with indirect service aspects (Goodrich and Cornwell, 2008), such as being able to have visitors, their trust in the doctor's ability, and the perceptions of staff awareness about patient needs. Therefore, it can be concluded that services differ in their organisational, process, and environmental aspects. While this does not appear to have direct impact on the theoretical framework of the Lean bundle concept, it does influence the alignment and content of the bundles. The major differences are already visible in the practical bundles of manufacturing industry (e.g. MacDuffie, 1995, Cua et al., 2001, Shah and Ward, 2003) and service industry (e.g. Hadid and Afshin Mansouri, 2014, Hadid et al., 2016), as discussed in sections 5.1.2.3 and 5.1.2.5 above.

In conclusion, the author will use an operational definition that considers Lean as a set of interrelated bundles, consisting of interrelated practices aligned to suit the organisational setting, processes, and environment of the targeted implementation area.

5.1.2.7 The transition of Lean service to the healthcare industry

(Bowen and Youngdahl, 1998) were the first to introduce the concept of Lean service to the literature. Specifically, they discussed the impacts of the implementation of Lean manufacturing practices in the service industry. While the literature has generated paradoxical results concerning the appropriateness of transferring Lean practices from the manufacturing industry to the services sector, the ongoing search for improvement methods has forced service organisations to try to adapt the methods to their needs (Hadid and Afshin Mansouri, 2014). Furthermore, this increased attention has resulted in advancements in the Lean service concept. However, the adaption of Lean practices has changed. At the outset, Lean practices were very much seen through the Lean manufacturing lens but over time this changed, and the differences between the industries became visible. (Hadid et al., 2016) used socio-technical system theory to explain the relationships in Lean service bundles, providing an alternative to (Shah and Ward, 2003)'s perspective of Lean. The studies were conducted in different industrial settings and so they do not contradict each other. However, (Hadid et al., 2016) presented new findings around Lean service bundles. Their work showed that the technical elements of Lean service did not have an impact on financial performance, while the social elements did have an impact on the operational and technical performance indicators. In addition, they identified 4 technical bundles and 2 social bundles. Their sample was focused on the service industry rather than healthcare and indeed, very few studies have investigated bundles in just healthcare (e.g. Costa and Godinho Filho, 2016). Clearly, the identification of Lean healthcare practice bundles will require their empirical validation. Traditionally, this is done by showing a link with a performance outcome of interest.

According to (Anuar et al., 2018), it remains unclear whether Lean practices lead to operational improvements. Some studies investigated relationships between Lean practices and performance indicators in healthcare. For example, (Aoun, 2015) showed that Kaizen, JIT, Andon, Jidoka, Poka Yoke, Kanban, Hoshin Kari, and quick change occur together, while standardisation by 5S remains independent. Both Aoun (2015) and Anuar et al. (2018) showed a positive impact on innovation management and, more specifically, the innovation rate. Healthcare organisations, like hospitals, tend to divide into a number of highly specialised subunits, even more so than other service organisations. (Chiarini, 2013) explored the relationship between value stream mapping, activity worksheets, and spaghetti charts on waste reduction (time and cost). The findings showed that more standardized assessments of A&E patients allowed for a more accurate evaluation of patients' health situations, facilitating decision-making about discharge or further hospitalisation.

These studies emphasise the strength and performance influences of technical Lean practices and stop short at discussing socio-practices. (Radnor, 2009) pointed out the specifically high occurrence in healthcare of Rapid Improvement Events (RIEs) in short term improvements, and Kaizen events in long term improvements. RIEs operate as drivers for Lean implementation, which they achieve through high staff involvement.

The following paragraph will summarize the main lessons from the general literature review and highlight how they lead to the motivation for this study.

5.1.3 Research Motivation

The Lean Production System has achieved success in a wide variety of nations, industries, sectors, environments and processes. It has shown that it can deliver continuously improvement in difficult contexts. Healthcare is through its complex and heterogenic operations one of these contexts. Shah and Ward (2003)'s understanding of Lean as a bundle of internationally consistent Lean practices enables a finer grain consideration such diverse operations.

However, closer insights into the relationships between Lean practices in healthcare remain wanting. While first reviews have touched Lean practices in healthcare, their lens was still strongly biased towards other research objectives or industrial settings. For example, Hadid et al. (2016) focused on the overall Lean service perspective, while Costa and Godinho (2016) gave little attention to healthcare's heterogeneity. Furthermore, only a few studies have operationalized Lean bundles in healthcare. In conclusion, while studies indicate that Lean service bundles may have similar relations within Lean healthcare, there is inadequate research to explain how Lean practices operate together in healthcare. In addition, virtually no research has considered the heterogeneity of healthcare facilities and their subunit diversity. This study intends to change that through an in-depth literature review identifying how combinations of Lean practices occur, followed by an empirical analysis of the performance impact of Lean practices in healthcare organisations.

5.2 Additional Literature Review

As previously discussed, the dearth of studies that identify combinations of Lean practices in healthcare while taking into account the industry's characteristics has resulted in the author's decision to conduct an additional, more focused, literature review. This second review aims to identify a variety of Lean practice combinations in healthcare. Given that the identification of practices and their relationships are well known objectives of scientific discussion, a structured review approach was chosen to strengthen the validity and coherence of the review.

The detailed review process applied within this paper was based upon the approach outlined by Tranfield (2003), albeit modified for the specific context of this research. An illustration of the review process applied within this research can be seen in Figure 1, below.

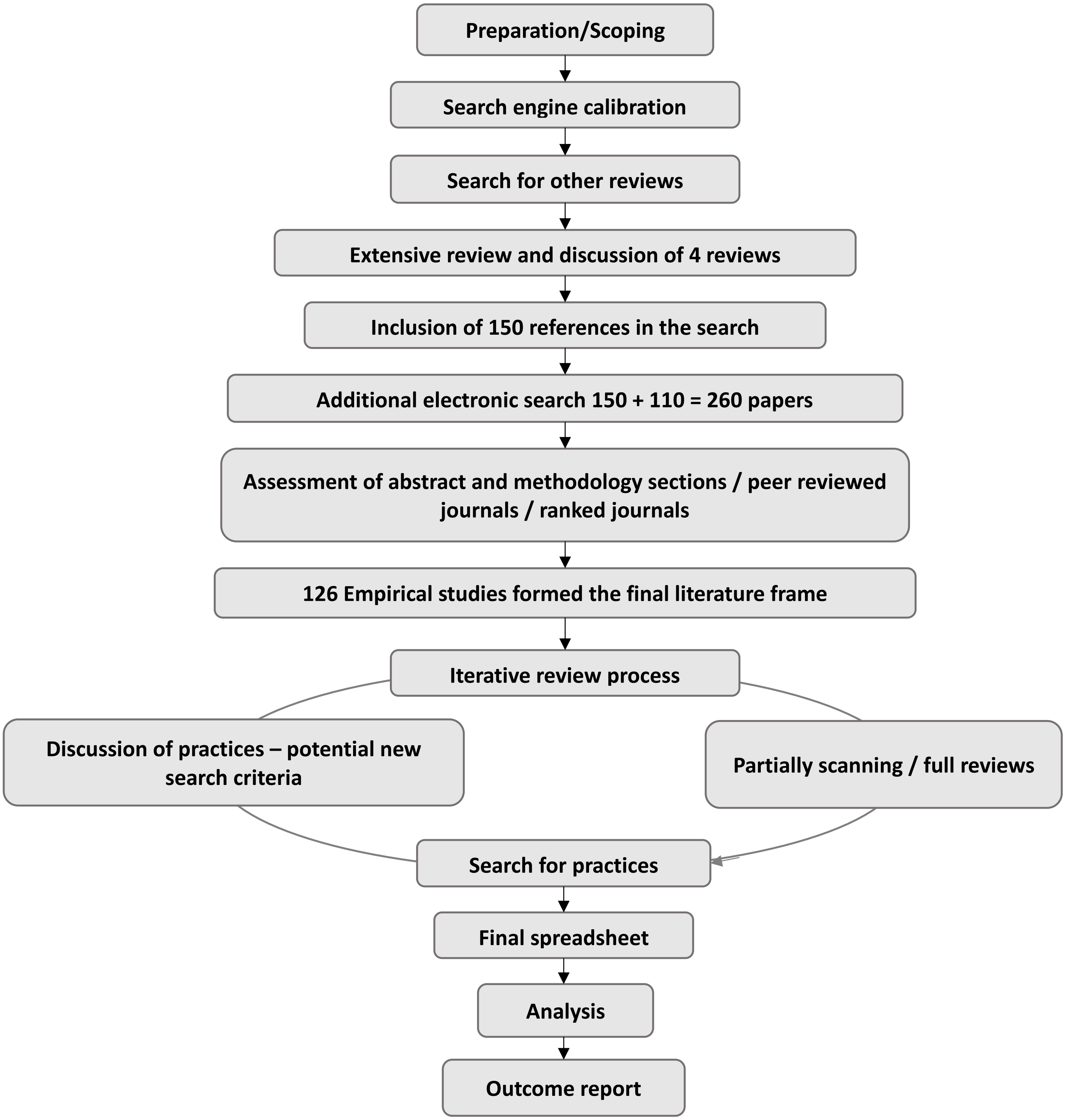

Figure 3: Review process

During the first stage of this process, the multidisciplinary background of the literature considered in this context presented a significant challenge. In particular, the term ‘practice’ is differently defined and understood according to the field and area of study. This led to careful considerations of inclusions and cross-comparison with other references ensuring consistency. Furthermore, limiting the scope of the review to management related journal databases was considered insufficient, given that the majority of the literature originates from medical or public health journals. Therefore, an initial review was conducted that including searching within different academic disciplines, fields, and using a broad range of additional search terms, as outlined below. As a result, all databases that provided suitable results were included and explored. Since no meta-database that linked all relevant subject databases was found during the review process, a Google scholar search algorithm was adopted and applied using a small self-written script with embedded database APIs. The script accessed MEDLINE, Web of Science, ScienceDirect, Embase, Proquest, Springer, ABI Inform and Emerald, Ei Compendex (Elsevier), and IET Inspec (Elsevier).

The search terminologies that were included can be divided into two categories. First, Lean or Toyota Production System terminologies were used in order to limit the scope of quality improvement approaches. Second, a variety of specifications in different combinations were added. These included practice, tool, routine, bundle, implementation, service, healthcare, hospital, ambulance, A&E, and ED. While papers written in English made up the core of the review, some French, Spanish, and German papers were also reviewed. These were small in number since different terminologies with unknown definitions were commonly used in these foreign language papers. To ensure coherence and consistency in the set of papers reviewed, only papers with comparable definitions were included. Hereby, it was revised if papers described processes from a management and/or operational lens. Papers taking a purely medical lens were excluded as they exceeded the scope of the study. However, studies taking a medical and managerial lens were included after a rigorous second read.

During the initial review, four similar literature reviews were identified: (Bucci et al., 2016, Costa and Godinho Filho, 2016, Hadid and Afshin Mansouri, 2014, Punnakitikashem, 2013). However, given their focus and scope as outlined below, none of these reviews provided the answer to our research question. For example, Bucci et al. (2016) investigated Lean implementations solely in emergency departments and rigorously excluded other studies. Hadid and Afshin Mansouri (2014) focused on service implementations in general, with health studies forming only part of the review. Conversely, Costa and Godinho Filho (2016) conducted a strong healthcare-wide literature review; however, their focus on mainly theoretical papers and their summary of a few smaller practices under one practice term created their own limitations. As a result, a decision was made to use the 151 references cited and their definitions of Lean practices in these reviews as a foundation for this review, with the addition of another 113 papers identified during the search process.

In the next stage, the abstracts and methodology sections of the identified papers were assessed on scientific relation to Lean in healthcare. Additionally, the papers' contents were scanned, and a second assessment of if the journals are peer reviewed and ranked was conducted. The assessment considered if the extent of provided information was sufficient to identify Lean practices, if the found keywords had any managerial relationship towards Lean practices and how was the balance between empirical and theoretical work. Most of the papers from the previous reviews, as well as the additional papers identified in the search process, were purely theoretical, thus reducing the number of empirical studies to 126. Given that the focus of this paper is on Lean practices and that the research question aims to highlight differences between theoretically constructed bundles and the applied configuration of Lean, a decision was made to focus on these empirical papers as the core of this review.

The next stage of the review process was based on an iterative full paper review. This approach was chosen for several reasons. First, some of the practices (such as process flow analysis and value stream mapping) are hard to distinguish in only the abstract or a cursory reading, especially if the Lean coach or provider uses vocabulary different to that used within the research papers. This decision was made after an observation during our empirical data collection that one hospital ward had a (replacement) Lean coach who used the terms 'Kaizen blitz' and 'Kaizen event' (for rapid process improvement workshops or RPIW).

Secondly, the initial reviews indicated a relationship between the higher ranked papers and their awareness of Lean healthcare investigations within other research fields. In most cases, papers in lower ranked journals had fewer citations from other fields as well as a more inconsistent use of terminologies. For example, many older medical publications (i.e., prior to 2006/7) focused mainly on technical Lean implementations, with few indications that social practices existed. However, by this time the management literature, especially the Operations Management literature, had already shifted its focus towards a more diverse Lean lens (Shah and Ward, 2003), even if the transition to a socio-technical lens for Lean, especially Lean health, was a few years away (e.g. Papadopoulos and Merali, 2008, Joosten et al., 2009).

Therefore, an iterative review process was used to identify and assess practices. If a practice was validated and shared links to past and/or theoretical papers, it was included in the practice set and all articles were checked again for this practice during the next iteration. In practice, this meant that the first papers were compared to original Lean work (e.g. the machine that changed the world) and if comparability was found the practices and papers were included in the list. With the rising amount of included papers, papers were compared internally. In total, the review required seven iterations to identify 58 practices. This allowed an in-depth analysis of combinations of practices by the inclusion of 'rarer' practices. The results of the review were stored and illustrated in Excel (see Appendix A6).

During the final stage of the review, descriptive analysis was applied to the raw data. Additionally, another small script was used to automatically extract the journal information (year of publication, subject area, H-index) from Scientific Journal Ranking database (SJR). The focus at this stage was on gaining an understanding about the research lens and its causal connection to terminologies. Following the initial investigation, the research question was addressed through a small-scale truth table analysis that presented the distribution of truth values and highlighted many possible configurations. This analysis uses Boolean logic to test the validity, dependency, and consistency of structured data (Wittgenstein, 2013, Anellis, 2012). Given the large number of practices and studies, truth table analysis would generate many solutions. Rather than searching for 'all' solutions, this study aims to provide an overview of the most frequently occurring combinations of practices and their relations. Therefore, the study focused first on the most frequently occurring practices and their set partners. It then searched for the largest possible combination covered by a minimum of two studies. The highest number of combination participants was five practices. In the next stage the combinations with 4 practices were investigated to identify the solutions with the largest study coverage. The process proceeded by reducing the number of practices (minimum 2) and increasing the required number of studies (minimum 5). This procedure was necessary as otherwise the number of solutions would have drastically exceeded the limits of feasibility.

The results were highlighted in three spreadsheets. The first gave a general overview about Lean practices applied in healthcare settings and thereby sharpened previous findings (Costa and Godinho Filho, 2016). Another table provided information about combinations of Lean practices; this uncovered insights into the setup of bundles as well as their rarity. The third spreadsheet classified each practice into (Radnor et al., 2012)'s types and their socio-technical background.

5.2.1 Review results

While the review was part of a continuous process, by summer 2019 114 main authors had been identified, who had published 126 papers. Many such papers had been published in health journals (more specifically public health, e.g., Social Science and Medicine). However, the highest number of citations of individual articles were achieved either by specialist medical journals (e.g., American Journal of Obstetrics and Gynecology) or management journals (e.g., Journal of Operations Management). Conversely, the citation originations have changed slightly over the last decades. Before 2012, 82% of citations originated from health journals and only 16% from management journals. Today, management journals—mainly Operations Management journals, including the Journal of Operations Management, the

International Journal of Operations & Production Management, the Supply Chain Management Journal, and Public Money & Management Journal—contribute 19% of traceable citations. This indicates a soft rise in the interest of Operations Management scholars in Lean implementations in healthcare. However, it should be mentioned that most authors investigated Lean implementations with only positive performance impacts.

5.2.2 Practices

Overall, the papers examined within this review provided information about a wide range of Lean practices, such that 58 practices were identified. Following the distinction of (Hadid et al., 2016), 36 practices were classified as technical and 22 practices as supportive (i.e., social). Appendix A4 provides a general overview of the number of named practices. Additionally, 18 practices appeared 10 or more times. Overall, the seven most cited practices were as follows:

Value Stream Mapping	71
Standardisation	55
Process/patient flow analysis	47
5S	36
RPIW / RIE	34
Visualisation	32
Integrated process redesign	32

These seven practices were chosen as they are included in a minimum of 25% of all studies. The most cited practice—Value Stream Mapping—appears in 56% of all papers.

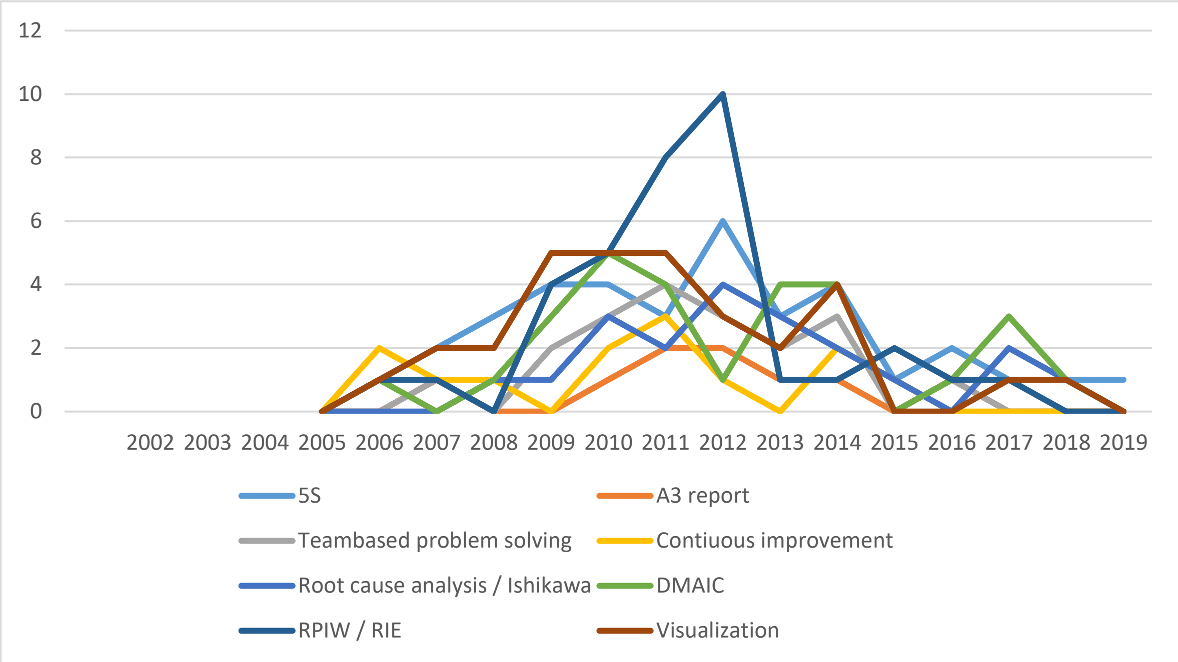

Figure 4: Occurrence of practices over time 1

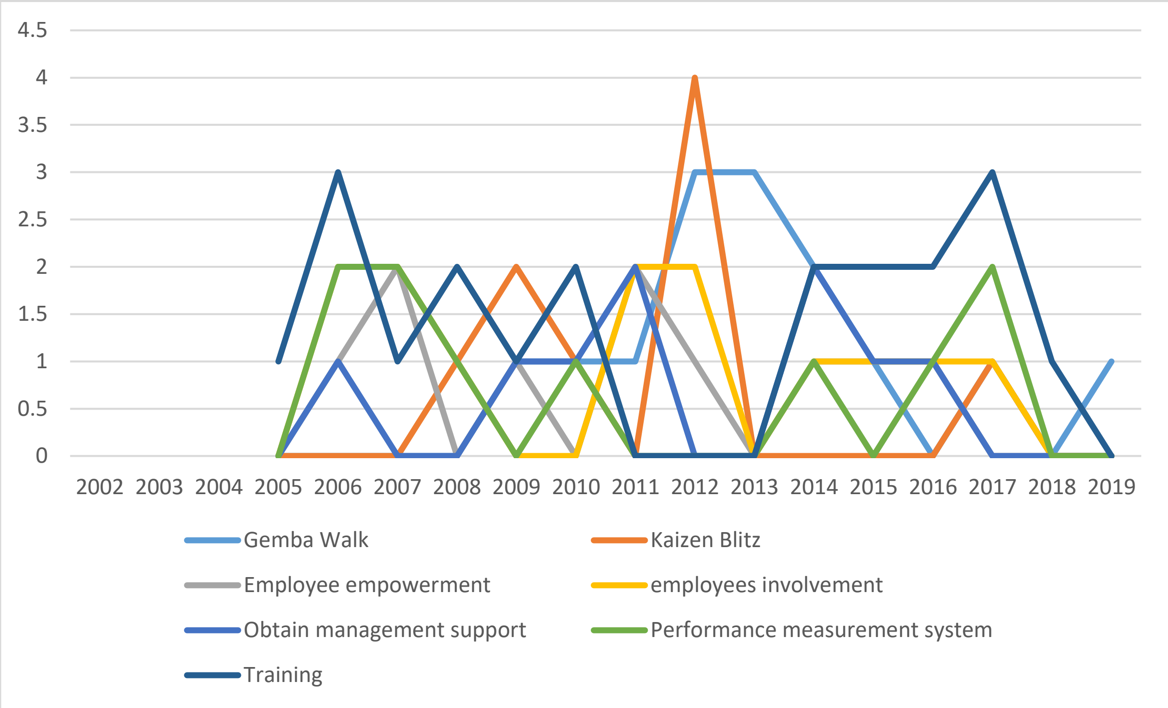

Figure 5 Occurrence of practices over time 2

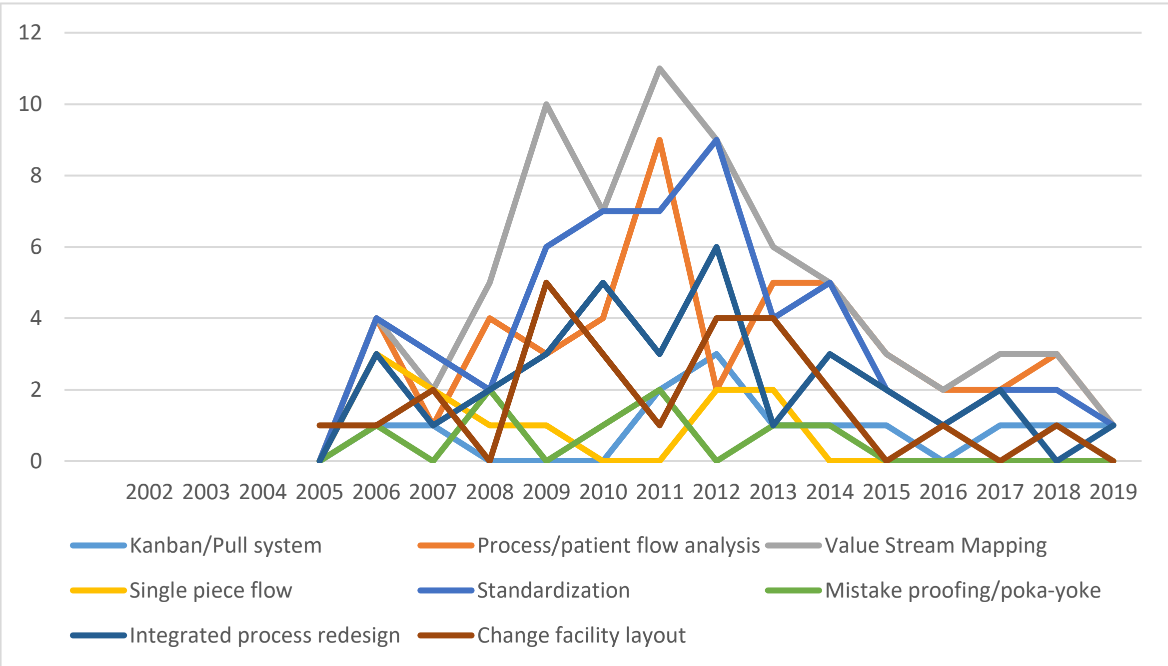

Figure 6: Occurrence of practices over time 3

Furthermore, the systematic review allowed for an investigation of the development of Lean practices over time.

In general, after 2009, social practices were increasingly discussed within the literature, with the discussion moving from a technical lens to a mixed perspective. For example, rapid improvement events (RIE) or rapid process improvement workshops (RPIW) received most of their citations between 2009 and 2012. Similarly, studies investigating team-based problem-solving peaked between 2010 and 2012. However, some supportive practices (e.g., training) rose and fell over the years but overall, their development remained constant over the entire period. Additionally, we did not find a single study that examined just one individual practice, with most studies investigating medium-sized combinations of practices.

5.2.3 Configurations

In most cases (32x) within the literature, similar combinations of two practices were investigated. Figure 7 shows the number of studies in which x-practice bundles occurred in the review. For example, the first block means that 32 bundles of maximum 2 practices were found. The three largest are one 19x (Manos et al., 2006) and two 16x practice bundles (Costa et al., 2017, Sloan et al., 2014). Given that larger bundles still allow causal connections between two practices to be derived, we decided to focus on bundles with more than 2 practices and their occurrences.

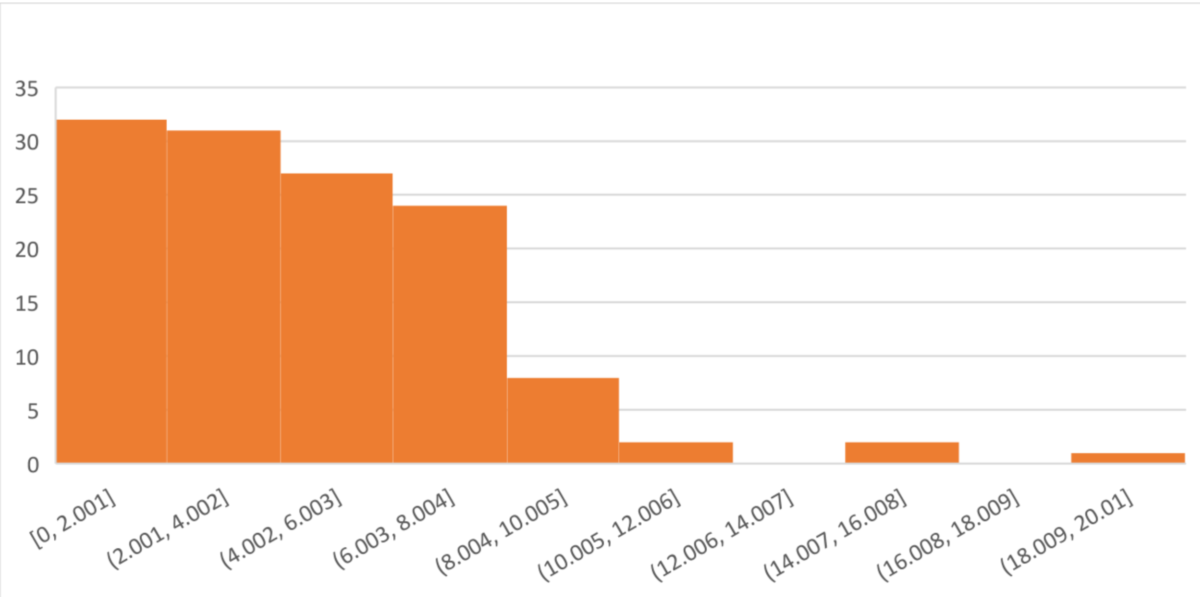

Figure 7: How often does an x-practice bundle occur?

For this reason, we compared the individual bundles for shared combinations of practices through truth table analysis. The summarised results can be found in Appendix A5.

Our review analysis generated the following identification of the most frequently occurring combinations of Lean practices:

5 Practices - combinations

3x [RPIW + Visualisation + Standardisation + process flow analysis + Kanban]

2x [Integrated process redesign + Visualisation + Standardisation + 5S + Process flow analysis]

4 Practices - combinations

5x [5S + VSM + Standardisation + Integrated process redesign]

5x [Training + VSM + process flow analysis + Standardisation]

3 Practices - combinations

14x [5S + VSM + Standardisation]

10x [RPIW + Visualisation + Standardisation]

This identification distinguishes between the different sized combinations of practices. Additionally, it shows the exact combinations and the number of studies in which they occur. The largest combination covered by the minimum of two studies consists of only five practices.

While both standardisation and VSM are the most commonly occurring individual practices (each >50), they also seem to frequently occur in combination. In the review, we found 27x times when both practices occurred together and 26x times when they were coupled with another practice. In 14x of these cases, they were combined with 5S and in 5x cases they were combined with integrated process redesign. These practices seem to play a strong linking role in Lean healthcare bundles. (Mazzocato et al., 2010) also identified these practices as two of the most relevant elements of a Lean implementation. Only a few studies have specifically investigated the relationships between them. An explanation for this may be that

VSM is used as an analysis practice while standardisation is the consequence of identified improvement potentials.

Furthermore, nearly all studies treated Lean bundles from the perspective of equally strong practices. Conversely, discussions around the superiority of individual practices or small sets were only done in the frame of socio-technical argumentation. While the socio-technical lens is essential for the understanding of Lean bundles as they allow to distinguish between specialisations of bundles as well as the combination of practices, it also limits the range of found combinations to results in the range of socio or technical practices. Thereby, results of individual practices or small sets functioning as performance drivers in a bigger bundle are constituted to the socio-technical argumentation, rather than the actual drivers.

Nevertheless, it is apparent that some practices occur more frequently than others, whether individually or in combination, even if few studies attempt to explain these occurrence differentials.

5.2.4 Organisational background

Following these initial findings, a decision was made to further investigate the service type, organisational size, and implementation links of these combinations in order to better understand their unknown drivers. Most reviewed studies examined A&Es, operation theatres, in-hospital pharmacies, community hospitals, and radiology wards. Several other departments/wards or subunits were found, indicating a far-ranging usage and adaptability of Lean in healthcare. Specifically, value-stream-bound implementations in, for example, cardiovascular care, have been identified as interesting within the existing literature (Schoonhoven et al., 2011) as they cross multiple departments and sometimes even multiple levels. Thus, such Lean applications seem to be more complex, as they require the inclusion of strategic practices and the enforcement of coherence between the Lean approaches in the individual subunits. Often, multiple subunit implementations come about because they start with an experimental or trial implementation in one subunit, which then extends to the surrounding subunits (Kim et al., 2009).

Comparing the identified combinations of practices, we identified similarities between the organisational backgrounds in which these studies took place. For example, the combination of the Lean practices training + VSM + process flow analysis + standardisation was found in the following settings:

Study	Implementation setting
(Sloan et al., 2014)	Multi-unit
(Costa et al., 2017)	Multi-unit
(Improta et al., 2018)	A&E
(Eriksson et al., 2016)	Nursing units
(McDermott and Venditti, 2015)	Value-stream (discharge)

Table 1: Different settings for training + VSM + process flow analysis + standardisation

Additionally, the bundle of 5S + VSM + standardisation + integrated process redesign was found in these settings:

Study	Implementation setting
(Costa et al., 2017)	Multi-unit
(Johnson et al., 2012)	Multi-unit
(Kane et al., 2015)	A&E
(Chadha et al., 2012)	A&E
(Ben-Tovim et al., 2008)	A&E > organisation-wide

Table 2: Different settings for 5S + VSM + integrated process redesign + standardisation

Both examples might prompt the conclusion that there is a relationship between either A&Es or multi-unit settings, and the choice of VSM + standardisation.

However, the systematic review also found studies in different organisational settings with the same two practices, and indeed the same setting with different practices.

As a result, we conclude that while the organisational background seems to have a certain impact on the choice of practices, the results of this review do not themselves support the causal argument that would strongly link the organisational background to the choice of practices.

5.2.5 Review discussion

The following paragraph will discuss the major findings of the additional literature review presented below in detail:

- Similar Lean practice combinations in different implementation settings
- VSM, process flow analysis, 5S, RPIW, process redesign and standardisation occur in 25% of all found practice combinations
- Similar Lean practice combinations have different performance impacts
- 75% of all combinations have less than 8 practices
- In healthcare, combinations of Lean practices exist in small to medium sized flexible bundles rather than big, fixed bundles
- Most of Lean papers in healthcare still become published in medical and public health journals rather than operational or management journals

The aim of this work was the creation of a structured review of the variety of scientific contributions that investigate the empirical impact of Lean configurations in healthcare organisations, thereby answering the first part of the research question of ‘which combinations of Lean practices lead to superior performance improvement in healthcare subunits?’. Overall, the literature review found sufficient indications for the first research question and limited evidence for the second research question.

Taken alone, the literature review therefore contributes to the existing discussion about sets of Lean practices generally, and in healthcare in particular. The results enhance, verify, and update past findings of other Lean practice literature reviews, especially the work of (Bucci et al., 2016), (Costa and Godinho Filho, 2016), (Hadid and Afshin Mansouri, 2014), and (Punnakitikashem, 2013) trough increasing the sample and requirements of empirical validation .

This review provides an updated and verified overview of the most frequently used combinations of Lean practice in the healthcare environment and sheds a spotlight on their relationship, as seen in the bundles in which they occur. Some combinations are interesting because their practices generally appear in a specific configuration, while other practices occur alone. This study provided a lens on empirically validated practice combinations, creating clarity about the rate of occurrence of each combination. At the same time, it also created a wider lens by doubling the number of included practices compared to previous literature reviews. The overview about the investigated practices can be found in Appendix A7.

The variety of distinct bundles identified in our review is consistent with previous studies that note the organisational heterogeneity in Lean supported healthcare settings. While our hierarchy of occurrences of Lean practices is as in line with previous work, it also shows the rise of practices that have attracted less notice. In particular, VSM, standardisation, and process/patient flow analysis are the most commonly occurring practices while training, RPIW, and visualisation are the fastest rising. One reason for this development would be the recent advances made in explaining and distinguishing between technical and social practices (e.g. Hadid et al., 2016), which have directed increased attention to the social side of Lean. Another reason would be the focus of previous literature reviews on the distinction between Lean service and Lean manufacturing, which meant that a narrower definition of Lean practices was used in this review.

It is worth noting that the lack of studies on Lean implementations with negative outcomes polarises the discussion about Lean and creates the illusion of a pure improvement system. It may be that switching from studies that focus on a single Lean implementation to those that incorporate multiple implementation studies might allow for a more critical perspective on the potential of Lean development. Comparing Lean studies is inherently difficult as there is a lack of clarity in the understanding of individual practices. Not all studies provide definitions of every individual practice. Therefore, the context, references and exact processes of the practices were compared. Additionally, there are a wide variety of training perspectives and types of Lean coach, which creates unknown biases. While the heterogeneity of healthcare organisations and their Lean implementations make an overly structured approach impracticable, an adjustable range of definitions might improve comparability between Lean studies. Another limitation of the selection criteria for this review is that its constituent studies primarily focus on operational outcome indicators. While the number of studies investigating medical problems have increased in recent years, there remains many additional fields to explore.

One of the strengths of Lean healthcare is that its practices do not just occur in narrow settings but often cross borders to multi-unit and multi-level implementations. Future studies investigating these relationships in more depth can not only contribute to the Lean healthcare discussion but also generate insights for overall Lean theory.

This additional literature review has a few minor contributions which directly impact this study. First the largest sample of Lean practice implementations enabled the author to identify Lean practices in the empirical work. Second it provided later explanations for empirical found practice combinations and how they work together. Third it provided a first indication that single or small practice sample implementations can occur.

The major contribution of this review is the identification of more advanced relationships in Lean combinations and, where included in the studies, aligned Lean bundles. While we found evidence of a wide variety of combinations of practices (See Appendix A6 and A7), we also noted that some configurations and practices play an apparently bigger role than others. Some combinations occurred more often or were more unified (See Appendix A5), compared to others that seem to prefer to exist alone. Although, this offer a new perspective on the sets of Lean practices it is nevertheless merely the starting point for the empirical evaluation of relationships between Lean practices in Lean bundles that follow this chapter.

5.2.5.1 *Summary of main articles and their contribution to this study*

Given the wide reach of this thesis, several references were evaluated and have an influence on this study. Of these, only a handful can be considered as having had an essential impact. Hence, we set out below the main theories and their influence on this work:

Theory	**Lean Bundles**	**Lean Service**	**Complex Lean Health**
Studies	(Shah, 2002, Shah and Ward, 2003)	Hadid et al. (2016)	(D'Andreamatteo et al., 2015, Ferreira et al., 2019)
Topic	Four Lean manufacturing bundles, typologies of Lean, bundle interrelations	Socio-technical bundles, relations between different bundles in the service industry	Lean in healthcare, current lens on Lean implementations divided
Implications for this dissertation	Conceptualises and investigates the relationship between Lean practices in sets. Additionally, these studies provide typologies that allow for categorisation of Lean organisations.	Offers advanced explanations for inter-bundle relationships and provides constructs of Lean service bundles and their content.	Enables a preliminary understanding of possible relations between Lean practices. Highlights complexity of the healthcare setting.

Table 3: Summary of main articles and their implications for this work

The fundamental Operations Management framework is formed by the Lean bundle theory of (Shah and Ward, 2003). In addition, Shah (2002)'s doctoral thesis provides typologies for Lean organisations. As her work is focused on the manufacturing industry, a validation for the service industry is required, and this was provided by (Hadid et al., 2016). However, this was in relation to the theoretical framework rather than the specific bundles. Given Hadid's exclusion of healthcare from the sample and the shortage of Lean healthcare studies, a further abstraction of the theory is required.

While (D'Andreamatteo et al., 2015) provided a preliminary grasp at some smaller relations between Lean practices through a system-wide lens in healthcare, more detailed examinations are the aim of this research. That study also improved understanding about the importance of multifunctionality and heterogeneity of tasks and processes in healthcare. This enabled (Ferreira et al., 2019)'s argument of a complex healthcare setting, which is what this research will discuss and address in the next chapter.

6 THEORETICAL AND METHODOLOGICAL FRAMEWORK

6.1 THEORETICAL FRAMEWORK

6.1.1 Structure – Chapter

This chapter will describe the theoretical framework of this study, especially the methodological background and the method itself. First, the chapter will shed light on the ontological and epistemological background of the study. Afterwards it will illustrate the complex situation in which the study takes place before exploring how both the organisational setting and the Lean bundle theory in healthcare put constraints on the research aimed at understanding them. Second, the suitability of several theorising modes for these settings will be discussed. Third, the chapter will argue in support of adopting a configurational lens for Lean and a subunit lens for healthcare organisations. Combining both perspectives provides an argument for fsQCA. An extensive introduction of this still novel methodology is then presented. All the subsections in this chapter will be supported by references from a wide range of literatures. Essential to this new method is the concept of cross-case comparison, which will be introduced. Additionally, the chapter will discuss different types of necessity and sufficiency tests, and their advantages and disadvantages. Finally, the chapter will close with the introduction of the procedure of calibration and the differently valued fuzzy sets.

It should be noted that this chapter shares stylistic similarities with the literature review and analysis chapters, begging the question of why this chapter was integrated with them. This was attempted several times. However, each time, the contextual and methodological backgrounds of the study vanished into the depths of the extensive Lean literature and the empirical fsQCA work. Therefore, it was decided to include both backgrounds into a separate chapter, specifically emphasising their links to each other. The contextual background provides the essential base for the application of fsQCA. Additionally, fsQCA and its analytical tools are still a novel methodology that requires an extensive introduction if the reader is to understand their underlying principles, such as the assumption of equifinality. If these principles are not understood, the reader will struggle to grasp the analysis and contribution of this work. These principles will be comprehensively discussed in the following pages.

Figure 8 illustrates the structure of this chapter:

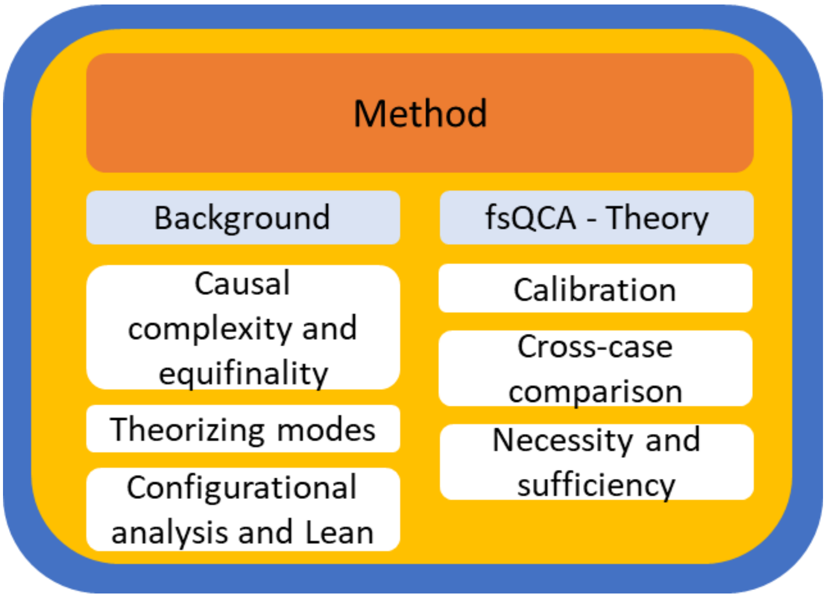

Figure 8: Structure – Theoretical framework

6.1.2 Ontological and epistemological background

Generally, research is unconsciously or consciously influenced by different categories of assumptions (Burrell & Morgan, 2017). These assumptions can be distinguished between ontological assumptions about the nature of realities, which the researcher may encounter and epistemological assumptions surrounding the human knowledge about the world (Saunders et al., 2009). These assumptions have a widespread impact on how research is done, which research questions and methods are chosen as well as how data is analysed and findings understood (Crotty, 1998). A few typologies have developed combining shared values and beliefs regarding the understanding and viewing of reality and knowledge. These are also known as paradigms (Kuhn, 1970). Patton (2002, p. 69) described them as "a way of thinking about and making sense of the complexities of the real world".

This research study can be categorized into the pragmatism paradigm. Originally, the notion of pragmatism developed from the ancient Greek term "pragma", meaning action (Pansiri, 2005). The paradigm of pragmatism means that human actions are always dependent on past experiences and beliefs (Kaushik & Walsh, 2019). In pragmatism reality is seen as continuously reshaping and being discussed based on a changing environment.

Pragmatists think that peoples' experiences are unique and that therefore their worldviews might be close but never the same. At the same time people might have shared experiences which might lead to similar beliefs and understandings. However, these overlaps may vary. This might lead to situations where people with similar backgrounds react differently or similarly in the same situation. This depends on the exact degree of shared experiences, beliefs and situational understandings. Consequently, pragmatism owns the ambiguity of worldviews which can be unique and/or intertwined at the same time (Morgan, 2014).

In the light of this study, pragmatism views the healthcare environment as complex and problem rich. Furthermore, it can be found again in the view of Lean practices which operate in improving ways that in one hospital ward may be successful but are not in another one. Finally, in the following paragraph the lens of pragmatism can be seen in the choice of an adaptive novel method with a mixed data set in a complex environment, trying to improve the understanding of process improvements in healthcare.

6.1.3 Causal complexity and equifinality

The literature review showed multiple different combinations of Lean practices in different settings, each with different performance outcomes. In the following section, these combinations and findings will be further discussed. Additionally, the notions of causal complexity and equifinality will be introduced.

Given that practical, empirical, and theoretical knowledge advances are intertwined, gaining a deeper theoretical understanding about the contents of a Lean bundle requires comparing the different configurations of Lean bundles, each with the other. This study will distinguish between the theory surrounding Lean bundles and practically driven Lean bundles. Nevertheless, the link between both remains object of interest for this study.

This study plans to address this by analysing different performance outcomes and linking them to specific sets of Lean practices. This study also aims to question the limits of the internal consistency of Lean bundles in practice by investigating the necessity and sufficiency of specific Lean practices in these bundles. In this case internal consistency describes how well individual practice fit together in achieving the bundles performance objectives. However, before these aims may be addressed, we must first take a step back and describe the theoretical framework that links these aims with each other and with the chosen methodology.

While the origin of Lean lies within the manufacturing industry, the numbers of Lean studies in the service industry have been rising, contributing new insights for Lean that are derived from the context of a distinct industrial environment. (Prajogo, 2005)'s work indicated that, when comparing the manufacturing with the service industry, the manufacturing sector is often seen as less differentiated, diversified, and specialised. As a result, specific differences can be found when comparing work specialisation, simultaneity, and intangibility as well as the level to which customers (or patients) are present during the production/service process

(Büyüközkan et al., 2011, Sampson and Froehle, 2006, Bowen and Youngdahl, 1998). Overall, these differences give rise to the assumption that the manufacturing industry is homogenous by nature, while the service industry tends to be heterogeneous.

While this study will specifically address set theory in later chapters, set theory has general links to Lean bundles, given that each bundle is also considered to be a set. In the literature review, several sets of different practices have been identified. However, the identified combinations of practices were often too diverse for a distinct linkage in common sets to be established. As several studies have highlighted the impacts that practices exert on performance, this study aims to conduct a wider performance comparison through linking some sets of practices to specific performance outcomes. Unfortunately, a comparison of all sets in terms of their performance impacts was not possible because too many used different performance variables or had a different organisational setting. However, in some cases it was possible to isolate and compare sets in terms of performance, and the results allowed a first glimpse of the complex relationships at play in the mix of practices and their performance impacts; this will be further discussed in later sections of this thesis. To illustrate these complex relationships, three studies from the review are presented in Table 4 below:

	(Kane et al., 2015)	(Ng et al., 2010)	(King et al., 2006)
Waiting time reduction	17%	29.7%	13.7%
Patient satisfaction improvement	4.1%	3.3%	
Lean practices:			
VSM	x	x	x
5S	x		
Process flow analysis			
Visualisation		x	
Standardisation	x	x	
RPIW	x	x	
Change facility layout	x	x	
Training	x		x
Segregating complexity			x
Obtain management support			x

Table 4: Differences between practice set performance

The above Lean implementations were conducted in different emergency departments in similar organisational environments. They additionally all included technical and social practices. To summarise, (Kane et al., 2015) and (King et al., 2006) achieved similar waiting time reductions (17% and 13.7% respectively), while (Ng et al., 2010) achieved a 29.7% reduction. The only clearly visible differences between the three studies would be the presence of visualisation and the absence of training in Ng's implementation. However, these cases can present only a limited explanation of the performance differences because the authors do not note why other practices were omitted or otherwise highlight their usage. However, it can be assumed that a certain, more detailed, dependence on other factors may exist, leading to the creation of different practice sets with complex performance links.

Although these studies had a shared organisational setting of the emergency department, this is often not the case in practice. Healthcare organisations, especially hospitals, consist of several highly diverse and specialised wards/units. Furthermore, the individual importance of a single unit to the organisation as a whole is different for manufacturing versus service versus healthcare organisations. For example, the assembly line of a factory is often the key element of the manufacturing process. Healthcare is very different. When a patient enters the hospital, the pathway that will be followed is unclear and will only become clearer as the patient progresses through the process. In such cases, the comparison must be carried out at subunit level to properly consider administrative intensity and vertical complexity (Carillo et al., 1991). While this allows for a more detailed and stable comparison, the different specialisations of these subunits (e.g. a comparison of A&E and rehabilitation) hinder the usage of traditional theoretical and empirical models.

Overall, industrial heterogeneity, setting dependence, the different bundle performance links, and organisational multiplicity add up to a difficult and diverse situation. This can also be described as causal complexity. The literature suggests several definitions of causal complexity. For example, 'the effect of one variable or characteristic can depend on which others are present' (Jervis, 1998, p.35). A somewhat different choice of words came from (Ragin, 2014, p.20), who describes situations in which 'an outcome results from several different combinations of conditions.' Based on these definitions, (Misangyi et al., 2017) identified that conjunction, equifinality, and asymmetry form the theoretical frame of causal complexity. Conjunction occurs because multiple conditions (here, Lean practices) share links and relationships with each other and with the outcomes of interest. The literature review in the previous chapter showed that different combinations of practices constitute different paths, leading to performance outcomes. This is known as equifinality. Similarly, the review indicated causal asymmetry by presenting different cases that achieve comparable outcomes. In some of these cases the same practice and/or combination was present while being absent in another. This further supports the assumption of a causally complex case. While a higher level of causal complexity is assumed here, theorising in such an environment might become difficult.

6.1.4 Modes of theorising

In the past, three different modes of theorising could be identified: universalistic theory, configurational theory, and contingency theory (Delery and Doty, 1996, Guest, 2011). Furthermore, (Colbert, 2004) proposed adding complexity theory to the list of modes.

Numerous scholars have discussed the differences between universalistic theory and contingency theory. Guest (2011, p.7)) referred to the debate thus: 'We might have thought that the debate about a universalistic versus a contingency approach would have been settled long ago in favour of contingency theory.' While contingency theory dominates the theoretical discussion, it is considered in the empirical dimension to be outperformed by universalistic theory (Combs et al., 2006).

According to Colbert (2004), the major differentiating characteristics of these theories are the level of explanation of interactions and system complexity. The two best known theories—universalistic and contingency—give scant attention to interaction effects. Both approaches were developed to explain a wide range of phenomena. The universalistic perspective, also called the best practice approach (Osterman, 1994, Pfeffer, 1994) allows for a direct link with certain corporate practices, e.g., HRM practices (Delery and Doty, 1996) with improved operational performance. While universalistic theory simplifies the process of identifying performance changes, it is disconnected from interaction effects within a set of organisational variables. It thereby provides an argument for the assumption of specific additive sub-group variables (Gerhart and Milkovich, 1990). Conversely, contingency theory suggests that alignment between corporate practices and other external and internal organisational elements is required. This is assumed to be necessary for coherent business strategy decisions (Delery and Doty, 1996, Lepak and Shaw, 2008, Marler, 2012).

In contrast to universalistic theory and contingency theory, configurational theory focuses on system interaction effects whereas complexity theory emphasizes interactions and assumes permanent change in the system, constituted by feedback loops.

The configurational mode of theorising suggests that certain sets of practices, like HR practices (e.g. Samnani and Singh, 2013), result in alignment within configurations (Delery and Doty, 1996, Martín-Alcázar et al., 2005) and provide competitive outcomes (Lepak and Shaw, 2008). The philosophical foundation of configurational analysis builds on the concept of a 'Gestalt'. Under the notion of Gestalt, quality can be understood as a product of uniformity of the whole construct and manifold of the construct's parts (Ehrenfels, 1890). According to (Venkatraman, 1989) the occurrence of low specificity among variables and criterion-free anchored specification can be used to create the concept of a Gestalt. From this observation, definitions emerged of the notion of a configuration. For example, (Miller and Friesen, 1978) described configurations as commonly occurring patterns of relations or attributes that share internal cohesion. The underlying idea behind this description is the assumption of different characteristics that commonly occur together. This assumption forms a fundamental part of configurational analysis, enabling a better understanding of interaction effects.

Generally, the configurational approach recommends focusing on the mutual influence of a wide set of variables (Meyer et al., 1993). It thereby attempts to model relationships through the use of typologies. Typologies are simplified constructs of sets of similar configurations.

According to (Miller, 1996) and (Burton and Obel, 2004), typologies are theoretically valuable for several reasons. On the one hand, the multidimensional nature of integrated typologies allows configurational thoughts to recognize the complex and individual nature of organisations. On the other hand, typologies can show that both competitive advantage and fit are sometimes connected to relationships between multiple characteristics instead of being connected to a single aspect.

While the typologies of (Miles et al., 1978) and (Mintzberg, 1979) assume ideal types, in practice, hybrid types—combinations of the initial ideal—can occur (Doty and Glick, 1994). This principle is called 'equifinality'. Thus, 'A system can reach the same final state from different initial conditions and by a variety of different paths' (Katz and Kahn, 1978, p. 30). Equifinality is one element of configuration theory, which suggests that organisational constructs consist of '…tightly knit and fairly stable constellations of mutually supportive elements put together into a thematic synergic whole' (Lamothe and Dufour, 2007, p. 68). Compared to linear cause-effect-relationships, equifinal typologies support the existence or creation of hybrids of ideal types of an effective organisation. Although linearity sees hybrids as ineffective organisations, equifinality argues that organisations use hybrids when they must respond simultaneously to conflicting contingencies (Doty et al., 1993).

Another mode of theorizing is complexity theory. It offers a view of organisations that differs from configuration theory. Complexity theory is founded on the assumption of a convoluted organisational structure, which includes dynamic systems of interactions (Colbert, 2004, Rivkin, 2000). Furthermore, interactions are considered to be adaptive rather than simply agglomerations of static entities (Mitleton-Kelly, 2003). However, while complexity theory supports organisational theory in explaining interrelations in the corporate structure, it is debatable whether the theory can explain the bigger picture in organisational theory and analysis (Fiss, 2007).

Both configurational theory and complexity theory are specific modes of theorising that have strong downsides; this limits their application in system-wide applications. Complexity theory asserts that systems are unpredictable, but the theory is self-limited by a set of order-generating rules (Burnes, 2005). Conversely, configurational theory lacks empirical support from studies that validate a wide variety of configurations instead of a small sample. Most studies have investigated between 1 and 5 configurations by reducing their samples towards the most likely configurations. Thereby, they neglect configurations that might gain importance in different settings.

This study aims to explain the relationships between Lean practices and their influence on performance. The Lean bundle theorem (Shah and Ward, 2003) allows the explanation of known relations in Lean, but it overlooks unknown relations. Given the context, this study requires a theoretical foundation before it can choose an empirical approach. Theories are designed to explain, frame and understand phenomena as well as challenge and advance existing perspectives within the limits of binding assumptions and constructs (Abend, 2013).

Both complexity theory and configuration theory are suited to fulfil this task. At first glance, complexity theory seems to have a better fit because it takes into account the unpredictable interrelations that have been neglected by traditional configurational perspectives (e.g., Doty and Glick, 1994). However, several recent advances of configurational theory (e.g. Fiss, 2011) appear to address this weakness.

6.1.5 Configurational analysis and Lean

(Shah, 2002) used to view the Lean manufacturing system as a Gestalt, a construct of internally consistent and coherent practices. The implementation of Lean practices could be conducted alone or in combination with other elements. Furthermore, she defined two archetypes of Lean implementations: one with a full complement of JIT, quality control tools, and employee-, supplier- and customer involvement, and the other with none of these Lean elements.

While her work was at, if not beyond, the cutting edge, it was also restrained by the founding stage development of the configurational lens (Delery and Doty, 1996, Miller, 1996, Mintzberg, 1979, Miles et al., 1978) and the limited acceptance of mixed method approaches (Ragin, 2014). Although (Shah and Ward, 2003) were able to deduct and prove the existence of Lean bundles, a deeper analysis of individual practices and their relationships beyond the boundaries of the bundles was denied to them. For example, (Shah, 2002) used a two-stage cluster analysis to investigate causal structures in Lean configurations. However, cluster analysis was found to be inferior to newer methods, such as QCA (Hotho, 2014), as it only gives restricted insights into typologies, their internal causal architecture and neutral permutations, and the asymmetry of causal connections with outcome indicators (Fiss, 2011).

Cluster analysis was most likely chosen because at that time there was a void of methodological tools able to apprehend causally complex problems (Fiss et al., 2013a). Furthermore, cluster analysis lacks detail when used to investigate the relationships between distinct organisational elements (Fiss, 2007). Specifically, cluster analysis is able to view only the patterns between combinations of causal elements rather than the elements themselves (Whittington et al., 1999). This constitutes an issue in that it might result in incorrectly viewing cases as very similar because of a limited set of important attributes for the configuration. However, there may be differences in the attributes that, although less important to the configuration, constitute the majority of attributes (Fiss, 2007).

The neo-configurational lens argues that the relationships between causal conditions that go beyond the boundaries of their constituting configuration require a stable link between theory and method (Misangyi et al., 2017). In the context of Lean, this could be a practice having primarily impacts on the bundle and secondarily impacts on another bundle. This would create a causally complex situation. QCA provides an approach considering causal complexity. While the method is similar to regression, ANOVA, or cluster analysis, it provides more robustness during the examination of causal configurations, encompassing their relationships as well as their impacts on outcome variables (Fiss et al., 2013a, 2007).

Hence, this study's choice of fuzzy-set qualitative comparative analysis (fsQCA) to investigate the inter-relationships in Lean practice bundles and the extent to which they exert (non-) superior influences. Data were gathered through an extensive documentation review, in-depth semi-structured interviews, and observations in 5 hospital-sized care facilities.

6.2 fsQCA – Theory

The following paragraphs will introduce the underlying theoretical framework of set theory, as well as describing the subsequent analysis procedure of fsQCA in this study. As fsQCA contains several quantitative and qualitative elements that operate in the background, this section will focus on the fundamental concept of the method; in particular it will shed light on parts that are otherwise unperceived by the practical analysis.

According to (Jech, 2013), set theory is characterised by membership, which amounts to a specific type of relationship. For example, an element K is a member of the set L ($K \in L$). In general, sets can be considered equal if they share the same constituting elements. However, set theory allows the consideration of relationships between constituting elements. Thus, where all members of a set M are at the same time also members of another set L, M can be considered a subset of L ($M \subseteq L$). According to (Klir and Yuan, 1995) this relationship is known as set inclusion. Both applied and theoretical studies have a wide variety of set typologies, allowing for investigations of completely distinct phenomena.

The first applications of QCA tended to use crisp sets (Ragin, 2014). A crisp set categorizes elements through a traditional quantitative distinction between full membership and non-existent membership. In comparison, more recent studies use fuzzy sets because these incorporate differing degrees of membership (Ragin, 2008, 2005, 2000). Both the configurational frame of the Lean bundle theory (Shah, 2002) as well as the causal complexity of Lean healthcare sets provide the foundation for the application of fsQCA here.

Through fsQCA, a phenomenon is investigated by classifying elements either as constituting attributes or causal conditions. Elements—internal and external—that influence outcomes of interest can be considered to be causal conditions (Rihoux and Ragin, 2008).

6.2.1 Calibration

In this study, the phenomena are high/low operational performance whereas the lean practices are the causal conditions. The attributes form independent sets.

During the calibration phase, each case (where subunit = hospital ward) in these sets receives a membership score. The membership score represents the extent of its membership in the set of attributes (Fiss, 2007). According to (Ragin, 2008), scores are based on the membership model. Normally, it is either an x-value fuzzy set with predefined categories or a 'continuous set'. The choice depends on the number of cases and logical reasoning. Cases are evaluated by the degree of set membership. Each set is regarded as a different typology of attribute configurations. (Ragin, 2008) suggested two different calibration approaches: the direct method and the indirect method. The first of these uses, in its unmodified form, three qualitative anchors, as follows:

Direct calibration

0 – fully out

0.5 – crossover point

1 – fully in

In other words, 0 represents the threshold for non-membership, 0.5 represents the point of maximum ambiguity, and 1 is the threshold for full membership.

The indirect method, in its unmodified form, uses six quantitative anchors to express the groupings. The method thereby permits the investigation of additional cases that are partly in or partly out.

Indirect calibration

0 – fully out

0.2 – mostly but not fully out

0.4 – more out than in

0.6 – more in than out

0.8 – mostly but not fully in

1 – fully in

However, in practice the most common calibrations use either the three-, four-, or six-values or the 'continuous' fuzzy set (Ragin, 2005). This last provides a finer-grained calibration but has the disadvantage of requiring a large sample size. In other words, the choice of the calibration involves a trade-off between very finely grained results and a high consistency in the resulting set of cases. In consequence, the decision about fuzzy set calibration is context dependent.

6.2.2 Cross-case comparison

This paragraph dives deeper into scientific discussion around the essential part of QCA, the cross-case comparison. First it introduces the method of cross case comparison, followed by different perspectives between methods and closes with recent scientific discussions and their impact on this work.

The previously discussed combinations of attributes form the foundation of cross-case comparison. Rihoux and Lobe (2009) offered an argument for why cross-case comparison is an advanced form of cross-case analysis. Under cross-case analysis, the research approach is already understood, assisting the progress of a comparison of the differences and commonalities of processes, events, and activities of interest (VanWynsberghe and Khan, 2007). Conversely, cross-case comparison combines the holistic conception of a case (in which it is seen as a complex entity) and the capability to investigate complex causality between multiple cases (Schlosser et al., 2009). Through this ability to examine multiple cases, the approach supports generalisation and replication. Thereby, it assimilates analytic elements while keeping hold of the holistic overview of the phenomenon of interest.

The combination of these features allows an in-depth investigation of several cases to be conducted, while still being able to see the bigger picture and generalize. Thus, fsQCA supports the configurational lens by providing a clearer perspective of multiple configurations of causal conditions that achieve the same outcome. (Katz and Kahn, 1978) referred to this scenario as 'equifinality'. Compared to 'unifinality' (Gresov and Drazin, 1997, Galunic and Eisenhardt, 1994), this theoretical vision enables the consideration of different combinations of Lean practices leading to high (low) performance. fsQCA extends the reach of equifinality by providing a tool to investigate the influence of each path (Ragin, 2006). Thereby, the consideration of equifinality combined with fsQCA enables unique scientific contributions.

Combining a qualitative and quantitative lens on causal combinations is a core element of fsQCA. Both perspectives are considered through a cross-case comparison of Boolean reduction (Lacey and Cohen, 2015), which supports the in-depth comparison of causal blueprints. As fsQCA may create exponentially large samples of causal combinations (e.g. qualitative cases) it requires a statistical tool to compare those combinations and find an underlying pattern in a reasonable size for the researcher. Furthermore, fsQCA can increase its focus on either the qualitative or quantitative side depending on the context of the cases and causal blueprints.

In this case the question becomes: How do intertwined Lean practices work or not work, and in what ways do they impact operational performance?

Fiss (2007) and (Kogut and Ragin, 2006) distinguished the following four main parts of Boolean algebra:

I. Using binary data;

II. combinatorial logic;

III. logical interpretation utilising Boolean algebra operators; and

IV. Boolean simplification minimising the presented paths of causal complexity.

Overall, Boolean reduction forms a pragmatic part of fsQCA as it cuts down to a manageable size the number of possible causal recipes by decreasing the arithmetic operations as well as the terms (Whitesitt, 2012), without losing the cogency of the results. Ragin (2008) distinguished the results as either complex, intermediate, or parsimonious; classifications that have recently attracted attention. (Baumgartner and Thiem, 2017) challenged the complex and parsimonious solutions, declaring them both obsolete. This was criticised in turn by (Duşa, 2019), who asserted that (Thiem, 2016) had misjudged the possibility of using necessary expression for the determination of incoherent counterfactuals, leading to wrongly weighted conclusions. Additionally, Dusa (2019) has suggested that (Baumgartner, 2009)'s CNA approach, a configurational comparative approach using Boolean logic to structure causality, to truth table analysis is a subvariant of QCA since it has similar quasi-counterfactual analysis and algorithms. Furthermore, he hints that Baumgartner and Thiem (2017)'s challenge of the parsimonious and complex solutions might be wrong. However, this challenge has not thus far been accepted by other researchers. Therefore, this study intends to consider and test all three possible solutions.

6.2.3 Necessity and sufficiency

fsQCA is not a static mixed method as it can be adjusted towards one method or another if so required by the empirical context, without losing its mixed character. This characteristic allows for the examination of structurally diverse sets that lead to the same consistent outcome. Hence, fsQCA supports the investigation of different levels of importance of set attributes. Specifically, fsQCA may test for the necessity and sufficiency of causal conditions and attributes in a configuration (Ragin, 2014) and, according to (Dul, 2016b), a necessary causal condition enables a specific outcome's existence. In reciprocity, this means that the lack of a necessary causal condition will result in the absence of an outcome. Conversely, a sufficient causal condition will create a specific outcome or can be described as the insurance of that outcome's existence. In some cases, the causes are so strong that their absence results in the absence of the outcome. While these causes might not be sufficient to cause an outcome on their own, they might still be a necessary element of the causal entity. This means that, irrespective of the other causal elements, the necessary cause is extremely likely to be present in the causal entity (Dusa, 2019). If a combination of causal elements is necessary for an outcome of interest, this outcome forms a subset of the causal combination. If, on the other hand, a causal combination is sufficient to achieve an outcome of interest, the outcome constitutes a superset of this causal combination. Causal combinations do not need to consist of multiple elements; a combination can be a single element or even the absence of elements. This study tries to identify a combination of Lean practices that are enough and/or required to enable superior performance improvement.

6.2.4 Uniqueness of fsQCA

In comparison with other methodical approaches, fsQCA empowers an asymmetric lens of configurational theory. (Fiss et al., 2013a, p. 192) defined causality with assumed asymmetry as the 'absence of causal conditions associated with an outcome not leading to absence of the outcome'. While the absence and presence of a phenomenon require different analyses, these results can frequently lead to an improved understanding of the phenomenon and its relations (Schneider and Wagemann, 2010).

Conversely, the combination of quantitative and qualitative analysis elements means that fsQCA has its limitations. For example, the heavy dependence on context-based subjective manual calibration leads to limited robustness (Krogslund et al., 2015). Additionally, too many causal conditions will decrease the meaningfulness of the analysis (Schneider and Wagemann, 2010) by pushing up the number of logical remainders as well as leading to complex empirical results. Later, we will address these issues through elicitation and sample focus.

Conventional research methods, mostly linear, specialise in the recognition of a link between one or several attributes that leads to one or more outcomes. Contrariwise, fsQCA moves beyond the examination of one link and provides a stronger understanding of these links and the relations between them. Consequently, fsQCA contributes several advantages to this study. For example, the expectations of equifinality (Ragin, 2014) and/or conjunctural causation (Ragin, 2008) strengthen the assumption that diverse combinations of Lean practices result in superior (or not) operational performance. In conclusion, this study assumes that isolated practices rarely achieve these kinds of performance outcomes.

6.3 Summary

This chapter has presented and discussed the theoretical framework of this study. It has also highlighted the causal complexity in Lean (Ferreira et al., 2019, Saurin et al., 2013) and in healthcare organisations. It has examined contingency theory, universal theory, configurational theory, and complexity theory as possible answers to the heterogeneity in healthcare and multiple practice combinations. This chapter then argued in support of a configurational theory, specifically neo-configurational theory, because it enables a finer grained lens.

Furthermore, it has proposed a subunit lens (Leatt and Schneck, 1982, Alexander and Randolph, 1985, Carillo et al., 1991) addressing the impacts of the complexity of the study's setting.

A strong argument in support of a fsQCA approach emerged from the proposed configurational lens. In this chapter, this still novel method was introduced and described. As the method contains several possible adaptations and sub-methods, the suitability of these to the expected setting was discussed in detail. The practical side of the analysis will be elaborated upon in the next chapter.

7 EMPIRICAL RESEARCH

7.1 STRUCTURE – CHAPTER

This chapter will present the analysis for both studies. First it will focus on study one, followed by a comment about the lessons learnt from that study before proceeding to discuss study two. The analysis of study one will begin with a detailed description about how access to the hospitals and their staff was obtained, followed by an illustration of the actual collection of reports and interviews. Thereafter, this chapter will explain how the collected data was transformed through data analysis into useable scores for fsQCA. Based on these scores, the paragraphs that follow will discuss the calibration of practice scores, as well as the key performance indicators. Finally, the analysis of the first study will close with the presentation of the truth-table analysis as well as the necessity tests.

The second study will use the experience gathered from the first study to incorporate adjustments to the content analysis and calibration to strengthen their validity. Study two's methodology will commence in a similar way to study one with a description of the data collection processes and the addition of observations. The elicitation phase is new to the second study, and it describes how the differences and distances between practice implementations are compared and transformed into scores. Based on these scores, the calibration paragraph describes the search for thresholds and the definition of the size of the fuzzy sets that are used. Finally, the analysis of the second study will finish by illustrating the superset/subset analysis and conducting tests for necessity and sufficiency.

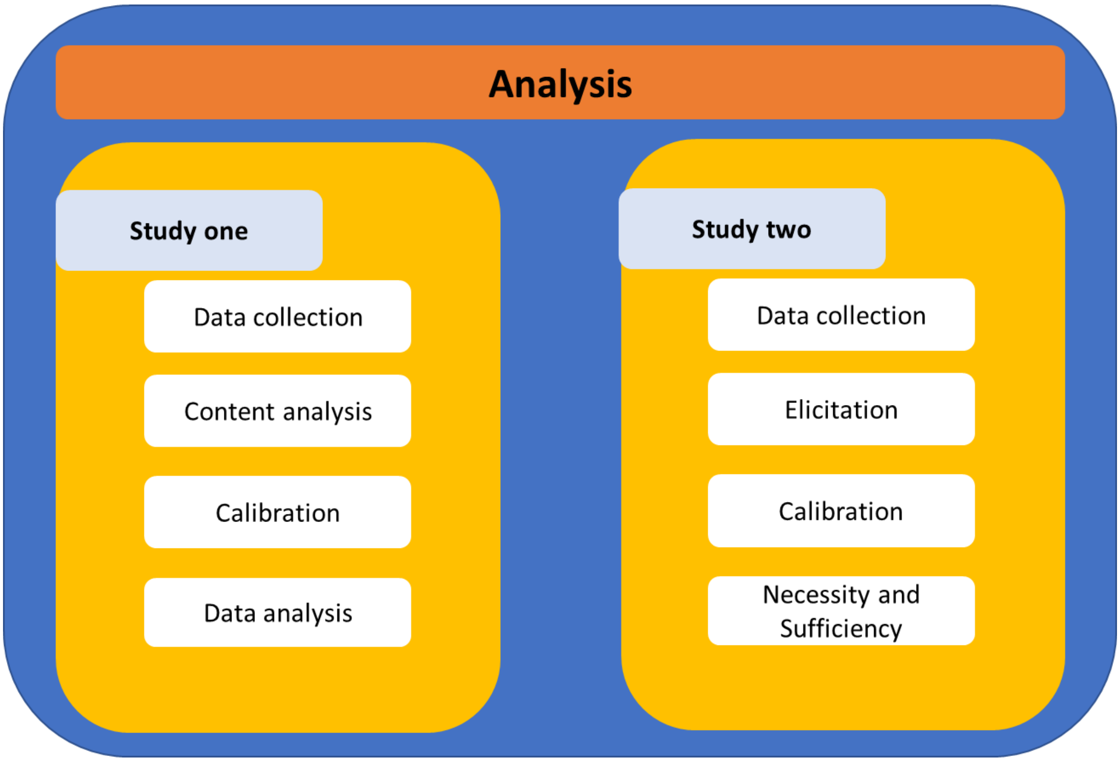

Figure 9: Structure – Analysis

7.2 Foreword

The previous chapter discussed how the heterogeneous nature of healthcare operations and processes, the rarely tested subunit lens, and the degree of causal complexity created by both can impact Lean studies. Furthermore, the still novel character of fsQCA, its sparse use in healthcare, and the unknown contextual factors were also potential barriers for this study. In consequence, the decision was made to split the study into two related studies, with the first operating as a pilot.

(Thabane et al., 2010) described a pilot study as a small-scale study conducted as the beginning of a larger-scale study. Pilot and larger studies share similarities in method and procedures that enable the justification and/or testing of the larger study (Jairath et al., 2000). The pilot thus becomes an analysis. In this case, both the pilot and subsequent studies used the fsQCA methodology. This has previously been utilised in the context of Lean and in the context of healthcare, but rarely in both contexts together. Therefore, it was difficult to predict the implications this might have on the data collection and context-dependent decision-making. According to (Connelly, 2008), a pilot study aims to guide future studies by alerting them to problems that could not have been prevented before the data were collected.

Since these studies used a rarely tested subunit lens to examine performance metrics of unknown complexity in a causally complex and heterogeneous healthcare environment, the empirical path was inherently difficult. Before the start of the empirical stage, this situation raised questions about the exact data sources (e.g., reports and participants), the sufficiency of the selected data for the scale of the method, and the context specifics needed for data transformation (e.g., calibration). Therefore, the pilot study was chosen to test the field so the subsequent study could follow in its footsteps. In general, pilot studies do not aim to test the effectiveness of the study but rather to control its feasibility (Jairath et al., 2000). In simple words, a pilot study investigates and identifies the barriers to a successful study.

7.3 Study 1 - Pilot

7.3.1 Data collection

7.3.1.1 Preparation

Prior to commencing the data collection process, this study needed to clarify which specific operational data were needed. First, a general assessment of the performance impacts of Lean practices required the identification of the practices and key performance indicators (KPIs) involved. In addition, other information was required to assess the status of the practice implementations and their potential impacts on KPIs. Therefore, this study distinguished the information of interest as either a core information set or a supportive information set. The core information set consisted of mainly operational and/or objective secondary data while the supportive information set consisted of a mix of subjective and objective information. The data was performance indicators were distinguished based on

Andrews et al. (2006)'s framework. Significant focus was given to the subjective assessments of the implementations, with the implementation team's leadership and the medical staff who were involved having the highest degree of accessible expertise for assessment.

Thus, the core set indicated the need for a collection approach that could collect structured data. Originally, structured interviews were considered but the idea was abandoned because the number of metrics and practices being assessed would lead to long structured interviews, which can be a burden for the interviewee (McCracken, 1988). Documentation reviews, on the other hand, enable the data of interest to be gathered from existing reports. Therefore, they are specifically suited to gathering background information and large structured data sets (Rohwer et al., 2014, Beelmann, 2006).

The limited sample of interviews implied that other approaches might also be required. The mix of structured and unstructured data, and subjective and objective data indicated the need for multiple data collection approaches. Since the amount of unstructured information required was limited, interviews were a suitable method of collecting it (King and Horrocks, 2010, Kvale, 2008). The nature of the possible answers provided an argument for semi-structured (Newton, 2010) and/or reflective interviews (Alvesson, 2003). Reflective semi-structured interviews were chosen because of the small sample and the topicality or interestingness. In detail, the study aimed to investigate how Lean is implemented in practice and therefore needed the possibility to interact with the participant if he/she provided unexpected insides. An interview protocol (see Appendix A8) was prepared that mainly addressed the implementation of practices. It was likely that the interviews would deliver enough information to address this study's aims, but there was a risk of participant bias because the implementation team were being asked to assess their own work. Therefore, it was decided to add participant observations of implementation workshops to smooth the potential impact of biases.

7.3.1.2 Data Collection

The data set is obtained multiple sources within the subunits/wards of a British hospital trust. Lean was implemented through a KPO team in collaboration with the operations and finance teams. Furthermore, the initiative came from the executive board and received its ongoing support. Data were collected via 2 semi-structured interviews, 7 observations, and 13 documentation reviews. All approaches shared the target of collecting information about the Lean implementation and measured performance indicators in specific subunits/wards. In order to establish comparability between our cases and support the generalisability of the potential findings, the following four selection criteria for subunits were chosen:

1. The subunit is a target of a Lean implementation;
2. A minimum of 3 months has elapsed since the implementation started;
3. No fewer than 10 different full-staff members work in the subunit over one month;
4. Consistent data can be ensured.

These criteria were chosen to guarantee the highest level of stability and validity. Given the aim of the study, only Lean implementations were chosen. The 3-month minimum was set to allow staff to assess how the implementation activities had been distributed during the first few months. After three months, it could be assumed that sufficient Lean practice essentials had been applied for them to be deemed to have had an impact on performance. Since subunits can differ in size and many run on part-time staff, it was necessary to ensure comparability by choosing only subunits with a minimum of 10 full time staff members. The maximum threshold had originally been set at 100 staff members, but it quickly became clear that this threshold was unrealistically high since no subunit exceeded 100 members. Last, all the gathered data were controlled for consistency. Some early-stage subunits used different performance measurement sheets and KPI. For this reason, the most up to date performance formats were used and, in the case of uncertainty, the subunits were excluded.

Of the 67 subunits originally identified, 28 were involved in the Lean implementation. Of these, 21 had been involved for longer than 3 months, 16 of them had more than 10 full time staff members, and 7 such subunits had consistent and complete data. While the sample size and extent of the data set is limited, fsQCA can still be applied. (Rihoux and Ragin, 2008) recommend a minimum sample of 5-7 cases for the application of fsQCA, preferably ~15.

Subunit	Hospital	# Staff	Square meters	Months since implementation start	Ward Task
Ward A	1	51	633	20	Renal
Ward B	1	70	441	20	A&E
Ward C	2	57	553	19	A&E
Ward D	1	60	1197	20	Surgery
Care Unit E	1	NA	101	14	Eye care
Ward F	1	57	585	23	Respiratory
Ward G	2	39	414	23	Respiratory

Table 5: Subunit overview

Table 5 presents the key information for the final sample. It shows the subunit's size (number of staff members + square meters) and the duration of the ongoing Lean implementation. It also describes the activity and speciality of the subunits. While the sample size is comparatively small, it covers all the main care activities of the trust.

Both the documentation review and the interviews provided sound information about the implementations. The interviews offered a general overview of the Lean implementation as well as information about the practical responsibilities of the Lean implementation elements. In addition, they served as an assessment tool for validating the content of the study (Brod et al., 2009). It has been noted that individual interviews not only allow an in-depth exploration of the topic of interest but also assist in validating the expertise of the interviewee (Cooper et al., 2006). For this purpose, the questions for the Kaizen Promotion Officer, who had charge of the Lean implementation, were designed to identify similarities and differences between the theoretical frame and the interviewee's understanding of Lean. The interview with the Chief Information Officer had the same purpose but was targeted at the interviewee's understanding of performance measures. The questionnaires, in addition to collecting data, were also used to validate that Lean approaches/Lean practices had been applied.

The observations supported the interviews by providing information about the practical implementation of the practices. In some cases, more than one theoretical practice had in fact been applied even though the name of the practice indicated otherwise. These distinctions were made based on a case-by-case comparison with practices discussed in the literature review.

While the interviews and observations delivered a general impression about the status of the implementation, as well as an overview of the sets of practices that were associated together, the documentation review was the real foundation for the data collection. RPIW reports provided information about the implemented practices and measured KPIs. These reports were created by the Kaizen Promotion Office. Furthermore, the performance was measured after 30, 60, and 90 days (thereafter quarterly) and these reports show performance changes over time. Moreover, monthly reports gave detailed information about additional supportive practices, which were not included in the RPIW reports since these focused on the workshops. Figure 10 summarises these data types.

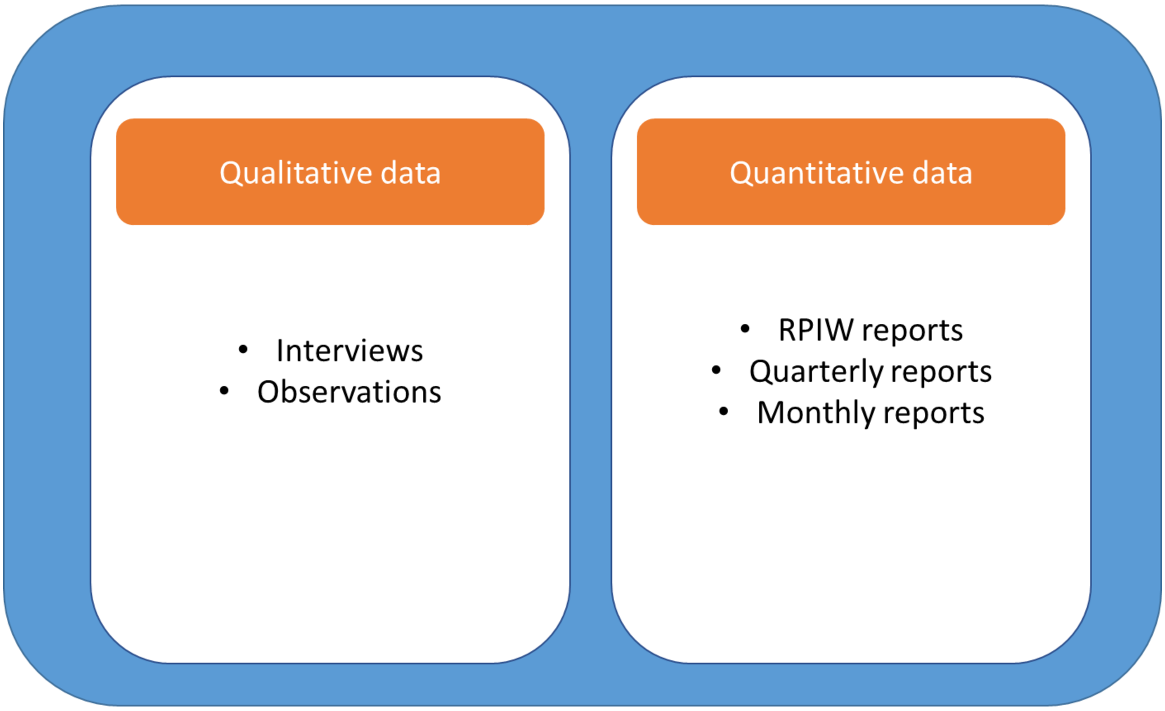

Figure 10: Overview main data types

7.3.2 Content analysis

7.3.2.1 Practice identification

In Operations Management research, interviews are often used in conjunction with content analysis of the frequencies of word usage in transcripts (Flynn et al., 1990). In this study, a summative content analysis approach helped to identify Lean practices in the documentation review. According to (Hsieh and Shannon, 2005), summative content analysis identifies keywords or text content by comparing and counting the word frequencies, as well as by analysing the surrounding context.

It was decided to take a continuous and concomitant approach to the summative content analysis during the documentation review. For example, it was counted how frequently a Lean practice was named in the comment or assessment of the reports. All identified constructs were compared to the practice definition in Lean service literature. The scarcity of existing studies narrowed the choice down to (Hadid et al., 2016)'s practices (see Appendix A2). The coded practices were identified during three different stages. The first phase can be described as the phase of 'direct fit', where practices that shared the same name and characteristics as the theoretical construct were simply copied. The second phase included practices that had similar characteristics but different names. These practices kept their name but received the theoretical description in an extra column. Management or OHRM practices were kept but separately marked. The most difficult cases were those practices that could not directly be identified as any kind of practice in the field of business research. In such cases, context analysis and expert judgement, between the supervisors and the author, provided the framework for the decision-making about their inclusion in the data set. An example would be the 'big huddle', which can be described as a combination of a daily huddle and a spontaneous RPIW. It has a longer time frame than the ideal type of the daily huddle but a lower degree of focus than a Kaizen event. Therefore, the name was retained and both framing theoretical practices were noted. However, none of the third phase practices were included in the data set as they occurred only once in the implementations and without a second or third practice it's difficult to validate a pattern for one practice.

7.3.2.2 Degree of practice implementation

The study has orientated itself on (Galeazzo and Furlan, 2018)'s assessment of the degree of Lean bundle implementations. Therefore, the lingual assessments of the practices in reports and interviews was used to determine an initial placement. However, assessing the level to which a Lean bundle (or in the context of this study, a Lean practice) is implemented is context dependent. Several scholars have investigated the assessment of Lean implementations as well as the status of organisational leanness. The scientific discussion has developed in several directions. First, there is a distinction between Lean in manufacturing industries (e.g. Saurin et al., 2011) and service industries (e.g. Laureani et al., 2010, Liker and Morgan, 2006). Also, Lean assessment studies have taken the form of a wide variety of theoretical and empirical work. Therefore, the methodological distinction between qualitative assessment procedures (Fullerton and Wempe, 2009, Doolen and Hacker, 2005, Panizzolo, 1998) and quantitative approaches (Behrouzi and Wong, 2011, Bayou and De Korvin, 2008) provides a strong framework for the assessment of Lean implementations. Quantitative approaches allow for a comparison across large numbers of practices while qualitative work can explain their individual impacts. However, while the implementation might be one of the most discussed topics in the Lean area, the discussion has been chiefly focussed on full implementations in an organisation-wide setting (e.g. Shah and Ward, 2003). Some scholars have chosen to focus on a more in-depth analysis (e.g. Fullerton et al., 2003) that might allow practices to be identified, but the complexity of the data and, indeed, the difficulty of accessing data on so deep a level, provides a significant challenge for most studies.

In the case of this pilot study, little of the previously discussed work was appropriate for forming the basis of assessment of the Lean practice implementations. Because existing studies have focused on the overall Lean implementation, they have generated only limited insights into the assessment of Lean practice implementations. Hence, this study's assessment was based on Pakdil and Leonard (2014)'s model of 8 dimensions and 62 assessment criteria. It therefore followed a structured and standardized procedure for assessing complex implementations by viewing the implementation from the perspective of time effectiveness, quality, process, and cost and compared with each other.

The practical assessment of the practice implementations used three multi-metrics, which were averaged to generate an overall implementation index per case. A 'case' in this pilot study was defined as an individual scenario of one practice implemented (or not) in one subunit. The three metrics consisted of one objective measure (time since start of implementation), one subjective measure (assumed depth of implementation), and one combined measure (implementation consistency). The objective measure is simply the amount of time that has passed since the implementation started. This does not necessarily mean that the practice was being implemented during the entire period but rather, since continuous improvement is one of Lean's fundamental assumptions, the secondary assumption could be made that the practices were still being applied over time. The subjective measure used the previously discussed dimensions of (Pakdil and Leonard, 2014) as the guide to assessing the depth of the implementation. However, since this approach was designed to access organisation-wide implementations, the limited pilot data prevented the full usage or adoption of the dimensions. In addition, few of the criteria were performance indicators, with most being practices. Consequently, the dimensions were used where the collected data allowed and were adapted accordingly. For example, while many staff members used Takt time reduction and several reports indicate a high usage of the practice, it was only used in basic organisational settings and operational tasks; these never addressed subunit-wide processes. Thereby, the lack of subunit-wide coverage resulted in low assumed depth. Finally, the combined measure of implementation consistency is an averaged score of (i) the ratio of subunits where the practice was applied compared to the total sample of subunits, and (ii) the expert-judged impact/quality of its implementation. The expert judgement was gleaned from the interview data where this was available, and from our own assessment in cases where it was not. For example, in a case where the interviewee described the leadership support as 'very strong', and this was supported by reports and observations, leadership was assigned an assumed depth score of one.

The collected data did not allow for the assessment of the assumed depth in every implementation case. However, time values were available for every subunit. Therefore, the overall depth metric for each practice was used and combined with the average time value for the relevant subunit. The results of the assessment are illustrated in Table 6 below. Values from 0 to 1.0 were used to evaluate depth and consistency. Both measures are partially objective and partially subjective, as they used both the subjective assessment of interviewee/researcher and the implementation duration. The 'Overall' column represents the average of both values.

Practice	*Assumed depth of implementation*	*Implementation consistency*	*Overall*
Process Redesign	0.5	0.3	**0.4**
Takt time reduction	0.3	0.7	**0.5**
Leadership / Obtaining Management Support	1	0.6	**0.8**
Process flow analysis	0.6	0.8	**0.7**
Just in Time/Quick setup reduction	0.8	0.6	**0.7**
Mistake proofing	0.7	0.6	**0.65**
Staff empowerment	0.9	0.4	**0.65**
5S	0.7	0.7	**0.7**

Table 6: Overview degree of the overall practice implementation

All measures were transformed into percentage scores to better calculate (i.e., average) the overall implementation scores.

All subunits have scores of between 0.8 and 1 in the objective criteria (time). The assumed depths of the practices range from 0.5 to 1, with only Takt time reduction scoring a value of 0.3. Management support achieved a perfect score of 1 because of the high presence of all involved managers, in terms of both the number of managers present and number of stages in which they were involved. The table shows a consistent score set with a limited variance; this will be transformed into a fuzzy set to calculate the causal condition of Lean practices.

7.3.3 Data cleaning

The data cleaning was conducted before the previously discussed evaluation of practice scores. However, it is discussed here as it links closely to the performance measures of interest.

Some subunit cases had to be excluded (e.g., eye care and recruitment) because of the lack of structural differences and/or lack of practices (e.g. huge differences in subunit size). Most of the eye-care wards did not fulfil the minimum inclusion criteria, which were defined as three practices applied and one minimum operational performance metric measured. The cut-off criteria were chosen in preparation for a feasible sample size of a QCA application and under consideration of overlapping data sources.

In the case of the recruitment subunit, the structural difference was simply too big in comparison to the other units, most particularly with regard to subunit size (square metres and number of staff) and process design. As regards performance measures, 5S was excluded because the collected data did not provide a sufficient explanation. Nevertheless, 5S remained in the practice set for comparison.

Originally, the content analysis found 42 KPIs. However, the measured KPIs addressed different levels of organisational structure and the metrics ranged from medical KPIs to financial performance measures. Some metrics were measured at subunit level, while others were measured at departmental level, or value stream level, or hospital level, and some were even organisation-wide. The purpose of the study was to identify the impact of combinations of practices that led to improved performance. Therefore, this required performance metrics measured at the subunit level. While the other levels (value-stream, organisation wide) were compared and tested for minimum sample size and data coherence, they did not provide a consistent data set or large enough sample size. For example, only one interview and one quarterly report indicated organisation-wide practice implementations, but no specific performance data was provided and it remained unclear what was the exact frame of organisation-wide.

After the first cleaning, 16 KPIs were left, which were finally narrowed down to 5 KPIs (data existing in a minimum of 3 cases) and 4 KPIs (minimum of 4 cases). Both categories are illustrated in Table 7 below. The first column presents all KPIs with only 3 cases and the second column presents KPIs with a minimum of four cases. The exact sample sizes are given in brackets.

Min 3 cases	Min 4 cases
Lead time	Lead time (7-8)
Quality defects	Quality defects (6-7)
5S Stages	5S Stages (6-7)
Set-up reduction	Set-up reduction (4)
Walking travel distance	

Table 7: Overview of included KPIs

The decision was made to exclude the KPI 5S Stages because it was a subjective qualitative measure without clear definition. In the light of combined subjective and objective metrics (Andrews et al., 2006), the 5S performance metric might become a subject of interest but would require additional information. Consideration of this was deferred to the second study.

- Lead time as well as quality defects are aggregated/averaged scores of processes that were not fully medical. As a result, the calculations and the processes have been checked for data consistency. While the use of averages has limitations for different sized processes, in this case the purpose and size were similar. For example, two practices involving Lead time were the board round and medical progress information gathering.

- Quality was measured as the number of operational (not medical) quality defects and transformed into quality improvement (%) due to comparison reasons and the positive bias of the data set (all variables highly positive – no deterioration).
- Lead time and the lead time delta were measured in minutes and, like the quality metric, they were transformed into lead time improvement (%).

7.3.4 Calibration

7.3.4.1 Theoretical concept of calibration

FsQCA distinguishes itself from other QCA approaches through the expression of degrees of set membership. In practice, these degrees are determined through calibration, which transforms attributes and outcomes to membership degrees of their ideal types.

In this study, the process of calibration started with a theoretically driven desire for a high level of granularity. The main factors for a calibration are context driven and, in this case, the major influences were the representation of the data spread, the accuracy of the data presentation, and the large improvement scores as representative of content consistency.

7.3.4.2 Lead time calibration

For the first round of calibration, a five-value fuzzy set was tested to get a fine-grained impression of the sensitivity of the data set. Neither a continuous nor a 7-value set provided an adequate fit for the small sample. However, using a five-value approach for seven cases was not optimal. In addition, the full membership threshold of 0.9 was above the highest achievement (88%), while the non-membership threshold was below the smallest improvement (19%). The next test used a four-value fuzzy set with new borders; this provided a more representative result that accounted for the maximum and minimum spread of improvement. However, the 'mostly in' and 'cross-over' thresholds did not accurately represent the distribution of cases. The reduction of both led to a finer representation of the reality. In the last two calibrations, the 'mostly in' threshold was abandoned because no score fell between this and the full-membership threshold. In addition, finer adjustments were made to the maximum and minimum values of the fuzzy set (fs) that respected the difference between the data interval boards. Therefore, three additional values above 0.8 had full membership. From a context perspective, 80% improvement should be considered as superior improvement.

Lead time	Full	Mostly Full	Cross-over	Mostly Out	Out	Decision	Reason
Round 1	0.9	0.75	0.5	0.25	0	Reject	No values above 0.88 & is 0.19 mostly out or out?
Round 2	0.85	0.75	0.5		0.2	Reject	Ward F (0.49) could be in. 0.75 doesn't represent the spread between the top and bottom values.
Round 3	0.85	0.65	0.45		0.2	Reject	Ward F is a difficult decision. No values around 0.65. Small sample + half the sample close > 4 value fs not perfect
Round 4	0.85		0.5		0.25	Reject	4 cases are close together. 3 are at 0.82. Difference to top is 6%. Difference to bottom 63%.
Round 5	0.8		0.5		0.25	Accept	

Table 8: Lead time calibration

Table 8 illustrates the calibration process for lead time. The first column lists the number of the particular round of calculations. The following five columns represent the degree of membership. ‘Full’ represents full set membership, ‘mostly full’ represents the mainly full set membership, etc. The penultimate column is the decision column, and it shows the general decision result, and the final column provides brief reasons for the decision.

The values can be understood thus:

0.8 – full membership of superior improvement > strong improvement

0.5 – cross over point > moderate improvement

0.25 – no membership of superior performance > fair improvement

Following the recommendation of numerous scholars, we added a constant of 0.001 to all values to avoid interpretation difficulties (Galeazzo and Furlan, 2018, Reimann et al., 2017, Santos et al., 2017, Greckhamer, 2016, Fiss, 2011, Greckhamer, 2011, Fiss, 2009, Ragin, 2008).

7.3.4.3 Quality improvement calibration

The calibration of the quality improvement measures had benefited from the experience gathered from the lead time calibration. Therefore, the slightly smaller sample led to the decision to start with a three-value fuzzy set. However, the limited sample size and the data spread provided difficulties for the fsQCA software calculations, such that a manual calibration through recoding was required. A four-value fuzzy set was also tested. While half the sample achieved full set membership, the other half was equally distributed. In the last calibration round, minor adjustments to the ‘mostly in’ and ‘fully out’ thresholds were conducted. Accordingly, a finer representation of the high performance of the lowest score was possible, as well as the full and mostly full members.

The small sample size led to the decision to use a gross calibration. In particular, a three-value fuzzy set was used for lead time and a four-value fuzzy set for quality. Another limitation is the exclusively positive data set, which in several cases showed

80% and more improvement. Therefore, the calibration was used to show the degree of membership of superior performance improvement.

For Table 9, which shows the overview of the quality improvement decision, the same layout was chosen as for the lead timetable. The quality improvement calibration is illustrated below:

Quality	Full	Mostly Full	Cross-over	Mostly Out	Out	Decision	Reason
Round 1	0.9		0.7		0.5	Reject	Sample to small. Spread to wide. Software has problems. Maybe recoding works better.
Round 2	0.91 - 1	0.81 - 0.9		0.71 - 0.8	0.61 - 0.7		Ward F (0.9) is difficult decision again. 10% up to max full. 18% down to mostly out. Maybe recalibrate mostly - values.
Recoded to	1	0.75		0.5	0.25	Reject	
Round 3	0.91 - 1	0.81 - 0.9		0.71 - 0.8	0.61 - 0.7		
Recoded to	1	0.85		0.5	0.35	Accept	

Table 9: Quality calibration

In the case of the KPI quality, the values were re-coded thus:

0.91-1 became 1 – representing full membership of superior improvement

0.81-0.9 became 0.85 – representing mostly in but not completely

0.71-0.8 became 0.5 – representing mostly out but not completely

0.61-0.7 became 0.35 – representing full non-membership

7.3.4.4 Practice Implementation Calibration

The sample representing the degree of practice implementation had 55 variables. It was therefore far bigger than either of the two other samples. The difference in sample size occurred because each level of implementation of each practice in each subunit was assessed. Based on the experience of calibrating lead time and quality, we decided to start with a five-value fuzzy set for the representation of the sample, which can be described as medium to large sized (Vis, 2012). The data limits of 0.32 and 0.77, as well as the 0.95-interval borders of 0.4 and 0.67, drove the decision-making for the first fuzzy set thresholds. However, a non-membership assessment of 0.55 for a 0.32 or 0.38 implementation assessment score seemed too high. Therefore, the out threshold was reduced to 0.35 in the second round. However, the use of a Pearson coefficient for Round B's fuzzy set still indicated a strongly negative skew. To establish a finer grained representation of reality, the decision was made to use a seven-value fuzzy set in Round C. Although different values were tested for slightly in and out in Rounds D and E, fuzzy set C remained the best option because it represented a small sample of full membership (0.7-0.79) and no non-membership scores. Table 10 below illustrates the calibration rounds for the practice implementations.

Practices	Full	Mostly Full	Slightly full	Cross-over	Slightly out	Mostly Out	Out	Decision	Reason
Round A	0.79	0.73		0.67		0.61	0.55	Reject	While only two cases were below 0.4 (0.32 and 0.38) they still seemed too high to be considered out.
Round B	0.79	0.72		0.65		0.55	0.35	Reject	Five value fs seemed fine enough but the mostly out and out values were still not looking like a solid fit.
Round C	0.71	0.61	0.51	0.41	0.31	0.21	0-0.2	Accept	
Round D	0.71	0.61	0.51	0.41	0.35	0.25	0-0.2	Reject	0.35 (slightly out) would have pushed the 0.32 value in the mostly out degree section.
Round E	0.79	0.65	0.51	0.41	0.31	0.21	0-0.2	Reject	Test to see the data sensitivity. High number (0.67) of cases didn't show changes. 0.79 a bit too high for full.

Table 10: Practice implementation calibration

7.3.5 Data Analysis

This study uses Qualitative Comparative Analysis. QCA recognizes patterns of set relations in a population of interest (Schlosser et al., 2009). In this pilot study, the population of interest are hospital subunits that have applied certain Lean practices (causal conditions), which exert specific impacts on performance measures (outcomes). For further assessment of the degree of membership, fuzzy set theory was used for the calibration. During the practical analysis, fsQCA Software (version 3.0) was used in the first stage. Later, more complex calibration approaches and sensitivity analysis were conducted through R and the software package QCA (version 3.4).

The fsQCA software tests for necessary and sufficient causal conditions or for combinations of them. While fsQCA is mainly designed to identify sufficient conditions, it can also show the existence of necessary conditions (Dul, 2016a). First, this study analysed which Lean practices might be necessary for the achievement of superior performance improvement. According to (Ragin, 2000), the outcome of interest is a subset of the causal condition. Therefore, the necessity of causal conditions can be tested by analysing if any causal condition is always present in the outcome of interest.

For this purpose, (Schneider and Wagemann, 2012) recommend testing the consistency scores of the recipes of interest to identify the statistically possible configurations of causal conditions. Each recipe is necessary for the achievement of the outcome of interest when it exceeds a consistency score of 0.9. This contrasts with the test for sufficiency, which requires a consistency score minimum of 0.8 (Ragin, 2008). Traditionally, the empirical evidence for sufficiency tests is not completely in line with set theoretical assumptions of consistency and coverage scores of 1 (Schneider and Rohlfing, 2016). A sufficiency test examines if a causal condition or conjunctural causation is enough to achieve by itself an outcome of interest. Thereby, the outcome can be understood as a superset of the combination of causal conditions. In practical terms and driven by the causal argument, the study tries to analyse which self-standing Lean practices are able to stimulate a superior improvement in performance.

In the case of a sufficiency test, QCA recommends the use of a truth table. The supporting algorithm (here, Quine and McCluskey) highlights all possible recipes of the chosen causal conditions (Fiss, 2011). An application of the test first requires the setting of values for consistency and case frequency. Both can be predefined through previous research. A raw consistency of 0.8 was defined as sufficient to show if a possible recipe or combination of Lean practices is connected to superior improvement. (Ragin, 2008) notes that the frequency is dependent on context and sample size. Under the condition of a medium sample size, 2-3 cases are recommended. However, the definition of a medium sample is somewhat vague. For example, (Fiss, 2011) and (Ren et al., 2016) used 3 cases as the solution frequency with a total sample of 56/160 cases. On the other hand, (Oyemomi et al., 2016) worked with a frequency score of 1 case, investigating a sample consisting of 28 cases, while (Cooper and Glaesser, 2016) used a frequency threshold of 12 on a sample of 6666 cases. Although the frequency is claimed to be driven by context and sample size, it was actually sample size and robustness that played the bigger role in most of the previous studies. In this study, sample and context were considered for the frequency determination. The small sample would support a low threshold. However, the inconsistency and incompleteness of the data, as well as minor differences between the cases, indicated that higher robustness would be beneficial to the study. Therefore, a frequency threshold of 2 cases was chosen.

Overall, truth table analysis enables the usage of all three solutions: parsimonious, intermediate, and complex. (Ragin, 2008) recommends usage of the intermediate solution as it includes only logical remainders that make sense according to their context. This study followed Ragin's recommendation.

7.4 Lessons from the Pilot

During the pilot study, major insights were gleaned regarding the data and collection procedure. First, it was noted that the Kaizen Promotion Offices (KPOs) seemed to hold an information monopoly about performance data and the implementation assessments. While this perspective was still not fully confirmed, as it did not touch related but not included subunits and staff, it allowed a clearer lens on possible participants and data sources. Second, it became apparent that wide-spread standardized performance reports were indeed a good foundation for a subunit lens. Third, the pilot emphasized that the KPI reports on their own did not comprise a sufficiently adequate documentation review for the context evaluation (calibration). Therefore, supporting (and most likely qualitative) data were required. The pilot also enabled the identification of an advanced core group of potential participants for interviews.

Overall, the pilot provided insights into how the suitability of the empirical setting for the study can be further improved. In recent years healthcare systems worldwide have been put under severe pressure. The NHS is one of the few fully nationally controlled systems with shared knowledge, practices, tools, procedures, leadership and HR. Through the pilot it became apparent that the study would need to make a trade-off decision between the need for access to very deep insights into the organisations and a variety of sources. Through the extension of the sample and the adjusted main study this should link the empirical setting to the overall study.

7.5 Study 2 – Main

7.5.1 Data collection

While a large number of British hospitals (NHS) have dabbled with Lean implementations, only a few have launched lasting organisational-wide production systems (Burgess, 2012). For this research, 7 care facilities were selected as representing Lean implementers in British healthcare. Three of these seven facilities were part of the first study. Through the updated dataset additional Lean implementing wards could be included for these three facilities.

However, as previously discussed, Lean practices are mostly implemented on a subunit level, with a few implementations being on a value-stream level that stretches across subunits focusing on one treatment path. Therefore, the decision was made to choose subunits as cases. Of the possible 183 wards/subunits, 56 subunits were found to have a substantial Lean implementation; these included the 41 subunits included in the sample. In the selection process the following four criteria were considered:

- The subunit has a minimum of two Lean practices.
- A minimum of 9 staff members spends no less than 50% of their work time in the subunit of interest.
- The implementation has been ongoing for at least 3 months.
- Performance data is consistent and available

While 41 cases were found, the exact number varied by the availability of data for the performance indicator under investigation. Unlike typical case study research, fsQCA's cross-case comparison is not limited to 5-7 cases (Rihoux and Ragin, 2008) with an effective sample size of within 12 to 50 cases (Ragin, 2008). FsQCA can be applied outside these boundaries and where there is, for example, a smaller sample, strong data intensity across cases can improve robustness that would otherwise be lacking. Alternatively, QCA can be combined with regression analysis to increase the case sample at the cost of a more limited sample for causal conditions (Fiss et al., 2013b). However, this study's sample is within the sample size suggested to be effective for fsQCA.

Originally, it was conceptualized that there would be a split between the core and the supportive data set, based on the differences between the subjective and objective data. However, the small-scale pilot study revealed that this division had even more importance for the validity of the data. Different hospitals used different Lean coaches, which caused staff to have different understandings of Lean. In one case, the staff team was trained by different coaches and as a result, this ward was excluded from the data set after testing. For these reasons, more attention was given to gathering data for the supportive data set in order to check and test the validity and understanding of the objective data.

The core data set consists of around 260 KPI reports. However, only 52 of them were used, as many reports included the previous performance results at a 30-day frequency for the first 6 months and quarterly thereafter. Reviewing these reports helped to identify the practices used and their related performance changes. In addition, 13 quarterly reports from the Kaizen Promotion Offices (KPO) to the executive boards were collected. Both report types were the object of documentation review.

Furthermore, the core data set was supported through reflective semi-structured interviews. Five interviews were conducted, of which only 2 were of significant relevance. The decision to not extend the sample further was practice-driven as, during the interviews and accompanying discussions, it became quickly apparent that there was a large imbalance in knowledge about the Lean implementations and their performance assessments. Although several staff members were extensively involved in the implementation, the KPO Leads kept an information monopoly. The majority of KPIs were introduced, measured, and assessed by the KPO teams. All other departments received the reports from the KPO team such that their feedback to the questions was consistent with the KPO Leads' answers, albeit less detailed. While the differences in power and personal experience could have provided a valuable contribution, the value for this study was considered not strong enough to justify the withdrawal active medical personnel from their service. Furthermore, the aim of the interviews was to identify the practical understanding of Lean in general and in the implementations. This aim was achieved with the two 4-hour interviews.

Additionally, 6 observations of Lean workshops, Kaizen, and RPIW events were conducted for validation purposes. Again, the sample was not further extended as the findings were consistent with findings from the other data sources and no indications of deviation were found.

Last, the supportive data set included a further 4 annual reports concerning the implementation. These reports provided detailed insights into the indirect assessments by staff members involved in the implementations. Specifically, the reviews provided an overview about supportive practices and analytical results of performance changes over the years. While the KPI reports included similar information, they were less detailed and illuminating.

7.5.2 Analysis

7.5.2.1 Practice identification

The standardized format of the analysed KPI reports simplified understanding of their content; there was a difference in understanding about the Lean practices across the healthcare facilities and/or subunits that required a structured approach to ensure content validity.

Content analysis was applied, which has '…the pretension to be intersubjectively comprehensible…' (Mayring, 2004, p. 161), thereby enabling reliability.

More specifically, summative content analysis was chosen to accompany the documentation review. This approach is based on the search for key words and phrases, tallying and comparing their occurrences, as well as the analysis of the context (Hsieh and Shannon, 2005). The occurrences of possible Lean implementation constructs were noted, counted, and compared with theoretical definitions where available. We controlled for cases where no definition was available by identifying a similar practice in one of the studies in the literature review. The intention was to assess whether the practice could be categorised as 'un-identified' if no fit could be found but, in the event, no such constructs were found. All identified practice constructs had previously been classified as Lean practices. In line with the literature review, the following practices and KPIs were identified in the data set:

Practices	KPIs
Effective Communication System	Leadtime
Change facility layout	Quality (Mistakes)
Kaizen / huddle	Setup time
Takt time	Walking travel distance
Total quality	Staff engagement
Multifunctional employee	Staff satisfaction
Customer involvement	
Standard work in progress	
Process redesign	
Staff empowerment	
Leadership	
Quick setup time	
Mistake proofing	
5S	
Simplification	
Process flow analysis	
Standard practice	
RPIW / Kaizen blitz	

Table 11: Practices and KPIs of study two

7.5.2.2 Degree of practice implementation

While the identified KPIs and Lean practices form the foundation of the data set, they require some transformation. Overall, one of the strengths of the fuzzy set method is that it can express different levels of set membership (Ragin, 2008). Nevertheless, the degree of leanness of a practice is highly dependent on its context.

This is particularly evident in the wide-ranging subfield distinctions seen in the scientific discussion surrounding the assessment and analysis of leanness. These are particularly diverse in qualitative analyses of differentiations and interlinks in the socio-technical stream of Lean literature (Fullerton and Wempe, 2009, Doolen and Hacker, 2005, Panizzolo, 1998) and complex quantitative discussions (Behrouzi and Wong, 2011, Bayou and De Korvin, 2008).

Similarly driven is the discussion about the relevance of the industrial environment to the assessment of an implementation, with a broad range of studies focusing on practice dependencies in a manufacturing (Saurin et al., 2011) or service setting (Laureani et al., 2010). Nevertheless, few studies have investigated Lean at an organisational subunit level (e.g. Fullerton et al., 2003), mainly because the accessibility and reliability of data can be difficult to obtain at this level.

It is therefore unsurprising that the socio-technical nature of the data, the setting dependence of the study, and the data reliability of Lean practices proved to be challenging for this study. However, these barriers were overcome by ensuring a diversity of data sources and combining expert judgement and elicitation based on subjective and objective scores.

7.5.3 Elicitation

7.5.3.1 Preference elicitation

After initial runs of the analysis, the decision was made to strengthen the conception of the practice implementation scores. Originally, these were created through expert judgements based on a combination of objective and subjective assessments. Some subjective biases in the expert judgements were particularly evident during the analysis. For example an over-prediction bias or under-prediction bias (Koehler et al., 2002) led to difficulties in assessing the relationships and distances between scores. Every time the decision makers judged different practice implementation scores, a preference decision occurred. For example, one of the interviewees strongly indicated that RPIW events were considered to be the most impactful practice. Effective communication systems were considered to have least impact, and so the distance between them was easy to assess. However, comparing takt time with staff engagement proved to be more difficult. Staff engagement was considered as impactful by the interviewees, but the secondary data indicated that takt time was involved in far more events. As a result, we were forced to conduct a preference judgement to weight these factors, as in this case, over- or under-predicting one of the factors could generate wrongly classified implementation scores. While these biases cannot be completely excluded, there are various decision-making tools that support a more structured and controlled judgement. In this research, preference elicitation was chosen as the decision-making tool because it allowed inclinations to be placed on a specific scale, decreasing the impact of subjective biases on the score distances.

According to (Simpson, 1998) preference values can be considered as fundamental elements of normative decision-making models. Nevertheless, it is doubtful that all values can be well defined. (March, 2006) argued in support of excavating pre-existing values whereas (Keeney, 1996)'s findings indicate that values have to be 'discovered'. While it is apparent that decision makers face fewer problems when dealing with intimate or personal preferences, in most decision-making processes, practitioners must assess the corporate environment. Corporate values are less personal and therefore less likely to be fully-formed. However, in other cases, the decision maker might have a personal relation to the preference but still cannot make a clear decision, such as in the case of a doctor faced with contradictory symptoms such that the treatment options cannot easily be identified. Furthermore, preferences can change over time and must therefore be considered to be dynamic. Here, the requirement for feedback becomes paramount and decision makers need to see the impact of their preferences on the problem of interest. The literature recommends a few tools for such a case. These include visual analogue scales (Simpson, 1998), standard gambles (Dupré, 1998), and proxy or natural scales (French et al., 2009). Given that the first two tools cannot accommodate the easing of preferences over time, this limits the efficacy of these tools, even though the easing of preferences is not a major condition of this study's data set.

As a result, for the purpose of determining the practice implementation scores, a natural scale with local scaling was chosen, with the scores representing the degree of implementation in percentages on a scale from 0 to 100 that occurred naturally without having to be manually set.

7.5.3.2 Practical elicitation

In this case, the determination of the subunit practice implementation scores was limited by the setting and the data frame. The objects of interest were the Lean practices and the subunits in which they were applied, and both required assessments. While the data set allowed conclusions to be drawn about implementations of the Lean production system in the subunits, it did not provide enough information in relation to the Lean practice implementations in the subunits. For example, the data set provided information about the duration of each Lean implementation by subunit and practice, but it did not provide this information for each Lean practice in each subunit. Therefore, it was decided to elicit both dimensions separately and combine their results.

In both cases, natural scales were used and, since the number of subunits exceeded 40, subclasses were used. These subclasses represented low, medium, and high values. For the first iteration, an anchor value was chosen as the starting point. However, in additional iterations, values were replaced, meaning that the anchor's binding became weaker with each iteration. Thus, the values became strongly linked to each other. In the subunit data dimension, there were a large number of values grouped in the high value area.

For the elicitation process, a local scale was set. Normally, local scaling extends the scale such that the initial borders are exceeded (French et al., 2009); however, this did not occur in this case.

Overall, the elicitation was based on different information strings, with one of them being a separate linguistic evaluation of the interviews and the other relating to the secondary data (i.e., the reports).

7.5.3.3 Linguistic evaluation

Following the recommendation of (Narayanamurthy and Gurumurthy, 2018) for the linguistic evaluation of Lean implementations, the following key words and descriptions were searched for and allocated to the relevant subunits/practices:

- Worst/Very low
- Very poor/Low
- Poor/Fairly low
- Fair/Medium/Average
- Good/Fairly high
- Very good/High
- Excellent/Very high

The evaluation was kept basic because going into more detail did not seem suitable to the purpose of the elicitation and the strength of the linguistic feedback. Other data sources were considered in a similar fashion.

For example, the practice RPIW was described as 'very prominent' and a multifunctional employee as 'average practice'. The length of text coverage in the reports supports this distinction. For these reasons, RPIW was placed on top of the list and multifunctional employee in the middle. Finer grained adjustments were done in iterations that considered and had regard to the other practices.

7.5.3.4 Elicitation of practices

According to (Ribeiro, 1996), the evaluation of preferences can be achieved through the investigation of attributes and their objects. In general, attributes can be defined as criteria for the assessment of decision options. There are two pathways through the elicitation process: intra-attribute preferences and inter-attribute preferences. The first path represents the value of each object to the attribute. In a Lean implementation, these might include budget and personnel. Both the weight and the value of an object represent a trade-off decision between the various objects of an attribute. For example, when a patient is faced with choosing a surgeon, does he/she choose the 0.2% error rate or the 7% error rate?

Inter-attribute preferences, on the other hand, represent the weight given to each individual attribute, which allows the least preferred feature to be distinguished from the most desirable. For instance, which attribute has a stronger impact on the strength of the Lean implementation: financial budgeting restraints or management training?

For this elicitation, the length of the implementation of each practice, the number and seniority of the staff members involved, the mentions in monthly executive reports, and the number of revisions were considered as attributes for decision-making. While there might be an overlap between some of the sources for this data set, there is no other information that would lead to contradictory assumptions.

The elicitation classified the practices in groups to generate a general overview. For example, process flow analysis was given high attention by the interviewees and it was frequently mentioned in reports and performance overviews, whereas change facility did not. Nevertheless, change facility was still specifically mentioned in one subunit as an essential element, which pulled it out of the 'very low' class into the 'low' class. Standard practice was placed in the 'high' class as it occurred far less often than the indicators for RPIW events. As a control element, individual practices were moved around the groups to see if they might belong to another class. If the preference decision was difficult, the practice was placed close to the border of the class.

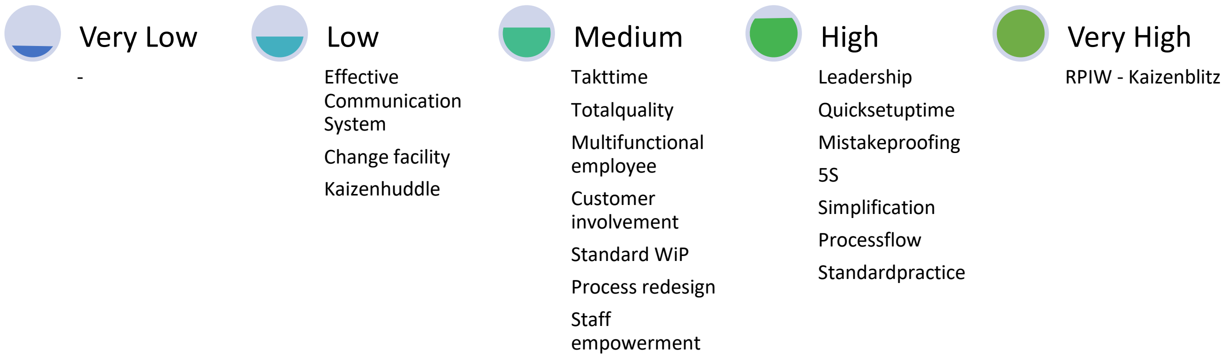

Figure 11: Elicitation practices

In the next stage, the practices were compared in their classes. As this comparison required more detail, the absence of evidence about the practice was also considered to be a factor. For example, takt time received a high frequency in performance overviews but a low frequency in monthly reports and was mentioned only once during the interviews. As a result, consideration was given to placing it in the low class, but its high frequency in performance overviews indicated regular usage, even if this was something of which the interviewees and reporting officers were potentially unaware. In the case of standard practices, the absence of performance measures in most overviews simply indicated that no performance directly connected with this practice was measured. Nevertheless, the very format of the performance sheet indicated that standard practice was applied. To control this assumption, performance overviews dating back to before the Lean implementation were requested and checked. These were either in another format or did not exist, indicating that no operational performance measurement in this subunit had been conducted prior to the Lean implementation. Hence, standard practice was placed ahead of process flow analysis. In Figure 11, the upper practices in each class represent the lowest position within the class. While there were individual practices that would have fallen into the very low class, they were excluded in earlier stages because of their limited sample or for other reasons. None of the excluded practices would have exceeded the medium class.

7.5.3.5 Elicitation of Subunits

In general, the elicitation of subunits followed the procedure outlined above for the elicitation of practices. Hence, the duration of implementation, the number of subunit staff members involved, and the mentions in monthly reports were considered. However, revisions could not be used because only practices were revised. Instead the frequency of reporting in each subunit, as well as the reporting between the subunit and the KPO office was considered. In addition, observations and the interviews provided more detailed feedback and assessments about the Lean implementations in the sub-units.

For example the subunits B1, B2 and A6 are wards in or close to the emergency care departments with similar tasks. While B1 and B2 are from a much smaller hospital than A6, the actual ward size differs. B1 is the smallest followed by B2 and A6. The last one has double the size and staff than B1. However, B2 applied Lean longer than A6 and has therefore more staff members involved in the implementation than B2. As result A6 and B2 were classified as high and B1 as Medium with A6 being one step higher than B2 because of the significantly larger number of involved staff members.

Below, the results of the subunit elicitation can be seen:

Figure 12: Elicitation subunits

The results of practice and subunit elicitations were numerated and combined to express the degree of implementation of one specific practice in one subunit.

7.5.4 Calibration

7.5.4.1 Procedure

In set theory, fsQCA distinguishes itself from other QCA approaches through its consideration of the degree of set membership (Ragin, 2008). Calibration can be defined as the method whereby outcomes and constituting attributes are transformed into scores that express the degree of set membership. Thus, the transformation process uses predefined ideal typologies. Calibration in practice uses not just theory but also in-depth case knowledge to pre-define these types (Rihoux and Ragin, 2008).

Two calibration approaches have distinguished themselves in the theory. According to Ragin (2008), the direct method, in its traditional form, uses three qualitative anchors that express ‘full membership’, ‘crossover point’, and ‘fully out’. The second approach is the indirect approach, which uses six anchors to additionally consider ‘mostly in’ and ‘mostly out’ cases.

In practice, calibrations are highly context dependent and become trade-off exercises between high levels of case consistency and high amounts of result detail. This commonly leads to usage of the ‘continuous’, three-, four-, or six-value fuzzy sets (Ragin, 2005).

For the calibration of this case's data and subsequent analysis, the QCA package in R was used and the results were validated through the fsQCA software, versions 3.0 and 3.1. The aim of the calibration was to achieve a high level of result fineness for the causal condition, and high stability for the outcomes. The degrees of leanness of each practice per subunit were defined by the context-dependent causal conditions. As previously discussed, the decision was made give more emphasis to context through the use of elicitation. It enabled the author to revise the impact of the duration of the implementation, staff members involved in the implementation etc. methodically. To further improve the contextual link, the structured approach was chosen. The calibration of the outcome values was mainly influenced by the accuracy of the presented data, its spread, and the slight imbalance towards higher scores and minor explanations in the KPI reports about KPI particularities in specific subunits.

Aligned with Ragin (2008)'s suggestion of transparency and presentation of face validity, the chosen thresholds are illustrated below.

	Full	Mostly Full	Slightly full	Cross-over	Slightly out	Mostly Out	Out	Fuzzy set
Lead time	0.8			0.5			0.25	3-value
Staff engagement	0.015			0.01			0.005	3-value
Staff satisfaction	0.25			0.15			0.1	3-value
Walking distance	0.7			0.55			0.4	3-value
Quality	1		0.85		0.5		0.35	4-value
Setup time reduction	0.95			0.8			0.65	3-value
Practices	0.79	0.65	0.51	0.41	0.31	0.21	0.02	7-value

Table 12: Fuzzy sets

The calibration was conducted on three groups. First, the smaller samples of <10 cases (walking distance, staff engagement, and satisfaction) and the medium sample of <15 cases (setup time reduction). The second stage consisted of the calibrations of the larger samples of <40 cases (lead time and quality). Finally, elicitation was used for a 7-value calibration. As the elicitation procedure has already been fully discussed, this paragraph will briefly explain the calibration of the first two groupings.

First grouping

- Walking distance/staff engagement/staff satisfaction/setup time reduction
- The first three KPI fuzzy sets had small samples, hence the decision to use 3-value sets.
- Setup time was tested as a 5-value fuzzy set but since most values bundled together in 3 groups, the 3-value fuzzy set was supported and used.
- The thresholds were chosen based on the assessment of which values represented superior improvement and the consideration of equal distances.

Second grouping

- Lead time and quality

- The increased sample led to an initial testing of 5-value sets. However, the data distribution prompted the use and refinement of the already tested pilot sets.
- Similar data structure and a shared organisational setting enabled this decision.
- The thresholds were marginally adjusted to the new data set.

Following the recommendations of (Greckhamer, 2011) and (Fiss et al., 2014), interpretation difficulties can be avoided by the addition of a constant of 0.001 to all scores.

7.5.5 FsQCA – Data analysis

This fuzzy set Qualitative Comparative Analysis follows the recommended procedure, as introduced and developed by Ragin (2008). This study investigated the impact of Lean practices (causal conditions) on key performance indicators (outcomes) in hospital subunits (population of interest). As this study aimed to identify practices in Lean bundles that have varying levels of importance, the sufficiency and necessity tests of fsQCA formed an essential role in the analysis.

(Schneider and Wagemann, 2012) suggest two measures to assess the fit of necessity and sufficiency tests: consistency and coverage. Both values are measured on a scale from 0.00 to 1.00. Consistency operates as a strength measure for the test. If the score is 1.0 it means that a present outcome is linked to the presence of a necessary causal combination. A decreasing score indicates an proportionately decreasing level of consistency. Coverage indicates the importance of a solution. A low coverage score means a low representation of cases in the solution.

That being said, empirical evidence has shown that test results are not fully aligned with the theoretical suggestion of consistency and coverage of 1.00 (Schneider and Rohlfing, 2016). While the evidence seems to be widely accepted, the conclusions differ. (Duşa, 2019) and Hugh (2019) support the rule of thumb that all results for necessity and sufficiency should be a matter of discussion if the consistency score is larger than or equal to 0.75 and the coverage score is 0.5. Although Ragin (2008) does not refute the evidence, he recommends consistency scores of greater than or equal to 0.9 for necessity and 0.8 for sufficiency. This study has taken on board the debate and therefore worked with both perspectives by selecting consistency and coverage scores based on a trade-off decision between test stability and result fineness, which is driven by context and setting.

The empirical testing of sufficiency happens through truth table analysis. A truth table presents all possible causal recipes (configurations) of causal conditions (Fiss, 2011). During the analysis, the sufficiency test was first conducted using the Quine-McCluskey Algorithm, followed by the necessity test for necessity.

As previously discussed, the substantial amount of research that links individual Lean practices to performance outcomes provides a strong theoretical basis for using easy counterfactual analysis, which is part of the intermediate solution (Ragin and Sonnett, 2005). This guided the decision towards the preference for an intermediate solution.

Additionally, this study followed Dusa (2019)'s recommendation to further investigate necessity and sufficiency through superset/subset analysis, automatically generating all mathematically possible necessary/sufficient conditions and comparing them with results that were previously theoretically driven. In this context, two measures become pertinent: relevance of necessity (RoN) and proportional reduction in inconsistency (PRI). RoN represents the relative importance of a causal condition as a necessary condition. Thereby, it allows for necessary conditions to be determined and for their different degrees of importance within the causal combinations to be established.

However, conducting superset/subset analysis can become very time-intensive as sample size increases. Some of the calculations took several hours. Using the recommended procedure by Dusa (2019), several hundred solutions were therefore excluded based on the following cut off values:

Leadtime

N – inclN = 0.9 RoN = 0.4 covN = 0.7 S – inclS = 0.7 covS = 0.51

Quality

N – inclN = 0.9 RoN = 0.6 covN =0.7 S – inclS = 0.7 covS = 0.51

The recommend cut-off value for RoN is 0.6. However, lead time did include a larger sample of cases since lead time is prominently present in several practices in the subunits. Therefore, the analysis used 0.4 as the cut-off criteria to consider all possible conditions for necessity.

Superset/subset analysis can also be used to identify sufficiency. One of the sufficiency test scores is PRI. This comes into play if simultaneous subset relations occur. According to Schneider and Wagemann (2012), this might happen if there is simultaneously one case that indicates sufficiency for the presence of the outcome, and at least one other case that contradicts this by indicating sufficiency for the absence of the outcome. In such a logically contradictory case, one of these two results must be declared causally sufficient. This can be done by determining a preferable solution. This involves calculating the product of the PRI score and the consistency score for both cases. The larger product is the preferable solution.

7.6 Summary

This chapter described the practical side of the analysis for both studies. It showed how interviews and documentation reviews were transformed through content analysis into useful information. This information was further processed to obtain KPI scores for practice implementation, lead time, and quality. Through calibration, the KPI scores were transformed into fuzzy scores, which enabled the truth table analysis. In addition, a necessity test was conducted. The analysis in the first study concluded with an acknowledgement that the study's small sample size and the only moderately robust calibration of the practice implementation scores were areas that required improvement, prompting the significant adjustments that were made to this second study.

The analysis for the second study began with a similar process of data collection. However, thanks to the lessons learnt from the first study, the sample size was increased, as was the length and rigor of the interviews. More documentation was gathered and observations were drawn to assess the validity of the other data sources. Additionally, the use of elicitation in the second study improved the robustness of the determination of the practice implementations' scores. Specifically, linguistic and practical elicitation were used to determine the differences between the implementations in different subunits and practices.

Further advances in the method supported the argument for an adaption of the data analysis. Superset/subset analysis was used to identify sufficient and necessary causal conditions for superior performance results for lead time and quality. The results for both studies will be presented in the next chapter.

8 RESULTS

8.1 STRUCTURE – RESULTS CHAPTER

In this chapter, the thesis results are presented. First, the results for the pilot will be presented, distinguishing between the results for improved quality and the results for improved lead time. The general lessons learnt from the pilot study will be highlighted, followed by a brief explanation of how these impacted on the second study. The results of the second study then follow, also distinguishing between the results for improved lead time and the results for improved quality. Finally, the chapter will close with the explanation of the results for walking distance and setup time.

It should be noted that the analysis and the result styles differ between the two studies. As the first study has a pilot character, some of the learned experience led to adjustments in the second study.

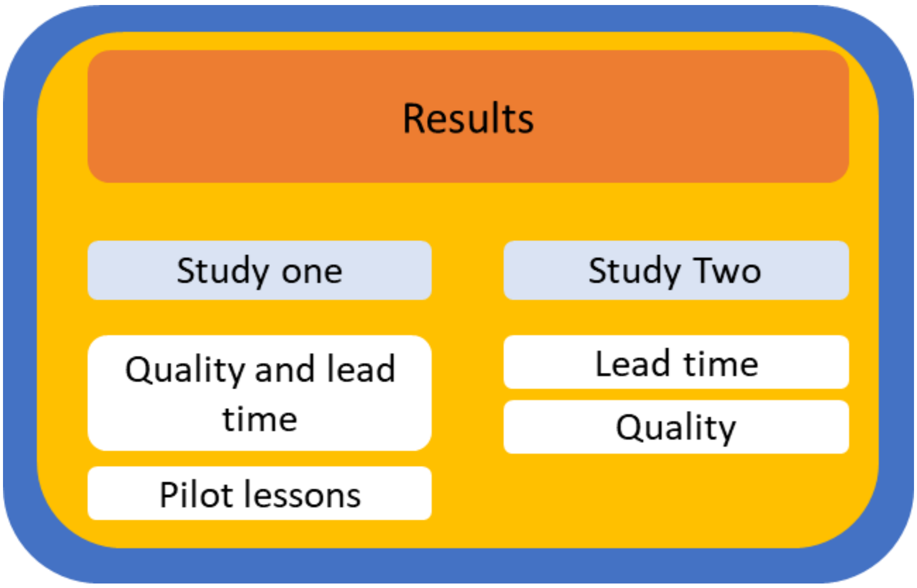

Figure 13: Structure of results

8.2 Results – Study 1

The first study was conducted in two hospitals. It investigated Lean implementations and their performance impact in seven subunits. Using fsQCA, a truth table analysis and separate necessity test was conducted. The results of the analysis of the first study can be seen in Table 13 below:

	Configurations	
	A	**B**
5S fz		
Takt fz		
Leader fz		
Mist fz		
Empower fz	●	
Redesign fz		
ProFlow fz		●
JITQuick fz		
Consistency	1.0000	0.9300
Raw coverage	0.1925	0.3043
Unique coverage	0.0007	0.0081

⬤ Necessary causal condition present

● Sufficient causal condition present

⊗ Necessary causal condition absent

⊗ Sufficient causal condition absent

Table 13: fsQCA results study one

The configurations A and B indicate that the presence of either strongly implemented process flow analysis or staff empowerment is sufficient to achieve superior performance.

While the solutions for superior performance improvement passed the threshold of 0.8, the solutions for inferior performance improvement did not.

The analysis of the pilot study recognized two configurations that lead to superior lead time improvement and no configurations that lead to superior quality improvement. It is noteworthy, however, that under a lower frequency threshold (<2) more than two dozen consistent complex solutions were found and even more intermediate ones. While that might stimulate a discussion of the appropriateness of the chosen frequency threshold, the results would potentially exceed the purpose of the pilot study.

The results imply that most Lean practices are not required to achieve superior performance improvement, except for process flow analysis and staff empowerment, which can lead to a stronger KPI improvement. This indicates a substituting relationship between staff empowerment and process flow analysis. Given the different socio-technical affiliation of both practices, it can be assumed that staff empowerment is more likely to occur in a setting (e.g. organizational and operational environment) with social issues, and process flow analysis in a technical setting. This relationship will be further discussed in the next chapter. It also cannot be concluded that the presence of strongly implemented practices will lead to inferior performance improvement, but nor will their absence.

Overall, it can be concluded that an implementation of the presented Lean practices can lead to performance improvement. However, except for the two configurations identified in Table 13, no implementation leads to high or low improvements.

8.3 Results – Study 2

This fsQCA study aimed to highlight the combinations of Lean practices that lead to superior performance improvement healthcare subunits. Furthermore, it intended to identify their relationships and their impact on operational performance.

The superset and subset analysis forms an essential element of set theory, as it enables the identification of the necessary and sufficient conditions for an outcome of interest. Both necessity and sufficiency are at the core of the QCA approach; hence the method can establish the necessary and/or sufficient Lean practices for the achievement of superior operational performance.

The analysis for both necessity and sufficiency values for lead time and quality was controlled through the testing of 20% cut-off criteria. No change to the results were found. However, the sufficiency tests did present contradictory results that surpassed the inclusion criteria of 0.7 inclS. For lead time, different results were found for leadership, mistake proofing, and kaizen huddle. In the case of the quality outcome, mistake proofing delivered a logically contradictory result. In such a case, Dusa (2019) recommends comparing the products of the Inclusion and PRI scores for both solutions to make a logical decision. This recommendation was followed.

8.3.1 Leadtime

The analysis of the performance impacts of Lean practices on lead time was done through superset/subset analysis. The results for the necessity and sufficiency tests are presented in Figure 14 below. This figure shows the configurations of different absent/present Lean practices and their relations. Every configuration is presented with three values, namely incl, RoN/PRI, and cov. For the necessity test, inclN stands for inclusion necessity; this can best be understood as the consistency score. The RoN refers to relevance of necessity, and the covN indicates the coverage of the necessity. Regarding the sufficiency tests, the inclS means inclusion sufficiency and represents an interpretation of consistency. The PRI score stands for proportional reduction in inconsistency and shows the influence of one configuration of its theoretically complete contradictory solution. Last, the covS score shows the coverage score of sufficiency, indicating the strength of the sufficiency.

Necessity Test	inclN	RoN	covN
N1 multifunctemployee*changfacility*kaizen huddle	0.915	0.439	0.773
N2 RPIWKAIZENBLITZ + ~5s	0.905	0.488	0.783
N3 processredesign+PROCESSFLOW	0.928	0.412	0.772
N4 processredesign+LEADERSHIP+MISTAKEPROOF	0.905	0.464	0.775
N5 processredesign+LEADERSHIP+STAFFEMPOWER	0.900	0.444	0.766
N6 processredesign+MISTAKEPROOF+STAFFEMPOWER	0.905	0.467	0.776
N7 ~5s+LEADERSHIP+standardpractice+MISTAKEPROOF	0.905	0.422	0.762

Notes: CAPITAL letter indicate the presence of a condition, small letters indicate the absence of a condition, also ~ indicates the absence of a condition, the sign + indicates OR, the sign * stands for combination of conditions

Sufficiency Test		inclS	PRI	covS
S1	RPIWKAIZENBLITZ	0.798	0.740	0.855
S2	standardwip	0.797	0.731	0.790
S3	~5s	0.864	0.810	0.619
S4	processredesign	0.781	0.720	0.874
S5	LEADERSHIP	0.893	0.849	0.170
S6	standardpractice	0.762	0.702	0.789
S7	PROCESSFLOW	0.850	0.800	0.630
S8	MISTAKEPROOF	0.965	0.945	0.339
S9	kaizen huddle	0.728	0.666	0.957
S10	STAFFEMPOWER	0.911	0.884	0.422
	Overall solution	0.779	0.667	0.989

Figure 14: Analysis results lead time

The superset/subset analysis testing for necessity found that 7 configurations of causal conditions were necessary for the achievement of superior operational performance (see above). Thereby, it was found that the (present) practices of obtaining leadership support, mistake proofing, and staff empowerment substitute (replace) each other, while process redesign needs to be absent in all relevant configurations. Although the configuration of the presence of RPIW with the absence of 5S scored the highest overall, this result does not seem logical as RPIWs usually introduce other tools during the workshop.

In addition, the QCA results show combinations of causal conditions that indicate complementary relations by rule of thumb (Dusa, 2019). In the case of three absent conditions, this is logically redundant.

Overall, the configurations achieved moderate RoN scores, indicating fairly weak relevance of necessity. This means that either that all solutions are not necessary, or they are quasi-necessary through their fairly weak triviality. As the RoN scores are specific for superset/subset analysis, these results enable a deeper understanding. Usually context-dependent RoN scores between 0.6 – 0.9 indicate a strong relevance of necessity. In this case, the results are too low to be considered strong but too high to be excluded. Thereby, these results become quasi necessary.

In addition, the superset/subset analysis testing for sufficiency identified 10 possible sufficient configurations of causal conditions. The results did not find configurations of more than one causal condition that fulfilled the inclusion criteria of 0.7 (inclS). Therefore, it can be concluded that individual practices dominate configurations of practices regarding sufficiency. The Lean practices of obtaining leadership support, mistake proofing, and staff empowerment achieved fairly weak coverage scores, expressing the degree to which a causal configuration explains an outcome (Schneider & Wagemann, 2012). This can be explained through the smaller number of cases with leadership and mistake proofing practices compared to the medium-sized outcome sample. Conversely, the absence of kaizen huddle, the absence of process redesign, and the presence of RPIW are the causal conditions most capable of explaining the superior lead time. The low coverage scores for mistake proofing and obtaining leadership support led to the rejection of these solutions.

Furthermore, it was tested whether the results would change when the analysis was conducted in the organisations impacted by only one of the KPO teams while excluding the other. No changes were found except for staff empowerment. Here, the test found that the subunit implementations in those subunits managed by one of the teams scored inclS 0.872 and covS 0.764, while the other had 0.931 and 0.218. Therefore the decision was made to not exclude the result.

None of the results was found to be strictly necessary, but the results S1, S2, S3, S4, S6, S7, S9, and S10 were found to be sufficient.

The findings therefore show that individual Lean practices have a higher impact on lead time than that exerted by their inter-practice relationship. To clarify, these relationships still have an impact; it is simply weaker than that that exerted by the contributing practices.

8.3.2 Quality

This paragraph will show the results of the fsQCA analysis of the performance impacts of Lean practices on quality. The table below shows the configurations of Lean practices found to be necessary and/or sufficient. No new scores were introduced.

Necessity Test	inclN	RoN	covN
N1 STANDARDWIP+processflow+MISTAKEPROOF	0.903	0.666	0.752
N2 STANDARDWIP+processflow+STAFFEMPOWER	0.907	0.652	0.746
N3 STANDARDPRACTICE+processflow+MISTAKEPROOF	0.904	0.630	0.733
N4 STANDARDWIP+~5s+LEADERSHIP+PROCESSFLOW	0.922	0.617	0.734
N5 STANDARDWIP+~5s+LEADERSHIP+STAFFEMPOWER	0.900	0.741	0.796
N6 ~5s+LEADERSHIP+PROCESSFLOW+CHANGFACILITY	0.909	0.625	0.733
N7 STANDARDWIP+~5s+LEADERSHIP+STANDARDPRACTICE +MISTAKEPROOF	0.916	0.679	0.765
N8 STANDARDWIP+~5s+LEADERSHIP+MULTIFUNCTEMPLOYEE +MISTAKEPROOF	0.914	0.747	0.804
N9 STANDARDWIP+~5s+LEADERSHIP+MISTAKEPROOF +KAIZENHUDDLE	0.913	0.750	0.806

Sufficiency Test	inclS	PRI	covS
S1 STANDARDWIP	0.951	0.824	0.584
S2 ~5s	0.852	0.626	0.632
S3 process flow	0.758	0.510	0.729
S4 MISTAKEPROOF	0.869	0.756	0.453
S5 STAFFEMPOWER	0.832	0.576	0.468
Overall solution	0.760	0.580	0.923

Notes: CAPITAL letter indicate the presence of a condition, small letters indicate the absence of a condition, also ~ indicates the absence of a condition, the sign + indicates OR, the sign * stands for combination of conditions

Figure 15: Analysis results quality

In relation to superior quality performance, the necessity test found 9 configurations of causal conditions. All configurations include logical disjunction OR (+) relations. A logical disjunction means that at least one of the criteria must be fulfilled to achieve the outcome of interest but that all can also occur.

The study also found evidence for substituting relationships. For example, standard work in progress, mistake proofing, staff empowerment, and leadership all substitute each other. The relevance of necessity scores in all configurations are solid (~0.6) to good (~0.8).

The superset/subset analysis investigating sufficiency found 5 configurations of the following single Lean practices: the presence of standard work in progress and staff empowerment, and the absence of 5S and process flow analysis. While configurations consisting of complementary Lean practices were found, they did not meet the inclusion criteria. Again, this indicates that individual practices dominate the configurations of Lean practices regarding sufficiency. Both mistake proofing and staff empowerment scored moderate coverage values, meaning that they are able to explain only around half of the quality improvement. As previously discussed, there were differences in the results for staff empowerment according to which of the two different KPO teams was in charge of the subunit implementation, with Team One's subunits passing the thresholds. Therefore, it was decided to continue to include staff empowerment. In addition, we will return to the scores for mistake proofing in the discussion section since they may indicate that the practice, although neither necessary nor sufficient, is nevertheless still strong enough to be of interest.

The results N1 to N9 were found to be necessary and the results S1, S2, S3, and S5 sufficient.

8.3.3 Walking distance & Setup time

The subset/superset analysis for walking distance and setup time did not provide any results that met the inclusion criteria. For walking distance, the first proposed solution had an inclusion score of 0.48 (cut-off 0.9). In the case of setup time, the inclusion score was 0.72 for a 0.8 cut-off score for sufficiency, and 0.8 for a 0.9 cut-off score for necessity. Therefore, the investigated Lean practices are neither sufficient nor necessary for the achievement of these KPIs.

8.4 Summary

This chapter presented the results for the first and second studies. It found the combinations of Lean practices that are necessary and/or sufficient for superior performance improvement in lead time and quality for both studies. Furthermore it showed differences in results between the two studies, and highlighted that other factors might impact Lean implementations on a subunit level. The implications of these findings will be discussed in detail in the next chapter.

9 DISCUSSION

9.1 STRUCTURE DISCUSSION

The following chapter will explain and discuss the findings of this thesis. The chapter will start with separate explanations of study one and study two. The discussion of study one will clarify how the pilot impacted and guided the findings, and this section will link the results of study one back to the literature. A discussion of the second study will follow. Here, the findings for lead time will be discussed and then the findings for quality. This distinction is necessary to avoid confusion since different configurations of practices lead to different outcomes. The chapter will then explain the links and differences between both studies, starting with an explanation of why only single practices meet the sufficiency and necessity OR configurations and closing with the explanation of context dependence. There will follow an in-depth discussion of the implications of adopting a subunit lens for this research on Lean. The overall limitations of this research will be noted, as will some recommendations for future studies. Finally, the section closes with an in-depth interpretation of the range and impact of these findings for scientific and managerial work. The discussion chapter might in places, and particularly with regard to the fsQCA methodology, echo the results chapter but this chapter will focus on the accepted results.

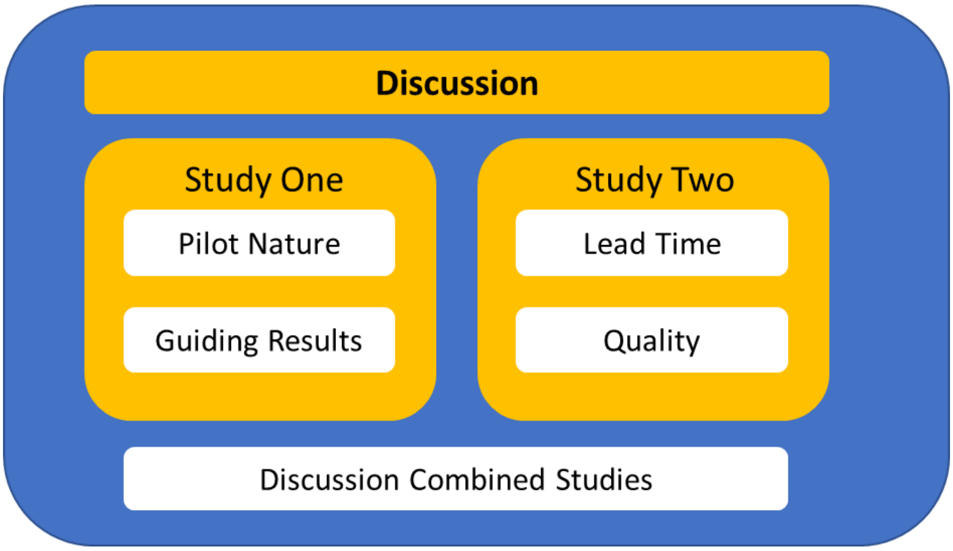

Figure 16: Overview of Discussion

9.2 Discussion – Study One

9.2.1 Pilot character

Pilot studies may explain and/or test future studies through their shared methodological and analytical framework (Jairath, Hogerney, & Parsons, 2000). The first study in this thesis was created as a pilot for the second study; both studies applied the still new fsQCA approach in the context of a Lean implementation in the healthcare industry. The data sources and sample were objects of interest for the pilot, which found that the KPO Leads hold a quasi-monopoly on information. The KPO Lead and the KPO Team decided not only which KPIs were collected, they were also responsible for their measurement. They are therefore the main corporate players responsible for planning and managing the implementation. All other organisational units receive their performance reports and evaluations about the implementation from the KPO Team.

While the first stage of data collection was relatively easy because only one access point was required, this also limited the collection process as in-depth analysis puts a bigger time burden on the participants in the long run. More importantly, the KPO Leads could not be left with sole responsibility for data provision because their subjective perspective would dominate the results. Supplementary objective reports and observations were therefore added to smooth the impact of this limitation.

Second, the pilot study found that the overall sample size of 15 to 20 subunits falls into the classification of a small sample (Ragin, 1987). This limited the capacity to investigate different KPIs, as their individual samples were between 5 and 10 (included) or below 5 (excluded). While these values are at the lower end of suggested sample size, they are supported by the strength of the information about each case, which is sufficient to generate a generalisation. Additionally, having a larger sample for the pilot study may have been counter-productive, given that the overarching aim of a pilot is to protect the future study from limitations that could not have been uncovered before conducting the data collection (Connelly, 2008); the small sample size turned out to be one such limitation. Since fsQCA usually operates with small (5-25) to medium (20-50) sized samples with a strong context dependence, the original sample was within the acceptable range for a pilot. Nevertheless, this limitation was a lesson learnt for the second study, resulting in an increased overall sample to cover smaller KPI samples.

9.2.2 Guiding Results – Study One

9.2.2.1 Different impacts of social and technical practices

This study investigated the performance improvement of the following Lean practices: 5S, takt time, leadership support, mistake proofing, staff empowerment, process redesign, process flow analysis, and JIT, by applying fuzzy set qualitative analysis. While all practices were found to have positive performance impacts, the truth table analysis demonstrated that only staff empowerment and process flow analysis were sufficient.

This means that superior performance improvement cannot be effected without the implementation of either process flow analysis or staff empowerment. Both practices improve Lean implementations, albeit via different paths.

9.2.2.2 Staff empowerment

Honold (1997) described staff empowerment as a multi-dimensional construct, consisting mainly of continuously trained, self-responsible and collaborative workers and teams, a staff developing and envisioning leadership, a flexible operational structure based on balances and checks and a fair reward systems. This multidimensional perspective enables the combination of KPO team efforts to engage staff and in return the actual commitment of staff members with Lean actions. While the combination of both sides was used, the trusts had a slightly stronger focus on the KPO teams' efforts to empower staff to participate.

This study's analysis found staff empowerment to be sufficient for performance improvement. This implies that creating a culture in which staff are motivated, supported, and inspired to improve care is sufficient to cause visible performance changes. At the observed hospitals, nurses engaged with Lean workshops during their lunch breaks because the creation of an atmosphere that welcomed feedback made them feel their suggestions were valued. During their careers, they had experienced niggles that negatively impacted on their care giving, but the environment was not conducive to striving for improvement. Furthermore, they felt rejected when they brought up improvement suggestions that were disregarded. Through the changed work atmosphere, staff members felt relieved and encouraged, and the feelings of rejection dissipated. This could also be seen in the increase of the participation rate of staff members at Lean events as well as the decrease of the hospital's overall sickness rate, which might relate to other facts but was also visible in the other participating hospitals. Staff empowerment increases the level of job satisfaction and motivation that, in healthcare, promotes a stronger commitment to patient outcomes and experiences (Neuhauser, 2011). Engaged medical staff can be considered as a pool of knowledge for the KPO Team. These professionals develop, discuss, and implement change from a front-line perspective (Fine et al., 2009). Furthermore, staff empowerment operates as an enabler for improvement by strengthening the receptiveness of staff to training and other initiatives (Henderson et al., 2013). It thereby forms synergies with other practices and enhances the improvements effected by these. The combined concepts of sufficiency and fuzziness probably found several different configurations with the shared condition of staff empowerment. In consequence the study found a very strong case for the enhancing presence of staff empowerment. Following up on these findings, the second study will be discussed more deeply the impacts of staff empowerment.

9.2.2.3 Process flow analysis

Medical staff are highly trained within their specialised fields so they can provide the best possible care while also being able to adapt to changes in medical conditions. However, they rarely learn about process improvement. Process flow analysis provides the tools for the visualisation, optimisation, and planning of waste free operations from the perspectives of all stakeholders, leading to improved overall process performance (Rahani and Al-Ashraf, 2012). In one of the subunits, a nurse explained that they had stopped their regular work to carry out an ad hob restock of medicines. A simple future process map was used to visualize the inventory and supply of the ward's medicine store. The process flow map enabled duplicated inventory reports to be identified, thereby preventing unscheduled supply runs. While the visualisation of process flow analysis improves the understanding of processes, it also equips its users with the tools for planning the necessary change to processes. Linking activities to resources and breaking down patient care into small steps allows costly/time consuming process elements to be identified and to focus the improvements on these. The multifunctionality and simplicity of process flow analysis combine to make this practice a technical all-rounder that, if fully implemented (i.e., such that it fulfils all the elements of visualisation, planning, process redesign, and waste reduction), it can be a free-standing source of performance improvements.

The second study will deliver further insights into process flow analysis, which will be elaborated upon in the combined discussion.

9.3 Discussion – Study 2

9.3.1 Lead time

The second study applied fsQCA to investigate the performance impact of different configurations of Lean practices in healthcare on lead time and quality. The sufficiency tests and necessity tests will be discussed separately.

9.3.1.1 Sufficiency results

The analysis found eight conditions to be sufficient for strong performance impacts. These include the three present practices of RPIW, process flow analysis, and staff empowerment. Also included are the five absent practices of standard work in progress, 5S, process redesign, standard practice, and kaizen huddle. This means that only one of these conditions needs to be met to achieve superior improvement in lead time.

Standardisation in healthcare is a contentious issue. While it has been shown to successfully reduce medical errors and knowledge loss in the frontline (Mannon, 2014), it also carries a negative connotation for the majority of senior medical staff, who are trained to function independently (Timmermans and Berg, 2003). Therefore, standard work implementations can only be essayed after making trade-off decisions between operational gains and social losses.

The second study only partially validated the findings of the first study regarding staff empowerment in the context of improving lead time. Overall, the scores did not pass the sufficiency test, but all the subunits of one hospital trust did pass. The major difference between the implementations lay in the focus given by the KPO teams to staff empowerment and how it was implemented. In the interviews, the KPO Team that had decided to focus on staff empowerment mentioned that the staff experience of improvement approaches, which typically took the form of short-lived management diktats, had instilled in staff a distrust of such practices. The other KPO Team was responsible for a newly formed care facilities group that had little experience of big improvement approaches. It was clear to the first KPO Team that implementing staff empowerment was crucial to the success of the overall implementation. It therefore focused on creating a culture of strong staff engagement. The other KPO Team paid lip service to staff empowerment by enacting some engagement exercises. As staff empowerment is an enhancer of lean implementations (Henderson et al., 2013), its strength depends on its links to other practices. If fully embraced, a strong staff empowerment culture can enhance the positive impact of other practices on lead time. Alternatively, if the staff have previously had a negative experience with improvement implementations and are therefore somewhat resistant to change, staff empowerment is a tool that can storm the barricades and enable a successful implementation.

Just as in the first study, the practice of process flow analysis was found to be sufficient to effect strong performance improvements. The second study found this to be case for lead time in particular. Process flow analysis practices, such as value stream mapping, provide simple tools to visualize and evaluate care processes. They link process steps to resources and thereby spark insights in staff members and promote consensus between them (Teichgräber and de Bucourt, 2012). Notably, the technical nature of the practice and the common outcome of the analysis (i.e., the reduction in process steps) directly or indirectly reduce the process time. For example, Johnson et. al. (2020) described how process flow analysis was used to investigate a compressed patient flow between the A&E, ambulatory emergency care and day-case units. The patient flow was rarely constant and ambulatory medical outliers would regularly end up blocking beds on the day-case unit. Following the process flow analysis the described hospital decided to increase the ambulatory capacity by transferring a small part of the day care unit over. Thereby, the bottleneck in the ambulatory unit could be relieved and the patient flow through the entire hospital was improved and significant resources were freed.

Although three single present and five single absent practices were found to be sufficient for superior lead time improvement, these practices were mostly implemented in sets of more than one. According to the KPO leads, the focus of resources and training tended to be concerned with a single practice as starting point of the implementation but often remained the sole practice implemented. Only small additions or adaptions were made if required.

9.3.1.2 Necessity results

This study found five configurations of several practices that passed the necessity and logical inclusion tests. Necessary configurations mean that superior performance impacts were achieved if one of these configurations is present. However, the configurations had only limited relevance to the necessity scores, implying weak necessity. Furthermore, all five configurations are internally linked through OR functions. An OR function means that all causal conditions of the configuration can occur together but that at a minimum, one of them needs to occur. Below, the identified necessary configurations are listed:

- processredesign OR PROCESSFLOW
- processredesign OR LEADERSHIP OR MISTAKEPROOFING
- processredesign OR LEADERSHIP OR STAFFEMPOWERMENT
- processredesign OR MISAKEPROOFING OR STAFFEMPOWERMENT
- ~5S OR LEADERSHIP OR standardpractice OR MISTAKEPROOFING

Note: CAPITAL letters show the presence of a condition, lower letters and (~) show the absence of a condition

For example, when the practice of process flow analysis is being implemented, it will not bring about superior improvement of lead time if it is combined with the practice of redesigning a process because the practical implementations of process flow analysis tools tend to redesign processes themselves. This is particularly evident in the value stream mapping tool, which develops a new process flow based on the experiences of the old one and the improvements to it. For example, in one of the quarterly reports the process of fetching and resupplying medicine and drugs from the medicine cabinet was described in greater detail. Retrieved medicine/drugs were noted physically at one of three positions in the ward. This increased probability of missing or late documentation entries, resulting in missing medicine which had to be organized from the store or another more distant medicine cabinet. The KPO team conducted process flow analysis and suggested the removal of two of the documentation points and focus it instead on the last remaining. So the process was minimized but technically not redesigned. Hence, process flow analysis substitutes for elements of process redesign and makes its full implementation situationally redundant. An attempt to implement the two practices together might result in a waste of resources, which is clearly counter-productive.

Additionally, the practices of staff empowerment, leadership and mistake proofing occur in different OR-combinations. Both KPO leads described the training and support of middle and upper leaders as essential to move faster. For example, the leadership committed to control agreed changes by either physical checks or performance checks of monthly or quarterly reports. Thereby they were continuously aware of Lean improvements and required resources, which increased the speed of the signing-off process of budgets. In one case an early RPIW event sparked the motivation of both attending middle management and staff members, leading to directly agreed change (moving of the welcome desk) and staff going into overtime to facilitate it. This led to slight improvements of lead time.

The combination of no single necessary solution dominating other solutions in terms of strength and the OR link between the practices means that the found configurations are strong enough to be used as a guiding example but not strong enough to be used as a one-way ticket to strong lead time. This makes an argument for taking RPIWs, process flow analysis, and staff empowerment individually and carefully combining them with other practices like leadership support and mistake proofing in the manner suggested by the found solutions to achieve stronger lead time improvements. The choice of the solution should be context dependent.

9.3.2 Quality

9.3.2.1 Sufficiency results

For the quality KPI, the same fsQCA approach was used as for the previously presented lead time KPI.

The investigation into quality found three present practices, namely standard work in progress, mistake proofing, and staff empowerment. However, all the sufficient present conditions achieved only moderate coverage scores, indicating that these practices also occur in several solutions that do not lead to superior quality improvement. Furthermore, the results showed the absence of 5S and process flow analysis.

The analysis for both lead time and quality found mistake proofing to be sufficient. However, only the quality analysis provided sufficiency coverage scores that passed the logical inclusion. While quality in manufacturing industries is established through product control and product exclusion, this is not an option for healthcare. Therefore, the trade-off decisions between quality and time are usually carried out differently. Factories avoid process stops because they cause fewer goods to produced and thereby impact on revenue; in healthcare, processes get interrupted to safeguard/improve quality. Hospitals apply mistake proofing to identify patterns that result in errors so that repetitions can be avoided. By implementing mistake proofing, a subunit can control, regulate, and improve its error rate and consequently improve the quality of care. For example, Grout & Toussaint (2009) described how the mistake proofing practice stop-the-line was used to shut down operating theatres after an increase of surgical related infections was found. The root-cause was investigated and found that physicians and nurses weren't following the protocols for hand-washing in adequate manners. In consequence, supervisors observed and controlled the hand-washing after the operations theatres were used again. In the long-term and after compliance with the handwashing rules was achieved, random assessments replaced the supervision. The combination of mistake proofing actions let a stop of surgical-related infection.

Standardisation was tested for sufficiency for superior performance improvement to lead time and quality. However, only the quality analysis gained support from the results. Standardisation creates process efficiencies through safety, quality control, and mitigation (Monden, 1983). This means that in hospitals, standardisation has a stronger impact on quality than it has on lead time. Although standardisation, like standardised work in progress sheets, can reduce process time through a better flow of knowledge, this has a smaller impact on the KPI than the impact exerted by preventing a medical mistake.

9.3.2.2 Necessity results

Similar to the lead time results, the quality results also found configurations of more than one practice during the necessity tests. Below, the found necessary configurations can be seen:

N1 - STANDARDWIP OR processflow OR MISTAKEPROOFING

N2 - STANDARDWIP OR processflow OR STAFFEMPOWERMENT

N3 - STANDARD PRACTICE OR processflow OR MISTAKE PROOFING

N4 - STANDARDWIP OR ~5S OR LEADERSHIP SUPPORT OR STAFFEMPOWERMENT

N7 - STANDARDWIP OR ~5S OR LEADERSHIP SUPPORT OR STANDARD PRACTICE OR MISTAKE PROOFING

N8 - STANDARDWIP OR ~5S OR LEADERSHIP SUPPORT OR MULTIFUNCTIONAL EMPLOYEE OR MISTAKE PROOFING

N9 - STANDARDWIP OR ~5S OR LEADERSHIP SUPPORT OR KAIZEN HUDDLE OR MISTAKE PROOFING

Note: CAPITAL letters show the presence of a condition, lower letters and (~) show the absence of a condition

The analysis found seven different configurations. All configurations included standard work practices. Furthermore, all configurations either had the absent condition of process flow analysis or 5S. Other configurations included mistake proofing, staff empowerment, leadership support, multifunctional employee, or kaizen huddle. However, all the configurations were OR configurations again, meaning that all conditions of the configuration can occur together but there must be a minimum of one.

Five of the seven configurations (N2, N4, N7, N8, and N9) combined social with technical practices. Social practices have an enhancing effect on Lean practices, especially technical ones (Hadid and Afshin Mansouri, 2014). The results showed that leadership support, kaizen huddle, multifunctional employee, and staff empowerment could be combined with standard work in progress, facility layout changes, mistake proofing, or standard practice in the given configurations. Following participation in Lean events, medical staff become more motivated and active in care improvement (Radnor, 2011).

One of the quarterly reports presented a case were leadership support worked in combination with process flow analysis, process redesign and a change of the facility layout. The improved subunit was an intense care ward focussing on post operational care. During a Gemba walk a junior executive was made aware by an older patient about the lack of enough emergency buttons or lines at the floor of the bed room, calling for help if a fall occurred. While the report indicated the presence of an emergency line reaching the floor, this line was on one side of the bed but if the patient fell down on the other side he/she would have needed to crawl around the bed. The next RPIW report showed that there was not just a lack of buttons/lines but also a slow reaction time until helping staff arrived at night. Through the early involvement of an executive manager and the presence in the RPIW the swap of a bedroom with the nurse station as well as the adding of emergency lines to the floor was quickly authorized and necessary construction work conducted. In addition, process flow analysis and process redesign for the entire subunit's operations enabled the restructuring of the daily task allocations. This enabled a second nurse to be present at morning/night shifts (doing preparations and planning) and react to alarms if required.

Previous research has found that leadership support and staff empowerment complement each other (White et al., 2013), whereas visible leadership (Steed, 2012) and showing dedication can convince sceptical and/or uncommitted employees (Mazur et al., 2012). Although guiding and training leaders will promote staff engagement, an engaged staff is more likely to offer improvement suggestions to leaders. Combining both practices with mistake proofing may (further) improve care quality. The practices of stop the line (which mitigates risks) and poka yoke (which prevents risks) depend on staff participation and training. Untrained and/or uncommitted staff are less likely to recognise and report quality risks.

9.4 COMBINED DISCUSSION

The combined studies contribute to multiple scientific discussions. Their main contributions lie in their identification of a path for gaining in-depth understanding about Lean in the heterogeneous and causally complex healthcare sector. The path involves applying fsQCA, adopting a subunit lens, and drawing on the single practice sufficiency. Furthermore, the studies contributed in minor ways to the literature by validating the QCA method in a novel sector, assisting with Lean's struggle to show superiority to other improvement methods, and identifying relevant practices for the achievement of superior improvement gains of quality and lead time. The structure of this section will follow a top-down approach, starting with the general discussion of pilot studies for fsQCA adoption in healthcare.

9.4.1 fsQCA pilot in healthcare recommended

The first study's pilot aim was to find limitations for the second study, which were preventable but could not be found until the actual method was applied. The pilot signposted the potential results from the subsequent study, and also showed that the theoretical framework, the planned methodology, and the practical setting were a suitable fit. In addition, it provided insights into data set consistency, the KPO Team information monopoly, and the restrictions in the samples that could have been a burden for the second study if they had stayed unknown. Also, the calibration process provided thresholds and identified the weakness and dependence of the practice calibration scores, which led first to adjustments in the calibration and also to the addition of elicitation. The final benefit gained from the pilot is the practical experience gained from conducting it.

That being said, the pilot required significant additional preparation time as it was laid out like a full study. Although the similarity of process in the two studies created synergies, accessing data and redoing interviews was a time burden. A pilot therefore becomes a trade-off between additional time, resources, and theorising actions, and the identification of barriers for the future study. Despite the inevitable time-cost, the complex and unpredictable environment of healthcare organisations tips the balance in favour of recommending pilot studies for future fsQCA adaptions in hospital settings.

9.4.2 Healthcare's complex environment

Implementing change in healthcare is met with barriers, resistance, and apathy. KPO teams struggle to succeed and are forced to adapt their implementations to meet organisational and operational challenges. To name but a few of these challenges, the overlapping responsibilities in management roles, the fire-fighting mentality of staff, and the more or less negative perception of Lean practices and terminologies (Souza & Pidd, 2011) are well known issues that put every Lean healthcare implementation at risk.

9.4.2.1 Lean in healthcare

One purpose of this research was to identify empirically validated combinations of Lean practices in healthcare subunits. The analysis did not find generalizable typologies of bundles in healthcare. Instead, several individual Lean practices were found to be sufficient for the achievement of superior lead time and quality. This apparently refutes (Shah and Ward, 2003)'s Lean bundle theory in relation to healthcare subunits, given that bundles consist of multiple interrelated practices rather than individually operating practices. However, this study does not come to this glib conclusion. Several configurations of Lean practices were found to have complementary and substituting relations. However, their solution recipes could not be empirically considered because they did not meet the consistency, coverage and/or inclusion thresholds. This means that configurations of several practices were found to have a fair impact on operational performance. However, these solutions were dominated by solutions consisting of individual practices. Furthermore, it remains unclear if bundles can consist of individual practices when their impact extends beyond the boundaries of one subunit. For example, in one hospital, staff empowerment started in ophthalmology, causing staff members from other units to ask the KPO Team if they too could participate in the Lean implementation. This indicates that the practice reduced barriers of resistance towards change in other subunits where Lean practices had been applied but without the support of staff empowerment. The studies showed that Lean bundles do occur on an organisational level in hospital and thereby validated the applicability of Lean bundle theory (Shah and Ward, 2003) to healthcare.

9.4.2.2 fsQCA in healthcare

Healthcare organisations and operations are burdened with complexity and heterogeneity. Traditional linear methods have struggled to cope with these issues. The studies addressed this by applying the still new method of fuzzy set qualitative comparative analysis (Ragin, 2000, Ragin, 2008). The methodology allowed for the consideration of asymmetric, causally complex, equifinal, and conjunctional relations between Lean practices and their organisational setting. The application of fsQCA in the lean and health environments further validates the suitability of the method, most particularly the application of superset and subset analysis for determining sufficiency and necessity (Duşa, 2019), something that has rarely been done in a Lean context. Contrary to standard fsQCA approaches, this procedure calculates all possible solutions and assesses their expressiveness. While a limited sample of studies have already applied fsQCA in a Lean bundle context (e.g. Galeazzo and Furlan, 2018), the use of superset and subset analysis allows the determination of solutions to be empirically rather than theoretically driven. The traditional fsQCA approach uses truth table analysis and minimisation to identify configurations of Lean practices. However, to do this the researcher must predefine a range of practices, which the software then uses to reduce the size of the solutions. This is mainly achieved through theoretical and context-led predefined typologies of Lean bundles. However, this risks overlooking less prominent Lean practices. Subset and superset analysis empirically derives all possible solutions of Lean practices and leaves the context-driven logical exclusion to the researcher.

Therefore, fsQCA investigations more accurately provide configurations by reducing the risk of overlooking possible solutions, which is particularly relevant given the combined complexity of Lean and health organisations. The superset/subset analysis embedded in fsQCA enables the researcher to better analyse and understand the phenomenon without overly sacrificing robustness and validity.

9.4.3 Performance impacts - similarities and differences in the study results

The first study found configurations of process flow analysis or staff empowerment to be sufficient for superior performance. These practices are good examples of socio-technical practices (staff empowerment being social, and process flow analysis being technical). They are well suited to address social or technical problems. Conversely, the second study confirmed the impact of process flow analysis on lead time but also showed that only one KPO Team's implementation of staff empowerment met the thresholds; this was the same KPO Team that was investigated in the first study but with a larger sample.

Staff empowerment seems to play a special role in the Lean practices implemented at subunit level. If implemented as an organisational culture, staff empowerment showed spill-over effects to other subunits, in which it proved to be more successful than a subunit focused practice. Nonetheless, a sample of only two KPO Teams, only one of which confirms the claim, is not enough for a generalisation. Even given the extensive in-depth data available for both teams, this sample is simply too small. However, it does strongly indicate the need for further research into the individual implementations of single practices (especially staff empowerment) in a complex environment. Future research should investigate how the characteristics of KPO teams and their organisations impact staff empowerment implementations.

9.4.4 Subunit lens

9.4.4.1 Only single practices sufficient

Both studies intended to find sets of Lean practices in all parts of the analysis. This aim was based on the extensive literature that finds and discusses full practice sets (e.g. Galeazzo and Furlan, 2018, Costa and Godinho Filho, 2016, Shah and Ward, 2003). Most of these studies investigated the organisation-wide impact of Lean practices, and assume that the Lean entity comprised of all the practices impacts the entire organisation. This work therefore approached the empirical work under the assumption that similar relations and sets would occur on a subunit level. Adapting the organisational lens for application to the subunit level (Whetten et al., 2009) supported fsQCA, bolstering data robustness by providing an in-depth perspective on the complexity of the healthcare organisations and operations. Looking at subunits instead of entire organizations enabled the researcher to understand the practical implications of Lean implementations on front line work in a greater detail.

However, both studies found only single practices (present and absent) to be sufficient. This means that the impact of Lean implementations in healthcare differs according to whether a subunit or an entire organisation is investigated. The literature review had found as many as 126 articles with empirically validated sets of Lean practices in healthcare, most of them on an organisational level. Hence, the two studies that comprise this research extend the discussion around Lean by shedding light on the organisational and subunit lenses.

The findings from these studies enable practitioners to focus their resources on single practices in their targeted subunits, rather than trying to simultaneously coordinate several practice implementations. Although the coverage scores (and several configurations of multiple practices excluded from the results) indicated that sets can occur, they do so only rarely. Consequently, in healthcare subunits single practices have the advantage of sets of Lean practices in achieving superior performance improvement.

9.4.4.2 Many OR configurations

Both the SLR as well as the fsQCA method contribute to the Lean bundle theory. Although the results of the sufficiency tests might call into question Shah and Ward (2003)'s theory, the necessity tests did find configurations of multiple practices and thereby provide evidence in support of the theory. Furthermore, the SLR identified numerous practically derived combinations of practices in healthcare settings.

Except for one configuration (combination of only absent practices), all necessary configurations had OR relations. The statistical OR relation means that all conditions can occur in any combination but at a minimum, one condition must occur. Simply speaking, only individual practices or small sets of practices achieve the highest performance improvement. These findings do not contest the sufficiency findings but rather specify them. While this means that individual practices can be both enough and required at the same time, other results are also possible. Thus, this work gives a strong indication of the presence of equifinality (multiple ways leading to one result), which speaks to the strength of the fsQCA model. This has two impacts:

First, while single practices dominate sets of practices, in some cases they may occur in these sets but take a leading role. Second, the OR condition in most accepted solutions implies that a vast variety of combinations of Lean practices leads to superior performance improvement. The tests only limited the choice of practices. Furthermore, this finding indicates that the layout of the configurations might be dependent on something else.

Practitioners striving for improved performance should start with one of the sufficient practices and focus all their resources and attention on it. During the progress of the implementation, they may later add additional practices selected from one of the necessary sets; these will vary with the implementation target and subunit setting.

9.4.4.3 Context dependence

This work found evidence for the context dependence of a Lean implementation at multiple points. The literature review had already identified that the motivation for the Lean implementation has an impact on the combination of practices but the previously discussed combinations of single sufficient practices and necessary OR configurations provided this work's first indication for a context-setting dependence. Moreover, the investigation of staff empowerment showed that two different KPO Teams in similar environments can achieve different results with the same practice; this clearly indicated the relevance of the KPO Team's background. Finally, the second study found that although the Lean practice standard work in progress is a sufficient condition for superior quality improvement, its absence is a sufficient condition for lead time. This research therefore establishes that the creation of a bundle of Lean practices in one subunit becomes a trade-off decision between performance indicators.

While the increasing number of fsQCA Lean studies (e.g., Garleazzo & Furlan, 2018) indicates an awareness of context dependence, only a few studies have actually specified it (e.g. Raab et al., 2006). This work further stresses the importance of future investigations into the dependence of Lean configurations on their surrounding context. The exact social and technical factors and their influences need to be identified to better understand how Lean bundles are configured.

9.5 Summary

This chapter has considered the impacts the results of these two studies may have on the scientific discussion as well as their implications for operations managers generally, and for those working in healthcare specifically. The first study showed that technical practices have a strong impact on technical KPIs. To be precise, process flow analysis plays a major role in respectively improving lead time and quality in healthcare.

Through the results of the second study, this chapter argued that different combinations of Lean practices lead to different performance outcomes. In addition, the second study shed light on the complex organisational situation in healthcare by adopting a neo-configurational lens that allowed the identification of typologies of Lean practices in healthcare. Furthermore, this study contributed to the still new fsQCA method by validating its use in a Lean/healthcare setting.

The combination of the neo-configurational lens, fsQCA, and the causal complex setting (i.e., a setting requiring a subunit lens) showed that Lean bundles occur more frequently on the organisational level than on the subunit level. As some lean practices at subunit level go beyond the boundaries of their organisational unit, the relationship between practices in Lean bundles becomes more complex, requiring a deeper understanding of the subunit/practice relations.

Both studies' results showed only OR configurations. This chapter argued that this indicates that Lean bundles on a subunit level do occur but that single practices are more often found.

Whether a practice is free-standing or exists in the construct of a bundle depends on contextual factors like organisational setting, staff experiences of improvement approaches, performance targets, and the motivation behind the implementation. Although this thesis provides a strong first look on Lean bundles as highly context-dependent, more research is required to evaluate the exact contextual factors and relations.

10 CONCLUSION

10.1 OUTLINE

This chapter presents the answers of this work to the following investigated research question:

❖ Which combinations of Lean practices lead to superior performance improvement in healthcare subunits?

The causal complexity of this study, especially the heterogeneity of healthcare operations and organisations, led to the decision to adopt a fsQCA approach. As the method is still relatively new in this research field, two studies were conducted, with the first being in the nature of a pilot. Based on experience gleaned from the first study, the second study increased the sample size and gathered data through interviews and documentation review. The analysis followed the recommendations of (Ragin, 2008) for the application of fsQCA and (Duşa, 2019) for the superset/subset analysis, finding several configurations of Lean practices that achieved high performance.

These results will be summarised in the next paragraphs. The first paragraph reflects on the methodological adjustments required by the organisational setting. This is followed by a paragraph that describes the links between Lean practices and operational performance in healthcare. The subsequent paragraph will review the subunit lens. Finally, this chapter will close by outlining the studies' limitations and the recommendations for future research.

10.2 REFLECTIONS ON THE ORGANISATIONAL SETTING AND THE METHODOLOGY

Healthcare organisations epitomise complexity like no other industry (Braithwaite et al., 2017). The investigation and analysis of processes and operations is hampered by the heterogeneity that characterises the sector. The complexity within as well as between Lean projects is yet another burden (Ferreira et al., 2019). Traditionally, these barriers were met with forced robustness and simplification by using linear approaches like the regression model (e.g. Hofer et al., 2012, Furlan et al., 2011, Eroglu and Hofer, 2011). When considering the relationships between bundles, a regression model assumes practices to be simply elements of bundles, with all characteristics shared. Thereby, it neglects interrelations within and around the bundles. By viewing Lean implementations as configurations of diverse combinations of Lean practices, this work has shed light on specific practices and offers an alternative explanation of the relationship between Lean services and operational performance by regarding it through a neo-configurational lens. It contributes to previous literature by expressing the impacts of practice implementations through the degrees of set membership, as well as diverse configurations of Lean practices on operational performance Thereby, it provides a finer-grained analysis of Lean implementations.

In this organisationally and causally complex setting, fsQCA provides the necessary theoretical and empirical tools to conduct a finer grained but still robust analysis. On the theoretical side, fsQCA allows equifinal, asymmetric, and conjunctional practices to be investigated, enabling it to reach deeper layers of detail than linear models. The application of fsQCA further validates the specific use of the method in healthcare. Furthermore, fsQCA was applied using the superset and subset analysis of Dusa (2019). This model calculates all possible recipes of causal conditions and assigns them scores for inclusion (quasi-consistency), relevance, and coverage. Hence it does not simply determine whether a condition is necessary and/or sufficient but also assesses its relevance and strength. In contrast to the traditional fsQCA approach, where a predefined range of causal conditions and context led typologies are enforced upon the method, superset/subset analysis follows an empirical path that identifies all possible solutions.

The traditional fsQCA approach can risk overlooking less prominent causal conditions. The use of superset/subset analysis reduces this risk and provides more accuracy. Calculating all possible solutions may require extensive computing power and while this was a considerable disadvantage back in 1987 when fsQCA was first introduced, subsequent improvements in technological capacity have alleviated this issue. Furthermore, the most common fsQCA software packages limited the number of possible causal conditions to a maximum of 20 or

30. The current software removed such limits, making it more likely that the result is under the proposed solution, thereby increasing the reliability of the analysis. This is specifically important as one of the major characteristics of fsQCA is the ability to generalize even when only a small to medium sample is used. The calculation of all possible solutions further strengthens this character. In addition, superset/subset analysis reduces the subjective bias of the researcher in the initial selection of the tested causal conditions and typologies.

In the face of the combined complexity of Lean and healthcare, all these methodological advantages become relevant as superset/subset analysis enables fsQCA to deliver stronger analysis and thereby improves understanding of the phenomena. In particular, the robustness, validity, and reliability of the analysis are improved.

10.3 LEAN CONFIGURATIONS ACHIEVING SUPERIOR PERFORMANCE

While several studies investigated Lean bundles in healthcare (e.g. Bucci et al., 2016, Costa and Godinho Filho, 2016, Hadid and Afshin Mansouri, 2014, Punnakitikashem, 2013), scant attention was paid to the importance of the specific practices in the bundles for achieving strong performance results.

This work has addressed this gap through necessity and sufficiency analysis via fsQCA. It has found that implementing process flow analysis is sufficient to effect lead time improvement and that in settings with staff who are resistant to change, implementing staff empowerment leads to superior lead time improvement. Furthermore, it is necessary to combine one of these two practices with the absence of process redesign, 5S, and standard practice and/or the presence of leadership support or mistake proofing in the configurations presented in the results section.

In additions, the analysis of configurations leading to superior quality found that under the condition of the presence of one of three practices (mistake proofing, staff empowerment, or standard work in progress) as well as the absence of the two practices of 5S and process flow analysis is enough to effect superior performance improvement. The previously discussed practices, when found in the configurations discussed in the results, become necessary for strong quality performance. They all contain standardisation practices (either standard work in progress or standard practice) to be present and some contain the absence of either process flow analysis or 5S. Additionally, other members of the configuration can be mistake proofing, leadership support, staff empowerment, kaizen huddle, and/or multifunctional employee.

These results allow Lean managers to focus resources on the practices that matter for the target performance indicators. They can avoid implementing vast numbers of practices and focus instead on a handful of performance drivers, splitting their attention between them according to the presented configurations. Furthermore, this focus will also increase the chance of implementation success as these practices/configurations were found to have the highest impact on the lead time and quality performance indicators.

10.4 SUBUNIT LENS IMPROVES LEAN

Most Lean studies investigated Lean through an organisational lens – viewing Lean bundles as operating within and impacting on the entire organisation. In the face of the operational heterogeneity and organisational multifunctionality of healthcare, this work opted to use a subunit lens – seeing Lean bundles as operating within a subunit and impacting on performance at subunit level, albeit with spill over effects and synergies for the entire organisation. It was assumed that the research would find full sets to be effective at subunit level. However, only individual practices were found to be sufficient and only OR configurations to be necessary. OR configurations mean that while all practices in the configuration can happen together, only one must occur. This finding is a strong indication of equifinality (i.e., several different paths may lead to one result) and implies that single practices may be both needed and enough to achieve superior performance. However, they are not the only way of achieving this since small set of practices with one that dominates might occur as well.

These findings provide evidence that the impacts of Lean practices are complex and context dependent. Lean operates on a subunit level in the form of individual practices or small sets, while larger sets are the norm at the organisational level. In consequence, single practices should become the focus of Lean implementations at subunit level with the caveat that if the context supports, small sets can be formed. However, the focus should still remain on one practice while using the other practices to support it.

This suggests that future Lean studies should investigate Lean through a multi-level lens, investigating how links between practices differ in and across subunits.

Furthermore, this work found several indications for the strong context dependence of the practice choice and its performance impact. Practices are chosen based on the motivation behind the implementation, the background and training of the KPO team, the organisation's history, and the staff's experience of improvement implementations, as well as trade-off decisions between performance indicators. This means that the relations between Lean practices as well as their impact on the performance indicators will differ widely according to the setting of the implementation.

While studies like (Galeazzo and Furlan, 2018) indicate that Lean's context dependence has been recognised, this study's findings highlight the importance of future studies of Lean addressing this issue. A possible way forward could be a standardized description of investigated implementations by answering the following questions:

- Which performance results were expected from the implementation?
- How do they prioritise performance indicators?
- Who trained the KPO team?
- How were the KPO team members trained?
- How were previous implementations received by staff?
- What are the functions, staff backgrounds, and size of the targeted subunits?

If these questions can be answered in future studies, there can be more effective comparisons between similar cases, improving understanding of how external factors influence the impact between Lean practices.

10.5 Overall limitations and future research

10.5.1 Overall limitations

This study was constrained by the medium-sized KPI data sample, the number of KPIs, and the KPI performance impacts being only positive, limiting the range of this work. While 41 hospital subunits in 7 care facilities provided a suitable size for fsQCA, the KPIs' performances were not measured equally by the organisations. This led to the forced exclusion of two KPIs. Furthermore, the limited KPI sample and the lack of coherent performance measurement across the subunits forced this study to investigate only lead time and quality. Future studies might test the findings with larger samples. Third, the subunit lens enabled an in-depth and greatly detailed understanding of Lean. However, tools like value stream mapping observe the entire value stream, crossing subunits and sometimes organisations. The subunit lens struggles with the consideration of spill-over effects and impacts of Lean practices which are below average but organisation-wide. While this can be partially compensated through a dense data availability and coverage like in this study, especially fresh Lean implementations rarely have that. The fourth limitation was that 97% of all performance improvements were positive – scarcely any negative changes were found in the collected data. The data was gathered from the KPO Teams responsible for improvement. While this might not particularly matter to this study because its general aim was to optimize the improvement method, it still takes away the possibility of learning from failed implementations. It has been acknowledged for some time that there is a dearth of studies into failed Lean implementations, but this knowledge has not sparked change. This study is also limited in that it is set in a highly specific service industry and its findings may not be generalisable to manufacturing industries.

10.5.2 Managerial implications

From the practitioner's perspective, the identified causal recipes provide strong insights into implementations of Lean practices, identifying essential conditions for superior performance. This is of interest to NHS Lean practitioners as it allows them to identify focus points for allocating time and resources. Since healthcare systems are proverbially battered by a lack of resources and staff, performance improvement measures that are more efficient and less resource-hungry will allow resources and staff to be freed up for problems elsewhere. In addition, this tight focus might improve the reputation of Lean and consequently reduce resistance to its implementation.

Most of all, this work has shown that practitioners should not follow a handbook implementation but instead carefully investigate the context and environment of the implementation and adopt their choice of Lean practice accordingly. The basic principle of Lean applies just as much to its own implementation – less can be more.

10.6 SUMMARY

In summary, this work promotes context-dependent combinations of Lean practices in healthcare subunits. The analysis found that individual practices, as well as small sets with one leading practice, tend to be enough to achieve superior lead time and quality improvement. Furthermore, the combined neo-configurational and subunit lenses open an entirely new perspective on Lean implementations in general and healthcare in particular.

11 REFERENCES

ABEND, G. 2013. The meaning of theory: sociological theory. San Francisco, CA: Berrett-Koehler Publishers.

ADLER, N. & DOCHERTY, P. 1998. Bringing business into sociotechnical theory and practice. *Human relations,* 51**,** 319-345.

ALEXANDER, J. W. & RANDOLPH, W. A. 1985. The fit between technology and structure as a predictor of performance in nursing subunits. *Academy of Management Journal,* 28**,** 844-859.

ALSMADI, M., ALMANI, A. & JERISAT, R. 2012. A comparative analysis of Lean practices and performance in the UK manufacturing and service sector firms. *Total Quality Management & Business Excellence,* 23**,** 381-396.

ALVESSON, M. 2003. Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research. *Academy of management review,* 28**,** 13-33.

ANDREWS, R., BOYNE, G. A. & WALKER, R. M. 2006. Subjective and objective measures of organizational performance: An empirical exploration. *Public service performance: Perspectives on measurement and management***,** 14-34.

ANELLIS, I. 2012. Peirce's Truth-functional Analysis and the Origin of the Truth Table. 33**,** 87-97.

ANUAR, A., SAAD, R. & YUSOFF, R. Z. 2018. Operational Performance and Lean Healthcare in the Healthcare Sector: Review on the Dimensions and Relationships. *International Journal of Academic Research in Business and Social Sciences,* 8**,** 276-292.

AOUN, M. 2015. The Mediating effect of Soft TQM on The Relationship between Lean Practices and Innovation Skills in Labanese Hospitals. *Universiti Utara Malaysia*.

ARTHUR, J. B. 1992. The link between business strategy and industrial relations systems in American steel minimills. *Industrial & Labor Relations Review,* 45**,** 488-506.

ASIF, M., JOOST DE BRUIJN, E., DOUGLAS, A. & FISSCHER, O. A. 2009. Why quality management programs fail: a strategic and operations management perspective. *International Journal of Quality & Reliability Management,* 26**,** 778-794.

BARTEZZAGHI, E. 1999. The evolution of production models: is a new paradigm emerging? *International Journal of Operations & Production Management,* 19**,** 229-250.

BAUMGARTNER, M. 2009. Inferring causal complexity. *J Sociological methods research* 38**,** 71-101.

BAUMGARTNER, M. & THIEM, A. 2017. Model ambiguities in configurational comparative research. *J Sociological Methods Research* 46**,** 954-987.

BAYOU, M. E. & DE KORVIN, A. 2008. Measuring the leanness of manufacturing systems—a case study of Ford Motor Company and General Motors. *Journal of Engineering and Technology Management,* 25**,** 287-304.

BEELMANN, A. 2006. *Review of Systematic reviews in the social sciences. A practical guide*, Hogrefe & Huber Publishers.

BEHROUZI, F. & WONG, K. Y. 2011. Lean performance evaluation of manufacturing systems: A dynamic and innovative approach. *Procedia Computer Science,* 3**,** 388-395.

BEN-TOVIM, D. I., BASSHAM, J. E., BENNETT, D. M., DOUGHERTY, M. L., MARTIN, M. A., O'NEILL, S. J., SINCOCK, J. L. & SZWARCBORD, M. G. 2008. Redesigning care at the Flinders Medical Centre: clinical process redesign using "lean thinking". *Medical Journal of Australia,* 188**,** S27-S31.

BICHENO, J. 2004. *The new lean toolbox: towards fast, flexible flow*, Production and Inventory Control, Systems and Industrial Engineering Books.

BJERKNES, G. & BRATTETEIG, T. 1995. User participation and democracy: A discussion of Scandinavian research on system development. *Scandinavian Journal of information systems,* 7**,** 1.

BONAVIA, T. & MARIN, J. A. 2006. An empirical study of lean production in the ceramic tile industry in Spain. *International Journal of Operations & Production Management,* 26**,** 505-531.

BORTOLOTTI, T., BOSCARI, S. & DANESE, P. 2015a. Successful lean implementation: Organizational culture and soft lean practices. *International Journal of Production Economics,* 160**,** 182-201.

BORTOLOTTI, T., DANESE, P., FLYNN, B. B. & ROMANO, P. 2015b. Leveraging fitness and lean bundles to build the cumulative performance sand cone model. *International Journal of Production Economics,* 162**,** 227-241.

BOURNE, M., NEELY, A., MILLS, J. & PLATTS, K. 2003. Implementing performance measurement systems: a literature review. *International Journal of Business Performance Management,* 5**,** 1-24.

BOWEN, D. E. & YOUNGDAHL, W. E. 1998. "Lean" service: in defense of a production-line approach. *International Journal of Service Industry Management,* 9**,** 207-225.

BRAITHWAITE, J., CHURRUCA, K., ELLIS, L. A., LONG, J., CLAY-WILLIAMS, R., DAMEN, N., HERKES, J., POMARE, C., LUDLOW, K. J. A., APPROACHES, APPLICATIONS & ACCOMPLISHMENTS. A WHITE PAPER. SYDNEY, A. A. I. O. H. I., MACQUARIE UNIVERSITY 2017. Complexity science in healthcare.

BRECH, E. F. L. & BRECH, E. 1975. *The principles and practice of management*, Addison-Wesley Longman Ltd.

BROD, M., TESLER, L. E. & CHRISTENSEN, T. L. 2009. Qualitative research and content validity: developing best practices based on science and experience. *Quality of Life Research,* 18**,** 1263.

BUCCI, S., DE BELVIS, A., MARVENTANO, S., DE LEVA, A., TANZARIELLO, M., SPECCHIA, M. L., RICCIARDI, W. & FRANCESCHI, F. 2016. Emergency Department crowding and hospital bed shortage: is Lean a smart answer? A systematic review. *Eur Rev Med Pharmacol Sci,* 20**,** 4209-19.

BURGESS, N. 2012. *Evaluating lean in healthcare.* PhD, University of Warwick.

BURNES, B. 2005. Complexity theories and organizational change. *International Journal of Management Reviews,* 7**,** 73-90.

BURRELL, G., & MORGAN, G. 2017. Sociological paradigms and organisational analysis: Elements of the sociology of corporate life. Routledge.

BURTON, R. M. & OBEL, B. 2004. *Strategic organizational diagnosis and design: The dynamics of fit*, Springer Science & Business Media.

BÜYÜKÖZKAN, G., ÇIFÇI, G. & GÜLERYÜZ, S. 2011. Strategic analysis of healthcare service quality using fuzzy AHP methodology. *Expert systems with applications,* 38**,** 9407-9424.

CARILLO, P. M., KOPELMAN, R. E. J. G. & STUDIES, O. 1991. Organization structure and productivity: Effects of subunit size, vertical complexity, and administrative intensity on operating efficiency. 16, 44-59.
CARMIEN, S., DAWE, M., FISCHER, G., GORMAN, A., KINTSCH, A. & SULLIVAN JR, J. F. 2005. Socio-technical environments supporting people with cognitive disabilities using public transportation. *ACM Transactions on Computer-Human Interaction (TOCHI),* 12, 233-262.
CHADHA, R., SINGH, A. & KALRA, J. 2012. Lean and queuing integration for the transformation of health care processes: A lean health care model. *Clinical Governance: An International Journal,* 17, 191-199.
CHIARINI, A. 2013. Waste savings in patient transportation inside large hospitals using lean thinking tools and logistic solutions. *Leadership in Health Services,* 26, 356-367.
COLBERT, B. A. 2004. The complex resource-based view: Implications for theory and practice in strategic human resource management. *Academy of Management Review,* 29, 341-358.
COMBS, J., LIU, Y., HALL, A. & KETCHEN, D. 2006. How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel psychology,* 59, 501-528.
CONNELLY, L. M. J. M. N. 2008. Pilot studies. 17, 411.
COOPER, B. & GLAESSER, J. 2016. Exploring the robustness of set theoretic findings from a large n fsQCA: an illustration from the sociology of education. *International Journal of Social Research Methodology,* 19, 445-459.
COOPER, D. R., SCHINDLER, P. S. & SUN, J. 2006. *Business research methods*, McGraw-Hill Irwin New York.
COSTA, L. B. M., FILHO, M. G., RENTES, A. F., BERTANI, T. M. & MARDEGAN, R. 2017. Lean healthcare in developing countries: evidence from Brazilian hospitals. *The International journal of health planning and management,* 32, e99-e120.
COSTA, L. B. M. & GODINHO FILHO, M. 2016. Lean healthcare: review, classification and analysis of literature. *Production Planning & Control*, 1-14.
CROTTY, M. F. 1998. The foundations of social research: Meaning and perspective in the research process. Sage.
CUA, K. O., MCKONE, K. E. & SCHROEDER, R. G. 2001. Relationships between implementation of TQM, JIT, and TPM and manufacturing performance. *Journal of operations management,* 19, 675-694.
DABHILKAR, M. & ÅHLSTRÖM, P. 2013. Converging production models: the STS versus lean production debate revisited. *International Journal of Operations & Production Management,* 33, 1019-1039.
DAL PONT, G., FURLAN, A. & VINELLI, A. 2008. Interrelationships among lean bundles and their effects on operational performance. *Operations Management Research,* 1, 150-158.
DE SOUZA, L. B. & PIDD, M. 2011. Exploring the barriers to lean health care implementation. *Public Money & Management,* 31, 59-66.
DELERY, J. E. & DOTY, D. H. 1996. Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of management Journal,* 39, 802-835.
DENNIS, P. 2016. *Lean Production simplified: A plain-language guide to the world's most powerful production system*, CRC Press.
DENZIN, N. K. 2017. *Sociological methods: A sourcebook*, Routledge.
DOOLEN, T. L. & HACKER, M. E. 2005. A review of lean assessment in organizations: an exploratory study of lean practices by electronics manufacturers. *Journal of Manufacturing systems,* 24, 55.
DOTY, D. H. & GLICK, W. H. 1994. Typologies as a unique form of theory building: Toward improved understanding and modeling. *Academy of management review,* 19, 230-251.
DOTY, D. H., GLICK, W. H. & HUBER, G. P. 1993. Fit, equifinality, and organizational effectiveness: A test of two configurational theories. *Academy of Management journal,* 36, 1196-1250.
DOW, D., SAMSON, D. & FORD, S. 1999. Exploding the myth: do all quality management practices contribute to superior quality performance? *Production and operations management,* 8, 1-27.

DUL, J. 2016a. Identifying single necessary conditions with NCA and fsQCA. *Journal of Business Research,* 69**,** 1516-1523.
DUL, J. J. O. R. M. 2016b. Necessary condition analysis (NCA) logic and methodology of “necessary but not sufficient” causality. 19**,** 10-52.
DUPRÉ, L. 1998. *Bugs in writing: a guide to debugging your prose*, Addison-Wesley Boston MA.
DUŞA, A. 2019. Set Theory. *QCA with R.* Springer.
DYER, J. H. & SINGH, H. 1998. The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *Academy of Management Review,* 23**,** 660-679.
D’ANDREAMATTEO, A., IANNI, L., LEGA, F. & SARGIACOMO, M. 2015. Lean in healthcare: A comprehensive review. *Health Policy,* 119**,** 1197-1209.
EHRENFELS, C. V. 1890. Über gestaltqualitäten. *Vierteljahrsschrift für wissenschaftliche Philosophie,* 14**,** 249-292.
ELBOURNE, E. T., HOME-MORTON, A. & MAUGHFLING, J. 1979. *Factory administration and accounts*, Arno Press.
EMERY, F. E. 1974. *Futures we're in*, Canberra: Centre for Continuing Education, Australian National University.
ERIKSSON, A., HOLDEN, R. J., WILLIAMSSON, A. & DELLVE, L. 2016. A Case Study of Three Swedish Hospitals’ Strategies for Implementing Lean Production. *Nordic Journal of Working Life Studies,* 6**,** 105-131.
EROGLU, C. & HOFER, C. 2011. Lean, leaner, too lean? The inventory-performance link revisited. *Journal of Operations Management,* 29**,** 356-369.
FANE, G. R., VAGHEFI, M. R., VAN DEUSEN, C. & WOODS, L. A. 2003. Competitive advantage the Toyota way. *Business Strategy Review,* 14**,** 51-60.
FELD, W. M. 2000. *Lean manufacturing: tools, techniques, and how to use them*, CRC Press.
FERREIRA, D. M. C., SAURIN, T. A. J. P. P. & CONTROL 2019. A complexity theory perspective of kaizen: a study in healthcare. 30**,** 1337-1353.
FINE, B. A., GOLDEN, B., HANNAM, R. & MORRA, D. 2009. Leading lean: a Canadian healthcare leader’s guide. *Healthcare Quarterly,* 12**,** 32-41.
FISS, P. C. 2007. A set-theoretic approach to organizational configurations. *Academy of management review,* 32**,** 1180-1198.
FISS, P. C. 2009. Practical issues in QCA. *Academy of Management*.
FISS, P. C. 2011. Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal,* 54**,** 393-420.
FISS, P. C., MARX, A. & CAMBRÉ, B. 2013a. CONFIGURATIONAL THEORY AND METHODS IN ORGANIZATIONAL RESEARCH: INTRODUCTION. *J Configurational Theory - Methods in Organizational Research,* 38**,** 1-22.
FISS, P. C., MARX, A. & RIHOUX, B. 2014. Comment: getting QCA right. *J Sociological Methodology,* 44**,** 95-100.
FISS, P. C., SHARAPOV, D. & CRONQVIST, L. 2013b. Opposites attract? Opportunities and challenges for integrating large-N QCA and econometric analysis. *Political Research Quarterly***,** 191-198.
FLYNN, B. B., SAKAKIBARA, S., SCHROEDER, R. G., BATES, K. A. & FLYNN, E. J. 1990. Empirical research methods in operations management. *Journal of operations management,* 9**,** 250-284.
FLYNN, B. B., SCHROEDER, R. G. & SAKAKIBARA, S. 1995. The impact of quality management practices on performance and competitive advantage. *Decision sciences,* 26**,** 659-691.
FOX, W. M. 1995. Sociotechnical system principles and guidelines: past and present. *The Journal of Applied Behavioral Science,* 31**,** 91-105.
FRENCH, S., MAULE, J. & PAPAMICHAIL, N. 2009. *Decision behaviour, analysis and support*, Cambridge University Press.
FUJIMOTO, T. 1999. *The evolution of a manufacturing system at Toyota*, Oxford university press.

FULLERTON, R. R., KENNEDY, F. A. & WIDENER, S. K. 2014. Lean manufacturing and firm performance: The incremental contribution of lean management accounting practices. *Journal of Operations Management,* 32**,** 414-428.
FULLERTON, R. R., MCWATTERS, C. S. & FAWSON, C. 2003. An examination of the relationships between JIT and financial performance. *Journal of Operations Management,* 21**,** 383-404.
FULLERTON, R. R. & WEMPE, W. F. 2009. Lean manufacturing, non-financial performance measures, and financial performance. *International Journal of Operations & Production Management,* 29**,** 214-240.
FURLAN, A., VINELLI, A. & DAL PONT, G. 2011. Complementarity and lean manufacturing bundles: an empirical analysis. *International Journal of Operations & Production Management,* 31**,** 835-850.
GALBRAITH, J. R. 1977. *Organization design*, Addison Wesley Publishing Company.
GALEAZZO, A. & FURLAN, A. 2018. Lean bundles and configurations: a fsQCA approach. *International Journal of Operations & Production Management,* 38**,** 513-533.
GALUNIC, D. C. & EISENHARDT, K. M. J. R. I. O. B. 1994. Renewing the strategy-structure-performance paradigm. 16**,** 215-215.
GERHART, B. & MILKOVICH, G. T. 1990. Organizational differences in managerial compensation and financial performance. *Academy of Management journal,* 33**,** 663-691.
GOODRICH, J. & CORNWELL, J. 2008. Seeing the person in the patient. *The King's Fund, London***,** 76.
GRECKHAMER, T. 2011. Cross-cultural differences in compensation level and inequality across occupations: A set-theoretic analysis. *Organization Studies,* 32**,** 85-115.
GRECKHAMER, T. 2016. CEO compensation in relation to worker compensation across countries: The configurational impact of country-level institutions. *Strategic Management Journal,* 37**,** 793-815.
GRESOV, C. & DRAZIN, R. J. A. O. M. R. 1997. Equifinality: Functional equivalence in organization design. 22**,** 403-428.
GROUT, J. R. & TOUSSAINT, J. S. 2010. Mistake-proofing healthcare: Why stopping processes may be a good start. *Business Horizons,* 53**,** 149-156.
GUEST, D. E. 2011. Human resource management and performance: still searching for some answers. *Human resource management journal,* 21**,** 3-13.
HADID, W. & AFSHIN MANSOURI, S. 2014. The lean-performance relationship in services: a theoretical model. *International Journal of Operations & Production Management,* 34**,** 750-785.
HADID, W., MANSOURI, S. A. & GALLEAR, D. 2016. Is lean service promising? A socio-technical perspective. *International Journal of Operations & Production Management,* 36**,** 618-642.
HARDCOPF, R. & SHAH, R. Lean and performance: The impact of organizational culture. Academy of Management Proceedings, 2014. Academy of Management Briarcliff Manor, NY 10510, 10747.
HENDERSON, A., SCHOONBEEK, S. & AUDITORE, A. J. N. M. 2013. Processes to engage and motivate staff. 20.
HINES, P., HOLWEG, M. & RICH, N. 2004. Learning to evolve: a review of contemporary lean thinking. *International journal of operations & production management,* 24**,** 994-1011.
HOFER, C., EROGLU, C. & HOFER, A. R. 2012. The effect of lean production on financial performance: The mediating role of inventory leanness. *International Journal of Production Economics,* 138**,** 242-253.
HOLDEN, R. J. 2011. Lean thinking in emergency departments: a critical review. *Annals of emergency medicine,* 57**,** 265-278.
HOLWEG, M. 2007. The genealogy of lean production. *Journal of operations management,* 25**,** 420-437.
HOMBURG, H. 1978. Anfänge des Taylorsystems in Deutschland vor dem ersten Weltkrieg. Eine Problemskizze unter besonderer Berücksichtigung der Arbeitskämpfe bei Bosch 1913. *Geschichte und Gesellschaft,* 4**,** 170-194.
HONOLD, L. (1997), "A review of the literature on employee empowerment", Empowerment in Organizations, Vol. 5 No. 4, pp. 202-212.

HOTHO, J. J. J. O. S. 2014. From typology to taxonomy: A configurational analysis of national business systems and their explanatory power. 35, 671-702.
HSIEH, H.-F. & SHANNON, S. E. 2005. Three approaches to qualitative content analysis. *Qualitative health research,* 15, 1277-1288.
HUBER, V. L. & BROWN, K. A. 1991. Human resource issues in cellular manufacturing: A sociotechnical analysis. *Journal of Operations Management,* 10, 138-159.
IMPROTA, G., ROMANO, M., DI CICCO, M. V., FERRARO, A., BORRELLI, A., VERDOLIVA, C., TRIASSI, M. & CESARELLI, M. 2018. Lean thinking to improve emergency department throughput at AORN Cardarelli hospital. *BMC health services research,* 18, 914.
JAIRATH, N., HOGERNEY, M. & PARSONS, C. J. A. N. R. 2000. The role of the pilot study: A case illustration from cardiac nursing research. 13, 92-96.
JECH, T. 2013. *Set theory*, Springer Science & Business Media.
JERVIS, R. 1998. *System effects: Complexity in political and social life*, Princeton University Press.
JOHNSON, J. E., SMITH, A. L. & MASTRO, K. A. 2012. From Toyota to the bedside: nurses can lead the lean way in health care reform. *Nursing administration quarterly,* 36, 234-242.
JOHNSON, M., BURGESS, N. and SETHI, S., 2020. Temporal pacing of outcomes for improving patient flow: design science research in a National Health Service Hospital. *Journal of Operations Management*, 66(1-2), pp.35-53.
JOOSTEN, T., BONGERS, I. & JANSSEN, R. 2009. Application of lean thinking to health care: issues and observations. *International Journal for Quality in Health Care,* 21, 341-347.
KANE, M., CHUI, K., RIMICCI, J., CALLAGY, P., HEREFORD, J., SHEN, S., NORRIS, R. & PICKHAM, D. J. J. T. J. O. N. A. 2015. Lean manufacturing improves emergency department throughput and patient satisfaction. 45, 429-434.
KARLSSON, C. & ÅHLSTRÖM, P. 1996. Assessing changes towards lean production. *International Journal of Operations & Production Management,* 16, 24-41.
KATZ, D. & KAHN, R. L. 1978. *The social psychology of organizations*, Wiley New York.
KAUSHIK, V., & WALSH, C. A. 2019. Pragmatism as a research paradigm and its implications for social work research. Social Sciences, 8(9), 255.
KEENEY, R. L. 1996. *Value-focused thinking*, Harvard University Press.
KIM, C. S., SPAHLINGER, D. A., KIN, J. M., COFFEY, R. J. & BILLI, J. E. 2009. Implementation of lean thinking: one health system's journey. *The Joint Commission Journal on Quality and Patient Safety,* 35, 406-AP4.
KING, D. L., BEN-TOVIM, D. I. & BASSHAM, J. 2006. Redesigning emergency department patient flows: application of lean thinking to health care. *Emergency Medicine Australasia,* 18, 391-397.
KING, N. & HORROCKS, C. 2010. *Interviews in qualitative research*, Sage.
KLIR, G. & YUAN, B. 1995. *Fuzzy sets and fuzzy logic*, Prentice hall New Jersey.
KOEHLER, D. J., BRENNER, L. & GRIFFIN, D. 2002. The calibration of expert judgment: Heuristics and biases beyond the laboratory. *Heuristics and biases: The psychology of intuitive judgment*, 686-715.
KOGUT, B. & RAGIN, C. 2006. Exploring complexity when diversity is limited: Institutional complementarity in theories of rule of law and national systems revisited. *European Management Review,* 3, 44-59.
KRAFCIK, J. F. 1988. Triumph of the lean production system. *MIT Sloan Management Review,* 30, 41.
KROGSLUND, C., CHOI, D. D. & POERTNER, M. J. P. A. 2015. Fuzzy sets on shaky ground: Parameter sensitivity and confirmation bias in fsQCA. 23, 21-41.
KUHN, T. 1970. The structure of scientific revolutions. Princeton University Press.
KVALE, S. 2008. *Doing interviews*, Sage.
LACEY, R. & COHEN, L. 2015. Using Qualitative Comparative Analysis (QCA) as a Descriptive Numerical Method in Support of Narrative Methods. *Handbook of Qualitative Organizational Research: Innovative Pathways and Methods*, 362.

LAMOTHE, L. & DUFOUR, Y. 2007. Systems of interdependency and core orchestrating themes at health care unit level: a configurational approach. *Public Management Review,* 9**,** 67-85.
LANDER, E. & LIKER, J. K. 2007. The Toyota Production System and art: making highly customized and creative products the Toyota way. *International Journal of Production Research,* 45**,** 3681-3698.
LAUREANI, A., ANTONY, J. & DOUGLAS, A. 2010. Lean six sigma in a call centre: a case study. *International journal of productivity and performance management,* 59**,** 757-768.
LEATT, P. & SCHNECK, R. 1982. Work environments of different types of nursing subunits. *J Adv Nurs,* 7**,** 581-94.
LEPAK, D. P. & SHAW, J. D. 2008. Strategic HRM in North America: looking to the future. *The International Journal of Human Resource Management,* 19**,** 1486-1499.
LEWIS, M. A. 2000. Lean production and sustainable competitive advantage. *International Journal of Operations & Production Management,* 20**,** 959-978.
LIKER, J. 2003. *The Toyota Way : 14 Management Principles from the World's Greatest Manufacturer: 14 Management Principles from the World's Greatest Manufacturer*, Mcgraw-hill.
LIKER, J. K. & MORGAN, J. M. 2006. The Toyota way in services: the case of lean product development. *The Academy of Management Perspectives,* 20**,** 5-20.
LORDEN, A. L., LIN, S.-H. & COTE, M. J. 2015. Measures of success: The role of human factors in lean implementation in healthcare. *Quality control and applied statistics,* 60**,** 127-130.
MACDUFFIE, J. P. 1995. Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. *Industrial & labor relations review,* 48**,** 197-221.
MALMBRANDT, M. & ÅHLSTRÖM, P. 2013. An instrument for assessing lean service adoption. *International Journal of Operations & Production Management,* 33**,** 1131-1165.
MANNON, M. 2014. Lean healthcare and quality management: The experience of ThedaCare. *The Quality Management Journal,* 21**,** 7.
MANOS, A., SATTLER, M. & ALUKAL, G. 2006. Make healthcare lean. *Quality progress,* 39**,** 24.
MANZ, C. C. & STEWART, G. L. 1997. Attaining flexible stability by integrating total quality management and socio-technical systems theory. *Organization Science,* 8**,** 59-70.
MARCH, J. G. 2006. Rationality, foolishness, and adaptive intelligence. *Strategic management journal,* 27**,** 201-214.
MARLER, J. H. 2012. Strategic human resource management in context: a historical and global perspective. *The Academy of Management Perspectives,* 26**,** 6-11.
MARTÍN-ALCÁZAR, F., ROMERO-FERNANDEZ, P. M. & SÁNCHEZ-GARDEY, G. 2005. Strategic human resource management: integrating the universalistic, contingent, configurational and contextual perspectives. *The International Journal of Human Resource Management,* 16**,** 633-659.
MATSUI, Y. 2007. An empirical analysis of just-in-time production in Japanese manufacturing companies. *International Journal of Production Economics,* 108**,** 153-164.
MAYRING, P. 2004. Qualitative content analysis. *A companion to qualitative research,* 1**,** 159-176.
MAZUR, L., MCCREERY, J. & ROTHENBERG, L. 2012. Facilitating lean learning and behaviors in hospitals during the early stages of lean implementation. *J Engineering Management Journal,* 24**,** 11-22.
MAZZOCATO, P., SAVAGE, C., BROMMELS, M., ARONSSON, H. & THOR, J. 2010. Lean thinking in healthcare: a realist review of the literature. *Quality and Safety in Health Care,* 19**,** 376-382.
MCCRACKEN, G. 1988. *The long interview*, Sage.
MCDERMOTT, C. M. & STOCK, G. N. 1999. Organizational culture and advanced manufacturing technology implementation. *Journal of Operations Management,* 17**,** 521-533.
MCDERMOTT, C. M. & VENDITTI, F. J. 2015. Implementing lean in knowledge work: Implications from a study of the hospital discharge planning process. *Operations Management Research,* 8**,** 118-130.
MCKONE, K. E., SCHROEDER, R. G. & CUA, K. O. 2001. The impact of total productive maintenance practices on manufacturing performance. *Journal of operations management,* 19**,** 39-58.
MELMAN, S. 1958. *Decision-making and Productivity*, Wiley.

METTERS, R. & VARGAS, V. 2000. A typology of de-coupling strategies in mixed services. *Journal of Operations Management,* 18**,** 663-682.
MEYER, A. D., TSUI, A. S. & HININGS, C. R. 1993. Configurational approaches to organizational analysis. *Academy of Management journal,* 36**,** 1175-1195.
MI DAHLGAARD-PARK, S. & PETTERSEN, J. 2009. Defining lean production: some conceptual and practical issues. *The TQM Journal,* 21**,** 127-142.
MILES, R. E., SNOW, C. C., MEYER, A. D. & COLEMAN, H. J. 1978. Organizational strategy, structure, and process. *Academy of management review,* 3**,** 546-562.
MILLER, D. 1996. Configurations revisited. *Strategic management journal***,** 505-512.
MILLER, D. & FRIESEN, P. H. 1978. Archetypes of strategy formulation. *J Management science,* 24**,** 921-933.
MINTZBERG, H. 1979. *The structuring of organization: a synthesis of the research*, Prentice-Hall.
MISANGYI, V. F., GRECKHAMER, T., FURNARI, S., FISS, P. C., CRILLY, D. & AGUILERA, R. 2017. Embracing Causal Complexity. *Journal of Management,* 43**,** 255-282.
MITLETON-KELLY, E. 2003. *Complex systems and evolutionary perspectives on organisations: the application of complexity theory to organisations*, Elsevier Science Ltd.
MONDEN, Y. 1983. *Toyota production system: practical approach to production management*, Engineering & Management Press.
MONDEN, Y. 2011. *Toyota production system: an integrated approach to just-in-time*, CRC Press.
MORGAN, D. 2014. Integrating qualitative and quantitative methods: A pragmatic approach. Thousand Oaks: Sage. Soc. Sci.
MORTON, A. & CORNWELL, J. 2009. What's the difference between a hospital and a bottling factory? *BMJ: British Medical Journal (Online),* 339.
MUFFATTO, M. 1999. Evolution of production paradigms: the Toyota and Volvo cases. *Integrated Manufacturing Systems,* 10**,** 15-25.
MUMFORD, E. 2000. Socio-technical design: An unfulfilled promise or a future opportunity? *Organizational and social perspectives on information technology.* Springer.
MUMFORD, E. & MACDONALD, W. B. 1989. *XSEL'S progress: the continuing journey of an expert system*, John Wiley & Sons, Inc.
NARAYANAMURTHY, G. & GURUMURTHY, A. 2018. Is the hospital lean? A mathematical model for assessing the implementation of lean thinking in healthcare institutions. *Operations Research for Health Care,* 18**,** 84-98.
NEUHAUSER, D. M. 2011. *Impact of staff engagement on nurse satisfaction/retention and patient outcomes of patient satisfaction and NDNQI® indicators.* Western Carolina University.
NEUMAN, L. W. 2002. Social research methods: Qualitative and quantitative approaches.
NEWTON, N. 2010. The use of semi-structured interviews in qualitative research: strengths and weaknesses. *Exploring qualitative methods,* 1**,** 1-11.
NG, D., VAIL, G., THOMAS, S. & SCHMIDT, N. 2010. Applying the Lean principles of the Toyota Production System to reduce wait times in the emergency department. *Canadian Journal of Emergency Medicine,* 12**,** 50-57.
NICHOLAS, J. & SONI, A. 2005. *The portal to lean production: Principles and practices for doing more with less*, CRC Press.
NIEPCEL, W. & MOLLEMAN, E. 1998. Work design issues in lean production from a sociotechnical systems perspective: Neo-Taylorism or the next step in sociotechnical design? *Human relations,* 51**,** 259-287.
NØRBY, M., ARLBJØRN, J. S., NORLYK, B., WIBORG, K., HOLM, N. & SIMONSEN, R. 2008. *Lean uden grænser?*, Academica.
OHNO, T. 1988a. Just-in-time for Today and Tomorrow. Taiichi Ohno with Setsuo Mito; translated by Joseph P. Schmelzeis. Productivity Press, Cambridge, MA.
OHNO, T. 1988b. *Toyota production system: beyond large-scale production*, crc Press.

OLIVER, N., DELBRIDGE, R. & LOWE, J. 1996. The European auto components industry: manufacturing performance and practice. *International Journal of Operations & Production Management,* 16**,** 85-97.
OSTERMAN, P. 1994. How common is workplace transformation and who adopts it? *Industrial & Labor Relations Review,* 47**,** 173-188.
OYEMOMI, O., LIU, S., NEAGA, I. & ALKHURAIJI, A. 2016. How knowledge sharing and business process contribute to organizational performance: Using the fsQCA approach. *Journal of Business Research,* 69**,** 5222-5227.
PAKDIL, F. & LEONARD, K. M. 2014. Criteria for a lean organisation: development of a lean assessment tool. *International Journal of Production Research,* 52**,** 4587-4607.
PANIZZOLO, R. 1998. Applying the lessons learned from 27 lean manufacturers.: The relevance of relationships management. *International journal of production economics,* 55**,** 223-240.
PANSIRI, J. 2005. Pragmatism: A methodological approach to researching strategic alliances in tourism. Tourism and Hospitality Planning & Development, 2(3), 191-206.
PAPADOPOULOS, T. & MERALI, Y. 2008. Stakeholder network dynamics and emergent trajectories of Lean implementation projects: a study in the UK National Health Service. *Public Money and Management,* 28**,** 41-48.
PAPADOPOULOS, T., RADNOR, Z. & MERALI, Y. 2011. The role of actor associations in understanding the implementation of Lean thinking in healthcare. *International Journal of Operations & Production Management,* 31**,** 167-191.
PATTON, M. Q. 2002. Qualitative evaluation methods. 3rd Edition. Thousand Oaks: Sage Publications.
PFEFFER, J. 1994. *Competitive advantage through people: Unleashing the power of the work force*, Harvard Business Press.
PIERCY, N. & RICH, N. 2009. Lean transformation in the pure service environment: the case of the call service centre. *International journal of operations & production management,* 29**,** 54-76.
POKSINSKA, B., SWARTLING, D. & DROTZ, E. 2013. The daily work of Lean leaders–lessons from manufacturing and healthcare. *Total Quality Management & Business Excellence,* 24**,** 886-898.
PORTER, M. E. 2011. *Competitive advantage of nations: creating and sustaining superior performance*, Simon and Schuster.
PRAJOGO, D. I. 2005. The comparative analysis of TQM practices and quality performance between manufacturing and service firms. *International Journal of Service Industry Management,* 16**,** 217-228.
PUNNAKITIKASHEM, P. Determining Lean Practices in Health Care Service Systems. The 19th International Conference on Industrial Engineering and Engineering Management, 2013. Springer, 945-954.
QURESHI, Z. H. A review of accident modelling approaches for complex socio-technical systems. Proceedings of the twelfth Australian workshop on Safety critical systems and software and safety-related programmable systems-Volume 86, 2007. Australian Computer Society, Inc., 47-59.
RAAB, S. S., ANDREW-JAJA, C., CONDEL, J. L. & DABBS, D. J. 2006. Improving Papanicolaou test quality and reducing medical errors by using Toyota production system methods. *American journal of obstetrics and gynecology,* 194**,** 57-64.
RADNOR, Z. 2011. Implementing lean in health care: making the link between the approach, readiness and sustainability. *J International Journal of Industrial Engineering Management*

2**,** 1-12.
RADNOR, Z. J. 2009. Implementing lean in health care: making the link between the approach, readiness and sustainability.
RADNOR, Z. J., HOLWEG, M. & WARING, J. 2012. Lean in healthcare: the unfilled promise? *Social science & medicine,* 74**,** 364-371.
RAGIN, C. C. 2000. *Fuzzy-set social science*, University of Chicago Press.

RAGIN, C. C. 2005. From fuzzy sets to crisp truth tables. Compasss Working Paper.
RAGIN, C. C. 2008. *Redesigning social inquiry: Fuzzy sets and beyond*, Wiley Online Library.
RAGIN, C. C. 2014. *The comparative method: Moving beyond qualitative and quantitative strategies*, Univ of California Press.
RAGIN, C. C. & SONNETT, J. 2005. Between complexity and parsimony: Limited diversity, counterfactual cases, and comparative analysis. *Vergleichen in der Politikwissenschaft.* Springer.
RAGIN, C. C. J. P. A. 2006. Set relations in social research: Evaluating their consistency and coverage. 14**,** 291-310.
RAHANI, A. & AL-ASHRAF, M. 2012. Production flow analysis through value stream mapping: a lean manufacturing process case study. *J Procedia Engineering,* 41**,** 1727-1734.
REIMANN, F., KOSMOL, T. & KAUFMANN, L. 2017. Responses to Supplier-Induced Disruptions: A Fuzzy-Set Analysis. *Journal of Supply Chain Management*.
REINERTSEN, D. 2005. Let it flow: how lean product development sparked a revolution. *Industrial Engineer,* 37**,** 40-46.
REMLER, D. K. & VAN RYZIN, G. G. 2010. *Research methods in practice: Strategies for description and causation*, Sage Publications.
REN, S., TSAI, H.-T. & EISINGERICH, A. B. 2016. Case-based asymmetric modeling of firms with high versus low outcomes in implementing changes in direction. *Journal of Business Research,* 69**,** 500-507.
RIBEIRO, R. A. 1996. Fuzzy multiple attribute decision making: a review and new preference elicitation techniques. *Fuzzy sets and systems,* 78**,** 155-181.
RICE, A. K. 1953. Productivity and social organization in an Indian weaving shed: An examination of some aspects of the socio-technical system of an experimental automatic loom shed. *Human Relations,* 6**,** 297-329.
RICE, A. K. 2013. *Productivity and social organization: The Ahmedabad experiment: Technical innovation, work organization and management*, Routledge.
RIHOUX, B. & LOBE, B. 2009. The case for qualitative comparative analysis (QCA): Adding leverage for thick cross-case comparison. *The Sage handbook of case-based methods***,** 222-242.
RIHOUX, B. & RAGIN, C. C. 2008. *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*, Sage Publications.
RIVKIN, J. W. 2000. Imitation of complex strategies. *Management science,* 46**,** 824-844.
ROHWER, A., SCHOONEES, A. & YOUNG, T. 2014. Methods used and lessons learnt in conducting document reviews of medical and allied health curricula–a key step in curriculum evaluation. *BMC medical education,* 14**,** 236.
ROLLIER, M. 1979. Taylorism and the Italian unions. *The Quality of Working Life in Western and Eastern Europe*.
SAMNANI, A. K. & SINGH, P. 2013. Exploring the fit perspective: an ethnographic approach. *Human Resource Management,* 52**,** 123-144.
SAMPSON, S. E. & FROEHLE, C. M. 2006. Foundations and implications of a proposed unified services theory. *Production and operations management,* 15**,** 329-343.
SAMSON, D. & TERZIOVSKI, M. 1999. The relationship between total quality management practices and operational performance. *Journal of operations management,* 17**,** 393-409.
SANTOS, J. N., MOTA, J. & BAPTISTA, C. S. 2017. Understanding configurations of value creation functions in business relationships using a fuzzy-set QCA. *Journal of Business Research*.
SAUNDERS, M., LEWIS, P., & THORNHILL, A. 2009. Research methods for business students. Pearson education.
SAURIN, T. A., MARODIN, G. A. & RIBEIRO, J. L. D. 2011. A framework for assessing the use of lean production practices in manufacturing cells. *International Journal of Production Research,* 49**,** 3211-3230.
SAURIN, T. A., ROOKE, J. & KOSKELA, L. J. I. J. O. P. R. 2013. A complex systems theory perspective of lean production. 51**,** 5824-5838.

SCHLOSSER, D., DE MEUR, G., RIHOUX, B. & RAGIN, C. 2009. Qualitative Comparative Analysis as an Approach. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. London: Sage*, 1-18.
SCHNEIDER, C. Q. & ROHLFING, I. 2016. Case studies nested in fuzzy-set QCA on sufficiency: formalizing case selection and causal inference. *Sociological Methods & Research,* 45, 526-568.
SCHNEIDER, C. Q. & WAGEMANN, C. 2012. *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*, Cambridge University Press.
SCHNEIDER, C. Q. & WAGEMANN, C. J. C. S. 2010. Qualitative comparative analysis (QCA) and fuzzy-sets: Agenda for a research approach and a data analysis technique. 9, 376-396.
SCHONBERGER, R. 1982. *Japanese manufacturing techniques: Nine hidden lessons in simplicity*, Simon and Schuster.
SCHOONHOVEN, M., KEMPER, B. P., BRILLEMAN, M. I. & DOES, R. J. 2011. Quality quandaries: Streamlining the path to optimal care for cardiovascular patients. *Quality Engineering,* 23, 388-394.
SHAH, R. 2002. *A configurational view of lean manufacturing and its theoretical implications.* The Ohio State University.
SHAH, R. & WARD, P. T. 2003. Lean manufacturing: context, practice bundles, and performance. *Journal of operations management,* 21, 129-149.
SHAH, R. & WARD, P. T. 2007. Defining and developing measures of lean production. *Journal of operations management,* 25, 785-805.
SHIMADA, H. & MACDUFFIE, J. P. 1986. Industrial relations and" humanware"--Japanese investments in automobile manufacturing in the United States.
SHINGO, S. & DILLON, A. P. 1989. *A study of the Toyota production system: From an Industrial Engineering Viewpoint*, CRC Press.
SIMPSON, L. 1998. *Supporting decision analysis: a pragmatic approach.* University of Leeds.
SINGLETON JR, R. A., STRAITS, B. C. & STRAITS, M. M. 1993. *Approaches to social research*, Oxford University Press.
SLOAN, T., FITZGERALD, A., HAYES, K. J., RADNOR, Z., ROBINSON, S., SOHAL, A., DROTZ, E. & POKSINSKA, B. 2014. Lean in healthcare from employees' perspectives. *Journal of health organization and management*.
STACE, D. A. 1996. Dominant ideologies, strategic change, and sustained performance. *Human Relations,* 49, 553-570.
STEED, A. 2012. An exploration of the leadership attributes and methods associated with successful lean system deployments in acute care hospitals. *J Quality Management in Healthcare,* 21, 48-58.
STENTOFT ARLBJØRN, J. & VAGN FREYTAG, P. 2013. Evidence of lean: a review of international peer-reviewed journal articles. *European Business Review,* 25, 174-205.
SUÁREZ-BARRAZA, M. F., SMITH, T. & DAHLGAARD-PARK, S. M. 2012. Lean Service: A literature analysis and classification. *Total Quality Management & Business Excellence,* 23, 359-380.
SWANK, C. K. 2003. The lean service machine. *Harvard business review,* 81, 123-130.
SWANSON, R. A. 1999. The foundations of performance improvement and implications for practice. *Advances in Developing Human Resources,* 1, 1-25.
TALIB, F., RAHMAN, Z. & QURESHI, M. 2013. An empirical investigation of relationship between total quality management practices and quality performance in Indian service companies. *International Journal of Quality & Reliability Management,* 30, 280-318.
TAYLOR, F. W. 1919. *The Principles of Scientific Management,* [S.l.], Harper and Brothers Publishers.
TAYLOR, W. A. & WRIGHT, G. H. 2006. The contribution of measurement and information infrastructure to TQM success. *Omega,* 34, 372-384.
TEICHGRÄBER, U. K. & DE BUCOURT, M. 2012. Applying value stream mapping techniques to eliminate non-value-added waste for the procurement of endovascular stents. *J European journal of radiology,* 81, e47-e52.

THABANE, L., MA, J., CHU, R., CHENG, J., ISMAILA, A., RIOS, L. P., ROBSON, R., THABANE, M., GIANGREGORIO, L. & GOLDSMITH, C. H. 2010. A tutorial on pilot studies: the what, why and how. *BMC medical research methodology,* 10**,** 1.
THIEM, A. 2016. Analyzing multilevel data with QCA: yet another straightforward procedure. *J Quality Quantity*

50**,** 121-128.
TIMMERMANS, S. & BERG, M. 2003. *The gold standard: the challenge of evidence-based medicine and standardization in health care* Temple University Press.
TRANFIELD, D., DENYER, D. & SMART, P. 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management,* 14**,** 207-222.
TREPO, G. 1979. Improvement of working conditions and job design in France. *The Quality of Working Life in Western and Eastern Europe*.
TRIST, E. 1981. The evolution of socio-technical systems. *Occasional paper,* 2**,** 1981.
TRIST, E. & MURRAY, H. 1990. *The Social Engagement of Social Science, Volume 1: A Tavistock Anthology: The Socio-Psychological Perspective*, University of Pennsylvania Press.
TRIST, E. L. & BAMFORTH, K. W. 1951. Some social and psychological consequences of the Longwall Method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human relations,* 4**,** 3-38.
VAGHEFI, M. R., WOODS, L. A. & HUELLMANTEL, A. 2000. Toyota story 2: Still winning the productivity game. *Business Strategy Review,* 11**,** 59-70.
VAN BEINUM, H. J. J. 1963. *Een organisatie in beweging.*
VANWYNSBERGHE, R. & KHAN, S. 2007. Redefining case study. *International Journal of Qualitative Methods,* 6**,** 80-94.
VENKATRAMAN, N. 1989. The concept of fit in strategy research: Toward verbal and statistical correspondence. *Academy of management review,* 14**,** 423-444.
VIS, B. 2012. The comparative advantages of fsQCA and regression analysis for moderately large-N analyses. *Sociological Methods & Research,* 41**,** 168-198.
WHETTEN, D. A., FELIN, T. & KING, B. G. 2009. The practice of theory borrowing in organizational studies: Current issues and future directions. *Journal of Management,* 35**,** 537-563.
WHITE, M., WELLS, J. & BUTTERWORTH, T. 2013. Leadership, a key element of quality improvement in healthcare. Results from a literature review of "Lean Healthcare" and the Productive Ward. *J The International Journal of Leadership in Public Services*.
WHITESITT, J. E. 2012. *Boolean algebra and its applications*, Courier Corporation.
WHITTINGTON, R., PETTIGREW, A., PECK, S., FENTON, E. & CONYON, M. J. O. S. 1999. Change and complementarities in the new competitive landscape: A European panel study, 1992–1996. 10**,** 583-600.
WILSON, J., FARRINGTON-DARBY, T., COX, G., BYE, R. & HOCKEY, G. R. J. 2007. The railway as a socio-technical system: human factors at the heart of successful rail engineering. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,* 221**,** 101-115.
WILSON, J. M. & MCKINLAY, A. J. B. H. 2010. Rethinking the assembly line: Organisation, performance and productivity in Ford Motor Company, c. 1908–27. 52**,** 760-778.
WITTGENSTEIN, L. 2013. *Tractatus logico-philosophicus*, Routledge.
WOMACK, J. P. & JONES, D. T. 2010. *Lean thinking: banish waste and create wealth in your corporation*, Simon and Schuster.
WOMACK, J. P., JONES, D. T. & ROOS, D. 1990. *Machine that changed the world*, Simon and Schuster.

12 Appendix

A1 – Lean practices of Shah and Ward (2003)

JIT	TPM	TQM	HRM
Lot-size reductions	Predictive or preventive maintenance	Competitive benchmarking	Self-directed work team
Continuous-flow production	Maintenance optimisation	Quality management programs	Flexible, cross-functional workforce
Pull system	Safety improvement programs	Total Quality Management	
Cellular manufacturing	Planning and scheduling strategies	Process capability measurements	
Cycle-time reductions	New process equipment or technologies	Formal continuous improvements	
Focused-factory productions systems			
Agile manufacturing strategies			
Quick changeover techniques			
Bottleneck/constraint removal			
Re-engineered production processes			

A2 – Socio-Technical practices according to Hadid et al. (2014) and Hadid et al. (2016)

No.	Technical practices	References
1.	5Ss	Ehrlich (2006), Holden (2010), Poksinska (2010), Arlbjørn *et al.* (2011), Burgess and Radnor (2010), Manos *et al.* (2006), Fillingham (2007), Esain *et al.* (2008), Emiliani (2004), Bushell *et al.* (2002), Suarez Barraza *et al.* (2009), Wayne (2005), Brewton (2009), Tiplady (2010), Finigan and Humphries (2006), Maguad (2007), Julien and Tjahjono (2009), Haque and James-Moore (2004), Keen (2011), Pedersen and Huniche (2011), Wenchao Song *et al.* (2009), Kaplan and Patterson (2008), Markovitz (2012), Chadha *et al.* (2012), Radnor *et al.* (2012), Schulze and Störmer (2012)
2.	A3 report	Holden (2010), Jimmerson *et al.* (2005), Doman (2011), Qudrat-Ullah *et al.* (2012)
3.	Automation	Holden (2010), Poksinska (2010), Bortolotti and Romano (2010), Manos *et al.* (2006), Ahluwalia *et al.* (2004), Lodge and Bamford (2008), Wayne (2005), Julien and Tjahjono (2009), Åhlström (2004), Carter *et al.* (2011), Wenchao Song *et al.* (2009), Bortolotti and Romano (2012)
4.	Change management	Manos *et al.* (2006)
5.	Continuous improvement	Dickson *et al.* (2009), Ehrlich (2006), Poksinska (2010), Piercy and Rich (2009a), Manos *et al.* (2006), Emiliani (2004), Alagaraja (2010), Maguad (2007), Kuriger *et al.* (2010), Yavas and Yasin (2001), Hagan (2011), Qudrat-Ullah *et al.* (2012)
6.	Eliminating loop-backs	Swank (2003)
7.	Group technology	Piercy and Rich (2009b), Ehrlich (2006), Holden (2010), Swank (2003), Arlbjørn *et al.* (2011), Nielsen and Edwards (2010), Burgess and Radnor (2010), Manos *et al.* (2006), Arbos (2002), Ben-Tovim *et al.* (2007), Alagaraja (2010), Hyer and Wemmerlöv (2002), Tatikonda (2007), Cuatrecasas (2004), Middleton *et al.* (2005)
8.	Changing the facility layout	Allway and Corbett (2002), Holden (2010), Poksinska (2010), Holm and Ahlstrom (2010a), Manos *et al.* (2006), Tonya (2004), Cuatrecasas (2004), Nelson-Peterson and Leppa (2007)
9.	Just in Time	Cooper and Mohabeersingh (2008a, b), Holden (2010), Poksinska (2010), Arlbjørn *et al.* (2011), Manos *et al.* (2006), Emiliani (2004), Alagaraja (2010), Åhlström (2004), Cuatrecasas (2004), Nelson-Peterson and Leppa (2007), Chadha *et al.* (2012)
10.	Kaizen blitz	Dickson *et al.* (2009), Holden (2010), Arlbjørn *et al.* (2011), Burgess and Radnor (2010), Hines and Lethbridge (2008), Suarez-Barraza and Ramis-Pujol (2010), Suarez Barraza *et al.* (2009), Kress (2008), Papadopoulos and Merali (2008), Montabon (2005), Graban and Swartz (2012), Papadopoulos (2012), Radnor *et al.* (2012)
11.	Kanban	Holden (2010), Poksinska (2010), Arlbjørn *et al.* (2011), Manos *et al.* (2006), Alagaraja (2010), Reinertsen (2005), Hagan (2011), Nelson-Peterson and Leppa (2007)

No.	Technical practices	References
12.	Mistakes proofing/Poka-Yoke	Ehrlich (2006), Cooper and Mohabeersingh (2008a, b), Holden (2010), Poksinska (2010), Manos *et al.* (2006), Alagaraja (2010), Finigan and Humphries (2006), Maguad (2007), Kuriger *et al.* (2010), Mirehei *et al.* (2011), Hagan (2011), Doman (2011)
13.	Model cell, roll out	Swank (2003), Graban and Swartz (2012)
14.	Outsourcing	Comm and Mathaisel (2005b)
15.	Point of use storage	Manos *et al.* (2006)
16.	Policy deployment/ Hoshin Kanri	Poksinska (2010), Swank (2003), Emiliani (2004), Alagaraja (2010), Pejsa and Eng (2011), Ball and Maleyeff (2003), Wayne (2005), Qudrat-Ullah *et al.* (2012)
17.	Process redesign	Piercy and Rich (2009a), McQuade (2008), Suarez-Barraza and Ramis-Pujol (2010), Yavas and Yasin (2001), Carter *et al.* (2011), Edwards *et al.* (2012), Chadha *et al.* (2012), Bortolotti and Romano (2012)
18.	Production levelling/Heijunka	Poksinska (2010), Emiliani (2004), Staats *et al.* (2011), Pedersen and Huniche (2011)
19.	Pull system	Ehrlich (2006), Cooper and Mohabeersingh (2008a, b), Poksinska (2010), Holm and Ahlstrom (2010a), Arlbjørn *et al.* (2011), Manos *et al.* (2006), Petersen and Wohlin (2010), Kuriger *et al.* (2010), Reinertsen and Shaeffer (2005), Reinertsen (2005), Kress (2008), Mirehei *et al.* (2011), Hagan (2011), Ball and Maleyeff (2003), Schulze and Störmer (2012)
20.	Quality circles	Swank (2003), Searcy (2009b)
21.	Quality function deployment	Emiliani (2004), Alagaraja (2010), Tatikonda (2007), Wang *et al.* (2012), Schulze and Störmer (2012)
22.	Quick set up time	Arlbjørn *et al.* (2011), Manos *et al.* (2006), Arbos (2002), Finigan and Humphries (2006), Maguad (2007)
23.	Root cause analysis	Ehrlich (2006), Holden (2010), Poksinska (2010), Jones *et al.* (1999), Petersen and Wohlin (2010), Searcy (2009b), Villarreal *et al.* (2009), Haque And James-Moore (2004), Yavas and Yasin (2001), Wang *et al.* (2012), Collar *et al.* (2012), Schulze and Störmer (2012)
24.	Segregating complexity	Holm and Ahlstrom (2010a), Swank (2003), Nielsen and Edwards (2010), King *et al.* (2006)
25.	Self-inspection	Manos *et al.* (2006), Maguad (2007)
26.	Simplification	Bortolotti and Romano (2010), Bortolotti and Romano (2012)
27.	Single piece flow	Poksinska (2010), Staats *et al.* (2011), Alagaraja (2010), Kuriger *et al.* (2010), Haque and James-Moore (2004), Kress (2008), Mirehei *et al.* (2011), Nelson-Peterson and Leppa (2007), Chadha *et al.* (2012), Bortolotti and Romano (2012)
28.	Small lots	Ehrlich (2006), Swank (2003), Manos *et al.* (2006), Arbos (2002), Brewton (2009), Kuriger *et al.* (2010), Reinertsen and Shaeffer (2005), Reinertsen (2005), Kress (2008)

No.	Technical practices	References
29.	Standardization	Allway and Corbett (2002), Ehrlich (2006), Holden (2010), Poksinska (2010), Holm and Ahlstrom (2010a), Sprigg and Jackson (2006), Swank (2003), Bortolotti and Romano (2010), Kosuge *et al.* (2010), Nielsen and Edwards (2010), Manos *et al.* (2006), Emiliani (2004), Bushell *et al.* (2002), Staats *et al.* (2011), Alagaraja (2010), Haque And James-Moore (2004), LaGanga (2011), Hagan (2011), Wenchao Song *et al.* (2009), Kaplan and Patterson (2008), Nelson-Peterson and Leppa (2007), Middleton *et al.* (2005), Doman (2011), Carlborg *et al.* (2013), Wang *et al.* (2012), Qudrat-Ullah *et al.* (2012), Chadha *et al.* (2012), Bortolotti and Romano (2012), Jaca *et al.* (2012)
30.	Takt time	Allway and Corbett (2002), Poksinska (2010), Holm and Ahlstrom (2010a), Swank (2003), Arlbjørn *et al.* (2011), Arbos (2002), Emiliani (2004), Haque and James-Moore (2004), Reinertsen (2005), Kress (2008), Cuatrecasas (2004), Middleton *et al.* (2005)
31.	Total preventive maintenance	Poksinska (2010), Arlbjørn *et al.* (2011), Manos *et al.* (2006), Arbos (2002), Emiliani (2004), Finigan and Humphries (2006), Maguad (2007), Åhlström (2004)
32.	Total quality	Kuriger *et al.* (2010), Mirehei *et al.* (2011)
33.	Use of new technologies	Jones *et al.* (1999), Hines and Lethbridge (2008), Comm and Mathaisel (2005b), Tischler (2006)
34.	Value stream mapping	Dickson *et al.* (2009), Piercy and Rich (2009b), Ehrlich (2006), Holden (2010), Poksinska (2010), Arlbjørn *et al.* (2011), Bortolotti and Romano (2010), Piercy and Rich (2009a), Nielsen and Edwards (2010), Burgess and Radnor (2010), King *et al.* (2006), Jimmerson *et al.* (2005), Ahluwalia *et al.* (2004), Fillingham (2007), Jones *et al.* (1999), Lodge and Bamford (2008), Bushell *et al.* (2002), McQuade (2008), Suarez-Barraza and Ramis-Pujol (2010), Hines *et al.* (2008), Ben-Tovim *et al.* (2007), Suarez Barraza *et al.* (2009), Staats *et al.* (2011), Alagaraja (2010), Tonya (2004), Wayne (2005), Searcy (2009b), Tiplady (2010), Maguad (2007), Villarreal *et al.* (2009), Julien and Tjahjono (2009), Haque And James-Moore (2004), Keen (2011), Kress (2008), LaGanga (2011), Pedersen and Huniche (2011), Papadopoulos and Merali (2008), Wenchao Song *et al.* (2009), Tischler (2006), Chaneski (2005), Doman (2011), Wang *et al.* (2012), Chadha *et al.* (2012), Bortolotti and Romano (2012), Schulze and Störmer (2012), Vlachos and Bogdanovic (2013)
35.	Vertical information system	Holm and Ahlstrom (2010a), Åhlström (2004)
36.	Visualization	Holden (2010), Poksinska (2010), Arlbjørn *et al.* (2011), Manos *et al.* (2006), Fillingham (2007), Emiliani (2004), Bushell *et al.* (2002), Staats *et al.* (2011), Alagaraja (2010), Wayne (2005), Brewton (2009), Finigan and Humphries (2006), Haque And James-Moore (2004), Keen (2011), Wenchao Song *et al.* (2009), Tischler (2006), Kaplan and Patterson (2008), Nelson-Peterson and Leppa (2007)
37.	Work load balancing	Swank (2003), Brewton (2009), Kuriger *et al.* (2010), Cuatrecasas (2004), Mirehei *et al.* (2011), Wenchao Song *et al.* (2009), Middleton *et al.* (2005)

No.	Supportive practices	References
1.	An appropriate rewarding system	Piercy and Rich (2009b), Ehrlich (2006), Holden (2010), Wayne (2005), Jaca *et al.* (2012)
2.	Customer involvement	Holm and Ahlstrom (2010a, b), Suarez-Barraza and Ramis-Pujol (2010)
3.	Effective Communication System	Allway and Corbett (2002), Holden (2010), Swank (2003), Manos *et al.* (2006), Hines and Lethbridge (2008), Suarez-Barraza and Ramis-Pujol (2010), Hines *et al.* (2008), Comm and Mathaisel (2005a), Pejsa and Eng (2011), Jaca *et al.* (2012)
4.	Employee empowerment	Holden (2010), Jones *et al.* (1999), Comm and Mathaisel (2005a), Graban and Swartz (2012), deHaan *et al.* (2012), Bortolotti and Romano (2012), Collar *et al.* (2012)
5.	Employees commitment	Dickson *et al.* (2009), Poksinska (2010), Carter *et al.* (2011), Bortolotti and Romano (2012), Schulze and Störmer (2012)
6.	Employees involvement	Piercy and Rich (2009b), Ehrlich (2006), Holden (2010), Swank (2003), Bortolotti and Romano (2010), Manos *et al.* (2006), Suarez-Barraza and Ramis-Pujol (2010), Hines *et al.* (2008), Tonya (2004), Julien and Tjahjono (2009), Kress (2008), Graban and Swartz (2012), deHaan *et al.* (2012), Bortolotti and Romano (2012), Collar *et al.* (2012), Schulze and Störmer (2012), Jaca *et al.* (2012)
7.	Establishing a long-term relation with suppliers	Swank (2003), Wang *et al.* (2012), Qudrat-Ullah *et al.* (2012)
8.	Establishing environment for change	Comm and Mathaisel (2005a), Graban and Swartz (2012)
9.	Having multifunctional employees	Dickson *et al.* (2009), Ehrlich (2006), Arbos (2002), Tonya (2004), Moayed and Shell (2009), Cuatrecasas (2004), LaGanga (2011), Chadha *et al.* (2012)
10.	Improving teamwork spirit	Suarez-Barraza and Ramis-Pujol (2010), Graban and Swartz (2012), Jaca *et al.* (2012)
11.	Leadership	Allway and Corbett (2002), Swank (2003), Suarez-Barraza and Ramis-Pujol (2010), Comm and Mathaisel (2005a), Jaaron and Backhouse (2011), Keen (2011), Qudrat-Ullah *et al.* (2012), Schulze and Störmer (2012)
12.	Modifying the terminology to suit services	Hines *et al.* (2008)
13.	Obtaining management support	Dickson *et al.* (2009), Allway and Corbett (2002), Piercy and Rich (2009b), Poksinska (2010), Holm and Ahlstrom (2010a), Swank (2003), Bortolotti and Romano (2010), Piercy and Rich (2009a), Burgess and Radnor (2010), King *et al.* (2006), Jimmerson *et al.* (2005), Hines and Lethbridge (2008), Suarez-Barraza and Ramis-Pujol (2010), Tischler (2006), Towne (2006), Graban and Swartz (2012), Papadopoulos (2012), Jaca *et al.* (2012)

No.	Supportive practices	References
14.	Performance measurement system	Piercy and Rich (2009b), Ehrlich (2006), Swank (2003), Burgess and Radnor (2010), Suarez-Barraza and Ramis-Pujol (2010), Bhasin (2008), Comm and Mathaisel (2005a), Comm and Mathaisel (2005b), Kennedy *et al.* (2007), Kress (2008), Bortolotti and Romano (2012)
15.	Posting performance results	Swank (2003), Middleton *et al.* (2005)
16.	Providing justifications for implementing the practices	Jaaron and Backhouse (2011)
17.	Training	Dickson *et al.* (2009), Piercy and Rich (2009b), Ehrlich (2006), Holden (2010), Poksinska (2010), Holm and Ahlstrom (2010a), Piercy and Rich (2009a), Burgess and Radnor (2010), Manos *et al.* (2006), King *et al.* (2006), Jimmerson *et al.* (2005), Suarez-Barraza and Ramis-Pujol (2010), Hines *et al.* (2008), Staats *et al.* (2011), Comm and Mathaisel (2005a), Comm and Mathaisel (2005b), Tonya (2004), Wayne (2005), Searcy (2009b), Keen (2011), Kress (2008), Cuatrecasas (2004), Mirehei *et al.* (2011), Carter *et al.* (2011), Tischler (2006), Graban and Swartz (2012), Schulze and Störmer (2012), Jaca *et al.* (2012)

A3 – SOCIAL AND TECHNICAL FACTORS ACCORDING TO HADID ET AL. (2016)

Practices
Reward system
Communication system
Management support
Performance measurement system
Training
Employee empowerment
Employee commitment
Employee involvement
Leadership

Lean social factors (Hadid et al., 2016)

Practices
Automation
Just in time
Pull system
Work load balancing
Quick setup time
Small lots
5Ss
Group technology
Improving facility layout
Visualisation
Kaizen blitz
Policy deployment/Hoshin Kanri
Quality function deployment
Value stream mapping
Root cause analysis
Total preventive maintenance

Lean technical factors (Hadid et al., 2016)

A4 – PRACTICE COUNT LITERARY REVIEW

Practice	Amount
Value Stream Mapping	71
Standardisation	55
Process/patient flow analysis	47
5S	36
RPIW / RIE	34
Visualisation	32
Integrated process redesign	32
DMAIC	28
Change facility layout	25
Root cause analysis / Ishikawa	20
Team based problem solving	19
Training	19
Kanban/Pull system	13
Gemba Walk	13
Continuous improvement	12
Employee involvement	12
Single piece flow	11
Performance measurement system	10
A3 report	9
Kaizen Blitz	9
Employee empowerment	9
Mistake proofing/poka yoke	8
Obtain management support	8
PDSA	7
ERP and medical record system	7
5 Whys	6
Employee commitment	6
Work load balancing	5
JIT	5
Having a multifunctional employee	5
Spaghetti diagram	4

Improving teamwork spirit	4
Huddle	4
Andon / Stop the line	3
Change management	3
Group tech.	3
Self-inspection	3
Effective communication system	3
Establishing environment for change	3
Lean for Leaders	3
Jidoka/Automation	2
Production levelling/Heijunka	2
Model cell, roll out	2
Quality assurance program	2
Segregating complexity	2
Simplification	2
Customer involvement	2
HR/Change of staff allocation	2
Eliminate loop-backs	1
Point of use storage	1
Policy deployment/Hoshin Kanri	1
Quality circles	1
Quick set up time	1
Small lots	1
Takt time	1
Total preventive maintenance	1
Establishing a long-term relation to suppliers	1
Posting performance results	1

A5 – Practice bundles of lit. review

5 Practices - bundle

3x RPIW + Visualisation + Standardisation + process flow analysis + Kanban

2x Integrated process redesign + Visualisation + Standardisation + 5S + Process flow analysis

4 Practices - bundle

5x 5S + VSM + Standardisation + Integrated process redesign

5x Training + VSM + process flow analysis + Standardisation

4x RPIW + Visualisation + Standardisation + process flow analysis

4x 5S + Training + VSM + Standardisation

3x VSM + Standardisation + Visualisation + Team based problem-solving

2x Process/patient flow analysis + VSM + Training + employee empowerment

3 Practices - bundle

14x 5S + VSM + Standardisation

10x RPIW + Visualisation + Standardisation

8x VSM + process flow analysis + Standardisation

7x Integrated process redesign + Visualisation + Standardisation

6x VSM + Standardisation + Training

6x Integrated process redesign + VSM + Standardisation

6x Integrated process redesign + DMAIC + Process flow analysis

5x 5S + Training + Standardisation

5x Root cause analysis + DMAIC + VSM

5x Change facility layout + VSM + Visualisation

5x Obtaining management support + Training + VSM

4x VSM + Single piece flow + Standardisation

4x 5S + Team based problem solving + Visualisation

3x Multifunctional employee + 5S + Standardisation

3x Continuous improvement + process flow analysis + integrated process redesign

2x JIT + Standardisation + Kanban/Pullsystem

2x VSM + Process/patient flow analysis + continuous improvement

2x 5S + DMAIC + Integrated process redesign

2x Multifunctional employee + VSM + Integrated process redesign

Main Author	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
Abuhejleh 2016														x	x																														x	x	x				x	x			x		
Al-Araidah 2010	x	x								x								x					x																																				
Al-Owad 2014			x								x						x																																										
Aronsson 2011																x																																										
Atkinson 2012						x					x			x					x																																							
Balle 2007		x	x															x																														x								x		
Barnas 2011		x	x					x	x		x	x				x		x	x					x																																	
BenTovim 2008		x				x									x	x		x	x		x							x																														
Bhat 2013		x							x	x				x	x			x						x																																	
Bhat 2014									x	x					x	x		x		x				x																																	
Bhat 2014		x							x	x				x	x			x					x	x																																	
Bhat 2016		x		x	x					x					x			x			x		x								x																				x			x				
Black 2009									x	x																																																
Bucort 2011															x	x																																									
Burgess 2010		x														x											x			x																								x	x		x		
Bushell 2002		x										x				x		x																																							
Cankovik 2009		x										x					x				x																																				
Carboneau 2010									x	x						x																																									
Carter 2012	x		x						x		x					x								x																									x										
Casey 2009											x					x						x																																			
Castle 2009																x							x	x																																	
Cerfolio 2019																x																																										
Chadha 2012		x					x									x	x	x			x								x																						x							
Chiarini 2012	x								x	x						x																																									
Chiarini 2013					x											x					x		x																																		
Cima 2011										x						x																		x												x		x					x				x	
Collar 2012	x																				x																																				
Cooper 2008a																				x									x																													
Cooper 2008b																				x									x																													
Costa 2017		x				x				x	x	x		x	x	x		x			x	x		x						x																		x						x		x		
de Souza 2011		x		x							x	x								x	x																																				
Deans 2011		x									x					x																																									
Dickson 2009 a				x							x	x			x	x					x		x																																		
Dickson 2009 c																x														x									x							x												
Dickson 2009b											x				x	x		x					x																																		
Edwards 2012				x														x																																							
Eller 2009		x										x				x		x																																							
Eriksson 2017									x	x					x	x		x																																		x				x		
Esain 2008		x																																																							
Fache 2009										x																																																
Fillingham 2007		x										x				x																																									
Fosdick and Uphoff (2007)				x							x				x						x		x																				x			x	x	x						x				x
Gayed 2013												x				x	x	x																																							
Gijo 2013									x	x					x			x						x																																	
Grout 2010								x	x		x	x	x					x		x	x																	x																			
Grove 2010 a		x														x																																									
Grove 2010 b				x			x									x		x			x				x														x								x										
Guimaraes 2013															x								x																																		
Hagan 2011							x							x				x		x																																					
Heitmiller 2010				x					x	x		x																																											x		
Hintzen 2009		x										x				x		x					x																														x			x		
Improta 2018		x										x			x	x		x																																x						x		
Isaac 2012		x		x				x			x					x		x					x							x																		x										
Johnson 2012		x									x			x		x		x			x		x																																		
Johnson 2013			x	x															x					x																																	
Kane 2015		x					x				x				x	x		x			x			x																				x			x		x					x			x		
Kaplan 2008		x										x						x																																							
Kim 2009		x														x		x			x																																				
Kim, Hayman, Billi, Lash, and Lawrence (2007)							x									x	x	x																										x					x									
Kim, Spahlinger, Kin, and Billi (2006)							x			x					x	x																		x																								
Kimsey 2010			x				x				x	x						x	x		x		x	x																																	
King 2006																x																					x																x			x		
Kullar 2010				x										x																																										
L'hommedieu 2010																x																																										
LaGanga (2011)				x							x				x						x																																				
Laing and Baumgartner (2005)																							x																																x		
Laureani 2013		x				x			x	x					x	x				x																																					
Leeuwen 2010										x	x	x			x						x																																				
Leslie, Hagood, Royer, Reece, and Maloney (2006)											x				x		x	x			x																																				
Lodge and Bamford (2008)																x																																							x		
Lummus, Vokurka, and Rodeghiero (2006)															x	x																														x		x											
Lunardini 2014				x										x																																										
Manos 2006		x					x					x	x	x				x		x			x			x		x	x			x				x		x		x		x			x			x								x		
Martin 2013																x						x																																		
Mazur and Chen (2008)												x			x	x																																					x			x	
Mazzocato 2012				x					x		x	x		x	x			x					x							x																x											
McClean, Young, and Bustard (2008)															x	x																																									
McDermott 2015											x				x	x		x																																					x		
Melanson 2009				x							x				x			x	x											x													x							x							
Migita 2011							x								x	x																																									x	
Miller 2015	x								x					x	x	x					x																																				
Molla 2018										x					x	x																											x							x							
Murrel 2011							x								x	x					x					x																		x			x	x	x										
Naidoo 2016	x	x	x								x					x																									x															x		
Naik 2011			x								x	x		x	x	x		x																																							
Nelson-Peterson 2007												x		x			x	x					x						x															x		x								x				
Ng 2010											x	x				x		x					x																																		
Nicholas 2012		x			x						x					x		x					x																																		
Nielsen 2010																x		x										x								x																					
Papadopoulos 2008																x														x																											
Papadopoulos 2011				x							x	x			x			x					x		x																																	
Papadopoulos 2011											x	x						x																																			x					
Papadopoulos 2012											x	x						x			x									x																											
Parks, Klein, Frankel, Friese, and Shafi (2008)									x	x					x						x																																		x	x	
Pejsa 2011																																	x											x													
Persoon, Zaleski, and Frerichs (2006)															x	x	x				x																														x			x		x		
Pocha 2010										x																																																
Raab et al. (2008)							x										x																		x			x									x	x										
Raab, Andrew-JaJa, Condel, and Dabbs (2006)																	x	x			x																																	x				
Radnor 2012		x									x				x						x				x					x																													
Rico 2013											x					x		x																																							
Schoonhoven										x					x																																										
Schoonhoven 2011				x						x						x		x																																							
Simon 2012									x		x					x	x	x						x																																	
Sloan 2014		x		x			x				x	x			x	x		x								x																					x	x	x			x	x	x			x		
Smith 2012		x									x					x			x					x																																	
Snyder 2009												x				x							x																																		
Sprigg 2006																		x																																							
Stanton 2014										x		x									x																															x	x	x	x	x	
Stonemetz 2011										x	x				x																																										
Tagge 2017									x	x						x				x																																x		x		
Tejedor 2014																x					x																																				
Toda 2019		x			x									x	x			x			x			x																																	
Towne 2006																															x																										
Ulhassan 2013		x		x								x				x	x						x																																		
Van Vliet 2011															x																																										
Vegting 2012												x						x																																							
Waldhausen 2010		x				x					x				x			x			x																																				
Waring 2010															x											x																							x							
White 2014		x					x					x			x			x			x		x																																		x	x
Wikma 2009										x																																																
Wojtys 2009																x		x																																							
Yeh 2011									x	x					x	x		x																																							
Yousri 2011																x																												x													
Yusof 2012			x												x					x																																				
Zhu 2018									x					x	x	x		x					x																																	x	

A7 – Practice = Number

#	Practice	30	Kaizen Blitz
1	5 Whys	31	model cell, roll out
2	5S	32	point of use storage
3	A3 report	33	policy deployment / Hoshin Kanri
4	Teambased problem solving	34	Quality circles
5	Spaghetti Diagram	35	Quality assurance program
6	Work load balancing	36	Quick set up time
7	Contiuous improvement	37	Segregating complexity
8	Andon / Stop the line	38	Self-inspection
9	Root cause analysis / Ishikawa	39	Simplification
10	DMAIC	40	Small lots
11	RPIW / RIE	41	Takt time
12	Visualization	42	Total preventive maintenance
13	Jidoka/Automation	43	Huddle
14	Kanban/Pull system	44	customer involvment
15	Process/patient flow analysis	45	effect. Comm syst.
16	Value Stream Mapping	46	Employee empowerment
17	Single piece flow	47	employees commitment
18	Standardization	48	employees involvement
19	PDSA	49	establishing a long-term relation to suppliers
20	Mistake proofing/poka-yoke	50	establishing environment for change
21	Integrated process redesign	51	having a multifunctional employee
22	Production levelling / Heijunka	52	Lean for Leaders
23	Change facility layout	53	Obtain management support
24	Gemba Walk	54	Performance measurement system
25	improving teamwork spirit	55	posting performance results
26	Change management	56	Training
27	Eliminate loop-backs	57	Introduction/ modification of Information system - Combined ERP and Medical record system
28	Group tech.	58	HR / Change of staff allocation
29	JIT		

A8 – Interview Protocol

RQ	Parts	Key references	Additions	Interviewee	Questions
Which combination of organisational factors and lean practices results in a superior (or not superior) performance outcome in healthcare organisations?	Lean practices	Shah & Ward (2003, 2007) Hadid et al. (2014, 2016)	Lean bundles	Lean Officer (Middle/Exec manager)	**Primary Question:** How do you progress with the Lean implementation? **Essential questions:** o Which Lean practices are applied? **Optional questions:** o Where practices were implemented or are planned to be implemented? o Why are you planning to implement them? o Which processes or techniques are you planning to improve? o How are you planning to implement the practices? o What was the start date of the Lean implementation (hospital/trust/subunit levels)? o When do you expect to finish the core implementation? o Which other key persons play a major role in the Lean implementation in your hospital? o Did any events occur that required a change of the Lean practices? If yes, what were they? o Which practices are influencing which subunits/wards? o How would you assess the implementation?

					○ What has changed so far?
	Performance indicators	Andrews et al. (2015)	Objective	Chief Information Officer	**Primary Question:** How do you measure performance in the lean wards? **Essential questions:** ○ Which standardized performance variables do you collect on a subunit/ward level? (For example: readmission rates) **Optional questions:** ○ If differences occur between subunit-performance data, which variables are collected on which wards? ○ Which anonymised measurement scores do you create on subunit level? ○ If differences occur, which measures are used in which wards? ○ What did you do if different measures showed different results? ○ How would you asses the Lean implementation? ○ Which changes were visible in the data?
Which elements are necessary and/or sufficient for	Configurational theory & QCA	Ragin et al. (2008)	Level of influence / Degree of membership	Lean Officer (Middle/Exec manager)	**Primary Question:** How did you perceive the impact of Lean practices? **Essential questions:** ○ Which factors do you think favourites the performance influence of the Lean implementation?

a higher impact on performance of the bundle?					**Optional questions:** o Which practices were used most? o Which ones did you perceive least relevant for the improvement (in which cases)? o Which ones did you perceive most relevant? o Which practices did you spend most time with? o Which practices did you spend least time with?