



Impact of lean interventions on time buffer reduction in a hospital setting

Oskar P. Roemeling, Martin J. Land, Kees Ahaus, Jannes Slomp & Wouter van den Bijllaardt

To cite this article: Oskar P. Roemeling, Martin J. Land, Kees Ahaus, Jannes Slomp & Wouter van den Bijllaardt (2017) Impact of lean interventions on time buffer reduction in a hospital setting, International Journal of Production Research, 55:16, 4802-4815, DOI: [10.1080/00207543.2017.1301687](https://doi.org/10.1080/00207543.2017.1301687)

To link to this article: <https://doi.org/10.1080/00207543.2017.1301687>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 13 Mar 2017.



Submit your article to this journal [↗](#)



Article views: 2752



View related articles [↗](#)




View Crossmark data [↗](#)



Citing articles: 8 View citing articles [↗](#)

Impact of lean interventions on time buffer reduction in a hospital setting

Oskar P. Roemeling^{a*}, Martin J. Land^a , Kees Ahaus^a, Jannes Slomp^b and Wouter van den Bijllaardt^c

^aFaculty of Economics and Business, University of Groningen, Groningen, The Netherlands; ^bFaculty of Technology, HAN University of Applied Science, Arnhem, The Netherlands; ^cSt. Elisabeth Hospital, Tilburg, The Netherlands

(Received 14 October 2015; accepted 19 February 2017)

This paper focuses on performance changes stemming from a series of lean interventions in a medical laboratory. This research is one of the first to link a series of lean interventions and performance over time. In a mixed-method case study, six years of patient-related throughput data, retrieved from a laboratory computer database, are analysed. Three distinct periods with significant differences in throughput time performance can be distinguished. Semi-structured interviews were held to investigate the lean interventions preceding the performance changes. Given the long-term nature of the study, the event history calendar method was applied to enhance the respondents' recall and reliability. A single lean intervention, among the hundreds that took place, was supposed to cause the main reduction in throughput times. It concentrated on improving process flow through the removal of batching, a source of artificial variability. A later major intervention, the introduction of flow-focused machinery, had mixed effects and initial performance gains were not sustained. The results show that ongoing series of interventions do not always lead to ongoing performance improvements in terms of throughput times but support theories emphasising the importance of variability reduction.

Keywords: lean manufacturing; Kaizen; leanness; performance analysis; Toyota production system; health care

Introduction

This paper investigates the relationship between the long-term application of lean principles and throughput time performance in a health care setting. We will focus on the process flow of a medical laboratory and study performance changes over six years. Health care is faced with increasing demands for excellent performance, the objective being to deliver increased quality while reducing costs. One of the approaches to performance improvement in health care is adopting a lean philosophy (Burgess and Radnor 2013; Clark, Silvester, and Knowles 2013). Holweg (2007) notes that the lean philosophy has gained increasing attention since its popularisation by Womack, Jones, and Roos (1990). Over the years, lean has been shown to be a facilitator of performance improvements in industry (Belekoukias, Garza-Reyes, and Kumar 2014; Khanchanapong et al. 2014; Chavez et al. 2015). This increased attention for lean in research has resulted in a growing amount of scientific studies aimed at lean applications in health care (Young and McClean 2008; Poksinska 2010; Holden 2011). Despite the promising outcomes (D'Angelo and Zarbo 2007; Serrano et al. 2010), there is still a need for more empirical research focused on the application of lean principles (Jasti and Kodali 2015).

Current studies investigating lean and performance in health care tend to focus on the short-term effects of lean applications. This is problematic, since Marodin and Saurin (2013) report that the adoption of lean in other sectors, such as health care, requires an effective performance of lean in those sectors. According to Mazzocato et al. (2010), the research field has yet to reach maturity. Additionally, Poksinska (2010) comments that there is only limited evidence of the complete lean philosophy being applied in health care. Rather, the concept is perceived as a set of tools and techniques aiming at process improvement. Based on the studies included in the review by Mazzocato et al. (2010), it seems that most reports of successful lean implementations in health care focus on single event outcomes with lean tools aimed at reducing waste.

While using tools that reduce waste can have a large impact on performance, it is not the only goal of the lean approach. Targeting direct (or obvious) waste can be seen as the first step in lean implementation, but one might also seek to reduce variability. The reduction in variability is of major importance, since variability has consequences for both patient service measures and health care costs (Helm, AhmadBeygi, and Van Oyen 2011; Salzarulo et al. 2011). In the longer term, the reduction in variability allows buffers to be minimised (Hopp and Spearman 2004). In hospitals, there would particularly be an interest in reducing time buffers since these are perceived by patients as undesirable

*Corresponding author. Email: o.p.roemeling@rug.nl

waiting time and may also increase health care costs due to longer stays in a hospital. Despite lean having received attention for a number of years, the actual effects of long-term adherence to the lean philosophy remain unidentified. Whilst there appears to be a trade-off between short- and long-term effects of improvement programmes (e.g. Sterman, Repenning, and Kofman 1997), the cumulative effect of series of lean interventions is unknown. Therefore, our main research question is: *how do series of lean interventions influence time buffer performance over time?*

To investigate the relationship between lean and time buffer performance, we have conducted an in-depth case study in a health care environment. For our study, we make use of data spanning a six-year period. We visualise lean performance over time, and show the effects of adhering to lean principles with the aid of statistical tests. In-depth interviews help identify the specific root causes that result in performance improvement. Our study is one of the first studies to investigate the effects of series of lean interventions on performance over such a lengthy period based on real-life data. We use triangulation, combining both quantitative and qualitative data, to strengthen the reliability of our results.

The structure of the paper is as follows. The first section introduces our theoretical framework. In the second section, the methodology is presented, and the case setting and chosen methods are outlined. Then, in our third section, covering our analysis and results, we analyse the changes in performance over time and investigate performance changes through in-depth interviews. The paper concludes with a discussion of our findings and the implications for future research. Ultimately, our study should help create a better understanding, both within academia and management, of the effects of prolonged adherence to lean principles in a health care setting.

Theoretical background

A substantial research base reports positive lean effects in various health care settings (Ben-Tovim et al. 2007; Shah et al. 2008; Dickson et al. 2009). For example, lean techniques have been successfully applied to improve a histopathology department (Raab et al. 2008). In health care, studies commonly use various lean definitions with a focus on the reduction of waste or the creation of value (e.g. Laing and Baumgartner 2005; Fillingham 2007). These definitions are then used as a theoretical basis for the approach adopted. Additionally, some applied studies (e.g. Raab et al. 2006; Grigg, Garrett, and Miller 2009) report on the usage of tools without providing a clear theoretical basis. In short, there seems to be no generally accepted lean theory.

Shah and Ward (2007) underline the importance of variability within lean with their definition of lean as an integrated socio-technical system whose main objective is to eliminate waste by concurrently reducing or minimising supplier, customer and internal variability. This reduction in variability is obtained through continuously improving processes (Radnor, Holweg, and Waring 2011). Still, despite variability clearly being an important aspect of lean, health care-oriented studies often focus on direct waste. To go beyond the focus on waste, and help build our theoretical lean framework, we adopt the description of lean by Hopp and Spearman (2004) and Hopp (2008).

From Hopp and Spearman's perspective, apart from reducing so-called 'direct' waste, reducing variability is of major importance. Typical for direct waste is that it is visible or obvious. Contrary to variability, direct waste is not hidden by buffers. According to Hopp and Spearman (2004), when Toyota implemented lean, they first switched the focus from inventory buffers, which 'hide' undesired variability, to capacity buffers. Then, by removing sources of variability through a continuous improvement process, both inventory and capacity buffers could be reduced. Ultimately, this makes an organisation lean since this should be accomplished with minimal buffering costs (Hopp 2008). Hopp (2008) distinguishes four phases of lean implementation: (1) target direct waste, (2) exchange inventory for capacity buffers, (3) reduce variability and (4) reduce capacity. Furthermore, one should cycle between phases three and four as part of continuous improvement initiatives. Whilst testing the existence of these phases in practice is difficult, they do provide a frame of reference to argue that an organisation with lean experience should focus on more than just direct waste and show attention to buffers and variability.

While we expect inventory and capacity buffers to be present in health care organisations, we do not see inventory appearing in the core health care process of providing patient care (Jack and Powers 2004). Since patients are themselves transformed in the health care process, it is impossible to stock the transformed resource, the patients. When a queue of patients or patients' samples occurs, this should be seen, in terms of Hopp (2008), as the use of a time buffer (the customer requirement waits) and not as an inventory buffer. As such, coping with variability in health care is mainly achieved through capacity and time buffers. This also means that we expect time buffers to fulfil the roles associated with inventory buffers, since both inventory and time buffers hide variability. Time buffers in health care most commonly relate to patients waiting for either diagnosis or treatment. Whilst capacity buffers relate to different capacity resources (such as staff or MRI machinery) waiting for jobs.

While we expect throughput time reductions to result from adhering to a lean philosophy, the impact of adhering to the lean philosophy in practice will only result from lean interventions. A lean intervention incorporates changes in

work routines at a specific moment in time. These changes should relate to either direct waste, or to variability and related buffering requirements. Even when variability is a logical consequence of customer-specific operations, which certainly applies to health care settings, lean approaches could still seek to minimise the buffers needed to deal with this variability (Thürer et al. 2014). Especially for interventions in the mature phases of lean adoption, we would expect a relationship between interventions and time buffer performance, i.e. decreased throughput times. In other words, the focus on lean interventions allows us to investigate the relationship between *actual* lean practices and their outcomes in terms of *actual* changes in performance. The method applied to investigate the relationship between lean interventions and the time buffer is elaborated in the next section.

Methodology

This research is a mixed-method case study in which we combined both quantitative data on throughput times and qualitative data resulting from interviews. The case-study approach is the preferred method when ‘the focus is on a contemporary phenomenon within some real-life context’ (Yin 2003). Here, the contemporary phenomenon comprises lean interventions and their impacts on time buffers in health care. Further, we are able to study the phenomena over a lengthy period which makes the case-study approach particularly valuable (Voss, Tsikriktsis, and Frohlich 2002). The source of our quantitative data lies in the past, which means that we cannot have influenced the original data, thus avoiding the Hawthorne effect (e.g. Leonard and Masatu 2006).

In order to investigate the influence of a series of lean interventions on throughput times, we required access to an organisation that had a comprehensive lean strategy and had been applying lean principles over a prolonged period. In this study, we therefore focused on a hospital in the Netherlands that had been one of the first to adopt lean approaches across the entire organisation, and is still considered a leader in lean.

Within the hospital we focus on the medical laboratory, the first department that implemented lean on a larger scale. The hospital has its own lean training programme for both new and existing employees, and it has in-house lean consultants that facilitate continuous improvement projects. The laboratory personnel perform approximately 50 lean interventions each year, which include both small-scale improvements and larger improvement projects.

The extent of the lean strategy in the laboratory is also reflected in the richness of its practices and policies. Lean practices include employees applying tools such as 5S and participating in daily meetings focused on continuous improvement opportunities. Policies include mandatory lean training for new employees, and only hiring people that show an explicit willingness to work in a continuously improving environment.

The analysis of the data consisted of two main steps. First, and central to this study, is the analysis of quantitative data in the shape of throughput times for patient samples. Patient samples, such as blood or saliva, are processed in the laboratory and form highly homogenous groups with large numbers of flow units. This should enable statistically sound performance changes to be identified. Here, we view throughput time as an indicator of a time buffer. Second, we identify the causes of throughput time changes through a qualitative exploration. When a change in throughput time is identified, we consider whether this change can be attributed to lean interventions.

The analysis steps will be discussed in detail after the case setting has been specified. First, we will further introduce the research setting, and then our attention switches to the data and the steps needed to enhance the data analysis possibilities. In the final part of this section, we elaborate on our statistical tests and the interview approach.

Case setting

The main task of the microbiology department consists of testing patient samples (e.g. urine) for the presence of bacterial infections, and determining which antibiotics are best suited to curing these infections. When patient samples are delivered to the laboratory, the first step is to label these samples with an identifier to establish an order. Next, laboratory analysts inoculate the sample on Petri dishes. After inoculation, the dishes are transferred to an incubator, which maintains a specific environment in which bacteria can flourish.

At pre-determined time intervals, dishes are inspected for bacterial growth. Samples that do not show the growth of clinically relevant bacteria after protocol times have passed are considered negative. Conversely, when samples do show clinically relevant bacteria, they are considered positive and undergo further testing. These later tests determine the type of bacteria, and their resistance to antibiotics. After either concluding that no bacteria growth is present, or after subsequent testing where the bacteria are identified, the initiator of the order receives confirmation from an analyst that the process is complete. The confirmation is the signal that no further processing is required for this order. After a final check by a microbiologist, the outcome is authorised, and the results become available to the applicant.

In our study, we focus on the laboratory process from the moment an order is created to when it is confirmed that the testing is complete. For our analysis, we divide the samples into two groups, positive and negative, since we anticipate seeing different and stronger effects of interventions in the positive group. Positive samples require additional handling and testing, and it is the waiting time in this period that can be influenced through interventions. In contrast, the negative samples have a larger part of their throughput times determined by medical protocols, i.e. the length of incubation, and should thus be less influenced by lean interventions.

Quantitative data

In this research, the focus is on the quantitative performance first, and on the underlying causes later. Since hundreds of interventions took place, there would always be performance fluctuations that could indicate positive effects of preceding lean interventions. However, the goal of this study is to identify lean interventions that resulted in structural performance improvement. We are not interested in random or short-term performance variation. Our approach starts with objective performance data to reduce the chance of self-fulfilling prophecies. In other words, we avoid the risk of *wanting* to find performance changes, solely because there *has been* an intervention.

Alongside lean interventions, external influences (such as legislation) can also influence realised throughput times. To counter the risk that such factors would confound our analysis, we followed the suggestion of Voss, Tsikriktsis, and Frohlich (2002) to use triangulation, and discussed our approach in meetings with laboratory specialists. The team of laboratory specialists did not identify any external factors that they thought might have influenced throughput times over the period of our study. Further, these meetings helped us in extracting the appropriate data from the database, helped in identifying appropriate data inclusion criteria and provided us with our initial data-set as depicted in Table 1.

Our focus in the analysis should be on the bulk of the orders that appear to be reasonably homogenous in that they follow the same route. Therefore, samples that followed deviating routes were excluded from our analysis. For example, faeces go through a special set of pre-processing steps because this sample type always contains bacteria. Further, sample types that are not processed at least once a month were excluded. Although these would have little effect on our results (due to the low number) they could potentially cause disruptive effects on the mix corrections that will be introduced later.

Figure 1 shows a Pareto analysis for the sample types included in the data-set and indicates that the top six types account for almost 90 per cent of the data-set.

Pattern recognition

We adopted the moving average as our centrality measure to visualise patterns of throughput time performance. Averages were taken over a moving 28-day period to allow patterns in throughput times to be picked up. Using a multiple of seven days, we ensure equal numbers (here, four) of each day of the week are included in each average value. This avoids fluctuations caused by different throughput times for samples arriving on certain days of the week. Periods shorter than four weeks were investigated but the noise in the data was such that it became impractical to distinguish trends from normal fluctuations. With even longer periods, it became impossible to link trends to underlying lean interventions because of the blurring of time, and because multiple interventions may underlie a single observed trend change.

Table 1. Data inclusion criteria and their effect.

Inclusion criteria/Procedural measure	Consequences for data-set		
	Records	Percentage	Sample types
Records for analysis after initial data cleaning (e.g. removal of deviating processes).	287,757	100	30
Removal of records with obvious errors (e.g. negative throughput times).	287,744	99.9	30
Removal of samples that are not processed at least once each month.	282,913	98.3	17
Removal of samples outside the upper and lower bounds put on throughput time.	260,323	90.5	17

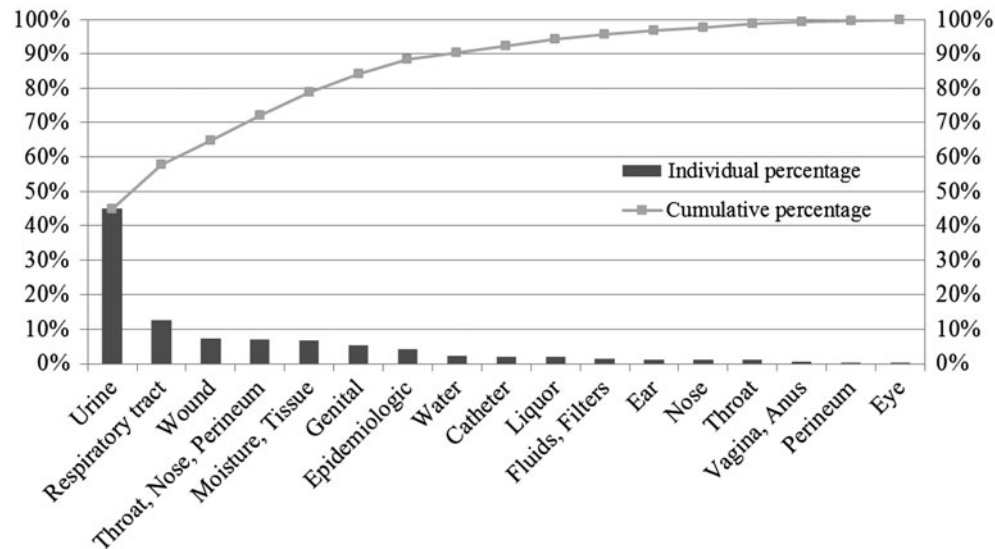


Figure 1. Pareto analysis of the samples included in the data-set.

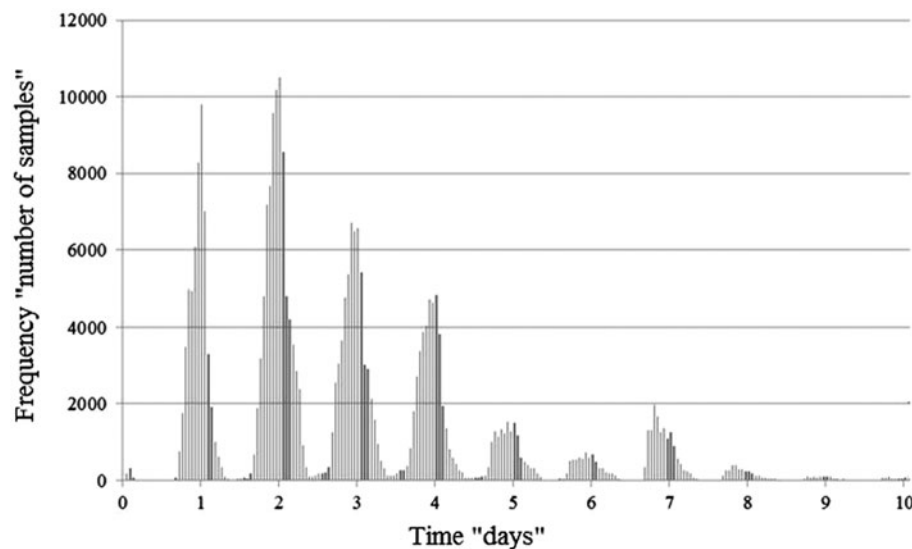


Figure 2. Distribution of throughput times across the samples.

To further guarantee robustness, we applied two operations in calculating moving averages:

- First, we focused on the key part of the throughput time distribution and removed outliers by only including throughput times between 16 h (0.67 days) and 156 h (6.5 days).
- We then corrected the averages for changes in the mix of samples processed over time. Since different samples require different incubation times, changes in the mix might cause changes in performance that are not related to interventions.

Figure 2 shows the distribution of individual throughput time observations and clarifies our choice in focus. Only samples with throughput times exceeding 16 h were included since this is the normal minimum incubation period for specimens. The upper limit was imposed because the laboratory executes a ‘sample clean-up’ procedure after seven days in which unused, forgotten or lost samples are discarded, and thus acquire a throughput time of approximately seven days or maybe longer. Including these measurements would distort our averages. In general, excluding the long

right-hand tail of the distribution is useful, despite some of the extremely long throughput times reflecting lengthy incubation requirements rather than genuine outliers. However, small incidental changes in these long throughput times might heavily influence our averages and therefore we adopted a conservative approach and set our upper limit at 156 h (6.5 days).

Correction for mix changes was reasonably straightforward. The distinguished sample types are referred to by a combination of two indices i and j : with i indicating the type of sample (1 = urine, 2 = respiratory tract sample, etc.), and j indicating whether it was negative ($j = 0$) or positive ($j = 1$). Then the corrected average throughput time for a certain period t is calculated as:

$$T^t = \sum_i \sum_j f_{ij} \cdot T_{ij}^t \quad (1)$$

with f_{ij} being the fraction of all throughput time observations that were associated with samples of type (i, j) during the measurement period and T_{ij}^t being the average throughput time in period t for sample (i, j) . Note that the set of samples considered and the periods were chosen such that each period t contained at least one observation for each sample (i, j) .

Our initial approach to visualising throughput time performance used the median as the centrality measure rather than the average, since the median would avoid the need to correct the data for outliers. However, as Figure 2 shows, the distribution of throughput times is multimodal. As the required incubation periods are multiples of 24 h, the distributions of throughput times also show multiple modes at intervals of 24 h. Consequently, while the median is normally a very robust centrality measure, it could lead to adverse effects with our multimodal data-set. That is, if the median was located between two modes, a small change in the data over time might result in the median value switching to either the preceding or the following mode. Conversely, if the median were to be in the middle of a mode, with many near-identical observations, even fairly major data changes would barely influence the median, which becomes 'stuck in the middle'. So rather than displaying its normal robustness, the median could, here, be either extremely sensitive or insensitive to changes over time, and so we adopted the average as our centrality measure.

Statistical analysis

Moving averages enabled us to visualise changes in performance over time. Where our visualisation suggested the existence of distinct performance periods, we carried out additional testing. First, Kruskal–Wallis tests were executed for both positive and negative samples to investigate if there were statistically meaningful differences between the periods identified. When the differences were meaningful, we applied Student's t -tests to further investigate these periods. For the statistical tests, we calculated weekly averages from the distinct performance periods to give a single (weekly) data point, since this enabled mix corrections to still be used. We took the corrected weekly averages for 33 consecutive weeks, from the first full week following 1 February 2007, 2009 and 2011. This avoided any influence of possible seasonal differences. We included only those sample types that were tested at least once each week. This left us with seven positive, and nine negative sample types to be included in the statistical tests. We applied the Bonferroni correction (McCabe and Moore 2005) to address the problem that the significance level is influenced by conducting multiple tests, and set α tightly at 0.005.

Qualitative data collection and analysis

To reach a better understanding of the events that led to performance changes and to triangulate our quantitative findings, we conducted five semi-structured interviews. These specifically focused on those moments when changes in time buffer performance were observable. Since our interviews were aimed at recalling knowledge from previous years, we adopted the Event History Calendar (EHC) approach. Here, we provided each interviewee with time lines and important headlines from the news. The EHC is seen as a way to increase reliability because it enhances recall (Emans 2002; van der Vaart and Glasner 2010; Glasner, van der Vaart, and Belli 2012).

The procedure for the interviews was as follows. One week prior to the interview, an interviewee received an email containing a reminder of the scheduled interview plus the EHC materials. These materials consisted of a timeline with selected headlines related to major events from the years 2007 to 2012. More detailed timelines with a higher density of events were provided for 2008 and 2010, since the preceding data analysis had shown the largest performance changes in these years. In the email, interviewees were asked to think about the events listed and recall events from their own experiences that had happened in these periods. The first four interview questions related to broad knowledge of the lean philosophy and the interviewee's background. Then the interviewee was asked to spend 10 minutes on the EHC task. This involved studying the EHC materials and adding meaningful personal events to the timelines. When this EHC task

was complete, the interview continued with 10 questions related to time buffer performance changes. The interview questions are available in Appendix I.

The interviews were semi-structured and the respondents were selected on the basis that they had been working in the laboratory sufficiently long to have witnessed all the events. In addition, respondents had a thorough understanding of the laboratory process, had experience with lean and would be informed about most of the lean interventions. These criteria resulted in the identification of five interviewees:

- a chief analyst, who was head of the bacteriology department and involved in most of the department's lean interventions;
- a database expert, professionally trained as an analyst and facilitating all the database requests for the laboratory;
- a lean expert, professionally trained as an analyst and involved in lean interventions throughout the hospital;
- two physician microbiologists who were part of the laboratory's management.

The respondents were all male, and had been working for the laboratory for 20.6 years on average (minimum 11, maximum 26). All the interviewees had received substantial training on the lean philosophy, and their experience with lean varied from four to eight years, with 6.4 years as the average. When asked to compare their lean knowledge with people in similar positions, all considered they had above-average knowledge. All the interviewees had an expert understanding of the laboratory processes. We refer to the respondents as interviewees A–E, with these labels randomly assigned to protect the respondents' identities.

The first step in analysing the recordings of the interviews was to transcribe them. Then, following the advice of Miles and Huberman (1994), the interview questions and related answers were positioned in a matrix. In the matrix table, answers are listed by interviewee, which allows for a quick comparison between the answers of the different interviewees. Effectively, based upon the constructed table, we assess whether our experts shared views on the causes of changes in throughput performance. If the same cause was identified by multiple respondents, we saw that as corroboration for the event having caused a performance change.

Analysis and results

Quantitative data: patterns and statistics

Figure 3 presents the 28-day moving average of throughput times over the entire studied period, and we distinguish between positive and negative samples.

From Figure 3, it is immediately clear that, even after eliminating irregularities and applying smoothing techniques, there is still quite some variability in the performance. However, two major shifts in throughput time performance are

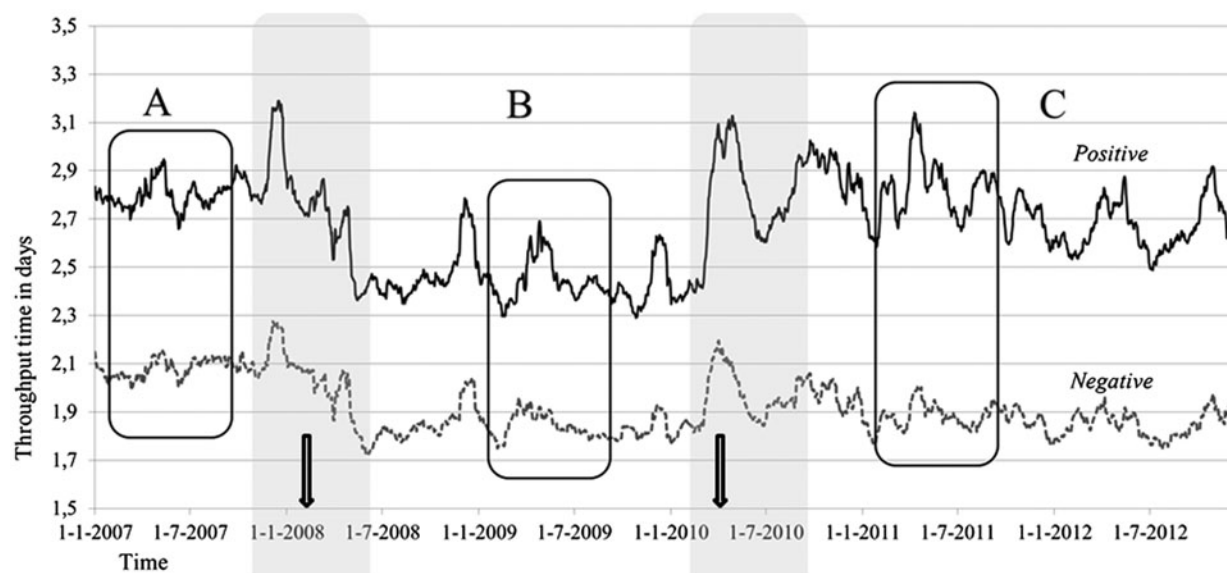


Figure 3. Moving average of throughput times: performance trend.

clearly present. The first is in early 2008 when throughput times started to become noticeably shorter than in the previous period. A second change occurs in early 2010, when throughput times increased sharply. As such, the six-year period studied can be typified by three distinct periods in terms of throughput time performance which we have labelled A, B and C. Period A is the year 2007. The second period, B, lasts from the middle of 2008 through to early 2010. The third period, C, starts in the middle of 2010 and continues at least to the end of our analysis in late 2012. Rather than step changes, there seems to be transient periods, which we have shaded in Figure 3.

To investigate these periods, we compare the means of three equal-length time periods, well away from the transient periods, within periods A, B and C. The tested periods are indicated by the rectangles in Figure 3. Table 2 shows the results of the Kruskal–Wallis tests carried out for both positive and negative samples.

Based on the Kruskal–Wallis test results, we conclude that the specific periods can indeed be typified by different distributions of throughput times and, therefore, additional tests are justifiable. To further investigate the differences between the periods, we carried out Student's *t*-tests on the means. The results of these tests are provided in Tables 3 and 4 for positive and negative samples, respectively.

The *t*-test results for the positive samples show that the mean throughput times in our chosen period are significantly higher in 2007 (period A, mean 3.040 days) and 2011 (period C, mean 3.040 days) than in 2009 (period B, mean 2.653 days). The *t*-test results for the negative samples show that the mean throughput time was significantly lower in our selected period of 2009 (period B, mean 2.112) than in 2007 (period A, mean 2.371). However, unlike with the positive samples, there was no statistically significant change in throughput times after 2009 for negative samples.

The mix corrections introduced earlier control for fluctuations in the types of samples. Table 5 shows that changes in throughput times also cannot be attributed to changes in total requirements. The total number of samples in period B is higher than that in period A, whilst throughput times were shorter in period B. In period C, the number of positive samples hardly increased relative to period B, but throughput times increased significantly.

Based on these quantitative results, we conclude that throughput times in the laboratory had indeed changed between various periods, and that these changes have been different for positive and for negative samples. Hence, we assume that lean interventions could provide a feasible explanation.

Qualitative data analysis and results: explanations

In order to explain the quantitatively identified changes in performance, we continued our research with a qualitative exploration of possible causes. Here, we focus on performance changes in 2008 and 2010, respectively, and we distinguish between positive and negative samples. Positive and negative samples were both influenced, albeit to varying degrees. Differences in throughput times between the positive and negative samples can be explained by the fact that positive samples require more employee involvement and additional testing.

Interviewee A: 'It is just more laborious. You have to take additional steps, have things wait an additional day.' and interviewee B: 'Positive samples just require more time, you'll be working for three days on a positive sample whilst negative samples are gone after one day'.

Performance improvement 2008

The laboratory professionals thought that a new way of working, which was in line with lean principles, to be the most plausible explanation for the improvements in performance early 2008. This new way of working was inspired by the one-piece-flow approach.

Interviewee A: 'I believe that it (the performance change) is due to the change in work methodology; we switched from working based on the patient's name towards a station-based approach' and interviewee B: 'One of the events that

Table 2. Results of the Kruskal–Wallis test for both positive and negative samples.

Kruskal–Wallis test	Positive samples	Negative samples
Null hypothesis	The distribution of weekly average throughput times is the same in the three distinguished time periods A, B, C.	
Significance	0.000**	0.000**
Result	Null hypothesis rejected	Null hypothesis rejected

**Sig. level < 0.005.

Table 3. Results of the student's *t*-test for positive samples.

Period	Mean	DF	<i>T</i>	Sig.*
A vs. B	3.040 (A)	64	10.621	0.000**
B vs. C	2.653 (B)	64	-7.249	0.000**
	2.653 (B)			
	3.040 (C)			

*Two tailed.

**Sig. level < 0.005.

Table 4. Results of the student's *t*-test for negative samples.

Period	Mean	DF	<i>T</i>	Sig.*
A vs. B	2.371 (A)	64	11.956	0.000**
B vs. C	2.112 (B)	64	-0.923	0.359
	2.112 (B)			
	2.133 (C)			

*Two tailed.

**Sig. level < 0.005.

Table 5. Distribution of positive and negative samples of the years.

	2007	2008	2009	2010	2011	2012
Period	A		B		C	
Positive	20,098	21,875	22,614	22,290	22,919	20,668
Negative	23,880	26,264	25,393	27,182	26,119	23,639
Total	43,978	48,139	48,007	49,472	49,038	44,307

should be considered when addressing the bacteriology department is adopting a different work method' and interviewee C: 'Well, that just means a different way of working, a new division of labour amongst our people'.

The laboratory had stopped the preprocessing steps that consisted of creating batches based on a patient's family name and hospital department. As such, the new working method meant that the laboratory had switched from needless grouping (or batching) to a process that was more in line with the 'one-piece-flow' principle. In essence, they had started to focus on introducing flow into the process and avoided the time needed for batching.

Interviewee B: 'Process-focused work – so we have our inserts, then the next step and so on; and we physically put these steps in their sequential order in the laboratory. So samples go from the front to the back of our laboratory, and at this point the process is finished'.

In the original set-up, a laboratory employee would perform all the processing steps for an order, for which they carried full responsibility. The new way of working meant employees became responsible for a specific part of the process and received standardised work quantities. This could be seen as the introduction of mixed modelling and Heijunka (production levelling).

Interviewee B: 'There was no understanding (in the previous situation) as to what kind of quantities of work there were for everyone'. Interviewee C 'so, eventually, we stopped separating samples, which meant people got allocated a standard number of petri dishes'. Interviewee D: 'Based on a lean approach, we just took everything (all types of samples, patients, departments) together'. This may explain another element of the reduced waiting times. Previously, some analysts might have a low 'utilisation', while others were highly utilised, causing samples to be delayed.

Based on the explanations of the laboratory employees, it is clear that the previous batching (separation in laboratory terms) cost unnecessary time and effort whilst failing to add value to the process. In applying their lean knowledge, they considered the creation of batches to be a form of direct waste.

Interviewee E: 'In the early days you had a separation of samples. In other words, we separated urines, we separated the IC (Intensive Care) which, together, costs a lot of energy. It was just over processing, the separation of samples and distributing them to different servers'.

This distribution existed because it was the way they had always worked and it was considered to be the 'state of the art' (Interviewee E). However, through their lean knowledge, they were able to boost performance by reducing direct waste and by balancing the workloads of the servers. In essence, the laboratory started to make better use of its available capacity, which resulted in improved throughput time performance.

Performance after 2010

In 2010, positive samples required longer times whilst negative samples' throughput times remained constant. The cause for this decrease in performance was linked to the introduction of a new machine (NM) for the incubation process. The laboratory viewed the introduction of the NM as part of their lean initiative, or a lean intervention, as the equipment required (and facilitated) working in a 'one-piece-flow' manner.

Interviewee E: 'Easy, (that's because of the) NM.' Interviewee C: 'That's solely the implementation of the system related to the NM.' Interviewee B: 'This should be the NM. Yes, that really was a major process change, a very different way of working'.

In order to clarify the changes due to the NM, a short explanation of these changes is provided next. Before the implementation of the NM, the laboratory employees physically inspected the Petri dishes for bacteria. The employees were responsible for transporting the Petri dishes to and from the incubators. The NM changed this by providing the employees with a conveyor-belt system that transported the inoculated Petri dishes to an incubator that had the ability to deliver photographs to facilitate bacteria assessment. If it was decided that a sample required additional testing, the NM would return the Petri dish to the employee. In principle, the way the NM was set up requires one and a half FTE fewer personnel although there have been no lay-offs. Overall, at first glance, the NM appears to improve the process.

However, there was little doubt amongst our professionals that the increased throughput time for positive samples was linked to the introduction of the NM. Two main reasons, experience and prioritisation, were advanced as to why the NM slowed the process.

First, employees had to become acquainted with the NM:

Interviewee B: 'You have to get used to it, it's a totally different way of working with images (instead of physically holding the petri dishes). You need explanations (on the working of the NM) and ask questions, and that does mean longer throughput times'. Interviewee C: 'Then you end up with an entire period where you have around twenty analysts having to learn how to work with this particular system'.

During the early phases, the machine also suffered some, what can best be described as, teething problems. Combined with the limited experience, this partly explains the increased throughput times. However, after a certain period of time, one would expect these issues to be resolved, and performance to return to the original level. Yet, this was not the case and, as such, the learning process cannot be the sole reason for the prolonged decrease in throughput time performance.

The second issue raised by the laboratory professionals related to prioritisation, and concerned the NM. The NM has to perform several tasks, i.e. incubate samples, take photographs of the samples and deliver Petri dishes to the employees for further processing. However, the NM is only capable of performing one of these tasks at a time, i.e. no parallel processing. Furthermore, certain tasks have priority over other tasks.

Interviewee A: 'My belief is that the NM was not properly set up and tuned, so you had to wait a long time for your (petri) dishes'.

The NM re-introduced batching, and now based on the prioritisation of tasks. For example, photographing samples has a higher priority than delivering Petri dishes. Thus, photographic tasks were batched and, as a consequence, the machine would be creating an inventory of pictures that would not necessarily be urgently required.

Interviewee C: 'If you consider the prioritisation, making a photograph has a higher priority than delivering a petri dish. So, we would have an analyst requesting dishes to process. However, because the NM was still busy with photographing dishes, requests for dishes would be overruled'.

In practice, this resulted in idle or waiting capacity. Employees sometimes waited for up to 18 h before they obtained the requested Petri dish. One aspect of this was that the NM, as soon as it became operational, was operating close to its maximum capacity, which caused even more delays. Essentially, the laboratory had created a highly utilised resource, which is against lean principles.

Interviewee B: 'It is slow, yes, the retrieval of dishes. They are improving it now, but it works slowly.' Interviewee C: 'Additionally, whilst we were working with the system, we started to notice that we were right at the peak of its possible capacity'.

Given that the NM takes longer to deliver the Petri dishes for further testing, we end up with a combined effect: positive samples inherently take more time to process, but they now also have increased waiting times because of the NM. Negative samples are not delayed further by the NM because these do not require further testing once the photograph has been taken. This provides a plausible explanation for the differences in throughput time performance between positive and negative samples.

Summarising these findings, we can state that laboratory employees identified the initial learning process as the main reason for the decrease in throughput time performance. However, learning to operate a new machine should not take years, and one would have expected to see the effects of learning on both negative and positive samples.

A more plausible explanation for the reduced performance relates to the time the new machine takes to deliver Petri dishes for further processing, and the fact that it has become a highly utilised resource. The NM increases the capacity buffer, and it introduces task standardisation, which reduces the processing variability that would normally stem from having different operators. However, these advantages did not result in decreased throughput times.

We were unable to pinpoint the exact timing of the interventions preceding the changes in throughput times. The arrows in Figure 3 indicate the approximate timing of the 'new working method' and 'new machinery' lean interventions that were revealed in the interviews as 'reasons' for the changes in throughput time performance.

Discussion

In this study, we set out to explore how a series of lean interventions influence throughput time performance over time. We expected interventions that reduce variability to allow successive time buffer reductions, and to see such interventions in the experienced case site. Yet, whilst there have been numerous interventions in the studied case, most of these interventions had either no effect on throughput time performance, or the effects were lost in the 'noise' of throughput time variability. Instead of a series of interventions leading to a gradual improvement of throughput times, two isolated lean interventions have had effects that were sufficiently substantial to be statistically significant.

The results indicate that the early changes in work methodology improved throughput times for both positive and negative samples. The batching applied originally can be seen as artificial variability in the terms of Litvak and Long (2000). Interestingly, this intervention was never explicitly intended to reduce time buffers and, further, addressing buffers and variability did not appear to be a conscious part of the laboratory's lean approach. Furthermore, these positive results were obtained early in the laboratories lean transition. Whereas, based on the model by Hopp and Spearman (2004) of the different lean phases, we would expect a focus on variability to only appear later, once an organisation obtains more experience or, in other words, achieves lean maturity. In that sense, the case investigated challenges conventional wisdom regarding progression towards lean maturity.

Continuous process improvement plays a central role in the laboratory, and in the lean literature (Radnor, Holweg, and Waring 2011). However, our results did not exhibit a pattern of continuously improving throughput times. One might have expected a declining pattern to be evident over time in Figure 3. Yet, after the first significant decrease, performance remains stable for negative samples, and even deteriorates for positive samples, after 2010. Whilst the continuous improvement of processes is often considered a hallmark of the lean approach (e.g. Chen, Li, and Shady 2010), and our case organisation does implement improvements on a continuous basis, these do not clearly result in continuously improving performance.

The decline in performance in the laboratory in the later years of our research period is difficult to explain if one relies on the common idea that lean automatically leads to improved performance (Womack, Jones, and Roos 1990). It could be questioned whether the adoption of the new machinery should be considered a lean intervention. However, the mere fact that a specific intervention does not result in improved performance is not sufficient to categorise it as 'not' lean, since such reasoning would result in a situation where only interventions that have positive outcomes are considered lean. In essence, we have to accept that applying and adopting lean ideas might, in some situations, have an unintended negative impact.

Given the laboratory's focus on continuous improvement, we would have expected the performance issues to have been picked up, and effort put into restoring the initial gains. However, we should not overlook the fact that this laboratory is a pioneer, both with the adoption of lean and in the use of the new machine. Additionally, until this study, the laboratory had never seen their actual throughput performance laid out this clearly.

The laboratory was aware that positive samples had longer throughput times, yet they had no structured methodology that could show them how much their performance had deteriorated. Moreover, as reflected in our initial methodical struggles, visualising performance is not that straightforward. Besides not being able to use the normally robust median values as a centrality measure in order to avoid correcting for outliers, we also had to correct for changes in sample types over time. Perhaps we cannot expect a laboratory to devote the time and skills necessary for such a complex analysis.

Based on the current study, we conclude that investigating the relationship between lean interventions and performance outcomes, in terms of reducing time buffers, is a complicated affair. Despite our study being based on a substantial long-term data-set that was collected in a short-cycle repetitive environment with homogeneous groups of flow items, a high level of variability still exists. This raises questions for studies aiming to focus on lean applications and resulting performance changes in other health care settings, such as nursing wards, which are characterised by much smaller and more heterogeneous populations of patients.

As all research, our study has some limitations. It is possible that a time lag occurs between applying lean interventions and their effects on performance. Perhaps the real benefits lie further ahead. Another limitation relates to the interviews that addressed events that had occurred three to five years earlier. Whilst we adopted recall-enhancing techniques, the human memory is not infallible. Additionally, there is always the risk that performance changes are attributed to the wrong root causes, despite our attempts to reduce this risk through triangulation with qualitative and quantitative data. The laboratory under study had extensive documentation on their executed improvements, but only kept meticulous records since 2009. In a follow-up study, the focus of improvement projects and the possibility to influence this focus have been studied in a field experiment (Roemeling, Land, and Ahaus, 2016). Finally, we opted for a single case-study approach because we had access to what is considered the leading lean application in the Dutch health care field, and this allowed us to study performance over time. The downside is that this obviously limits the possibilities for generalisation.

Conclusions

In this study, the focus was on how a series of lean interventions would influence throughput time performance over time. This research applied a robust and structured approach to determine changes in performance. During the review of the literature, it became clear that there is a lack of consensus on what constitutes lean performance. In our opinion, there remains a need for a generally accepted measure of lean performance. From the perspective applied in this study, time buffer reductions should play an important role in evaluating lean applications in the service sector as it should reflect the effect of variability reduction.

The current study has several contributions.

- It is one of the first to investigate how the long-term application of lean practices influences time buffers and shows that series of interventions do not automatically lead to a continuous improvement of performance;
- It shows that the reduction in artificial variability can lead to significant time buffer reductions;
- It indicates that the role of variability is neither well known nor understood in health care, where in the case under study there was almost no explicit *focus* on variability despite the experience with lean;
- It explores a novel approach to investigate performance over time when dealing with multi-modal data, and combines this historical quantitative data with the EHC approach to explain changes in performance.

Ultimately, our results show that an organisation is able to increase its throughput time performance through lean interventions. However, not all interventions lead to satisfactory results, and later interventions can distort or reverse earlier positive results. This raises questions as to the value of lean-oriented studies that focus on short-term effects. We would argue for more studies that investigate lean applications and performance over long periods of time, since one of the main challenges for organisations introducing lean ideas is to sustain the initial gains whilst moving forward.

Acknowledgements

The authors would like to thank SAG for facilitating this research project.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Martin J. Land  <http://orcid.org/0000-0003-1704-9874>

References

- Belekoukias, I., J. A. Garza-Reyes, and V. Kumar. 2014. "The Impact of Lean Methods and Tools on the Operational Performance of Manufacturing Organisations." *International Journal of Production Research* 52 (18): 5346–5366.
- Ben-Tovim, D. I., J. E. Bassham, D. Bolch, M. A. Martin, M. Dougherty, and M. Szwarcbord. 2007. "Lean Thinking across a Hospital: Redesigning Care at the Flinders Medical Centre." *Australian Health Review* 31 (1): 10–15.
- Burgess, N., and Z. Radnor. 2013. "Evaluating Lean in Healthcare." *International Journal of Health Care Quality Assurance* 26 (3): 220–235.
- Chavez, R., W. Yu, M. Jacobs, B. Fynes, F. Wiengarten, and A. Lecuna. 2015. "Internal Lean Practices and Performance: The Role of Technological Turbulence." *International Journal of Production Economics* 160: 157–171.
- Chen, J. C., Y. Li, and B. D. Shady. 2010. "From Value Stream Mapping toward a Lean/Sigma Continuous Improvement Process: An Industrial Case Study." *International Journal of Production Research* 48 (4): 1069–1086.
- Clark, D. M., K. Silvester, and S. Knowles. 2013. "Lean Management Systems: Creating a Culture of Continuous Quality Improvement." *Journal of Clinical Pathology* 66 (8): 638–643.
- D'Angelo, R., and R. J. Zarbo. 2007. "The Henry Ford Production System: Measures of Process Defects and Waste in Surgical Pathology as a Basis for Quality Improvement Initiatives." *American Journal of Clinical Pathology* 128 (6): 1055–1022.
- Dickson, E. W., Z. Anguelov, D. Vetterick, A. Eller, and S. Singh. 2009. "Use of Lean in the Emergency Department: A Case Series of 4 Hospitals." *Annals of Emergency Medicine* 54 (4): 504–510.
- Emans, B. 2002. *Interviewing. Theory, Technique, and Training*. Groningen: Wolters-Noordhoff. (In Dutch).
- Fillingham, D. 2007. "Can Lean save Lives?" *Leadership in Health Services* 20 (4): 231–241.
- Glasner, T., W. van der Vaart, and R. F. Belli. 2012. "Calendar Interviewing and the Use of Landmark Events – Implications for Cross-cultural Surveys." *Bulletin of Sociological Methodology/Bulletin De Méthodologie Sociologique* 115 (1): 45–52.
- Grigg, S. J., S. K. Garrett, and M. K. Miller. 2009. "Helping a Hospital Shine: Workflow and Space Utilization Aid South Carolina Residents (Company Overview)." *Industrial Engineer* 41 (10): 24–29.
- Helm, J. E., S. AhmadBeygi, and M. P. Van Oyen. 2011. "Design and Analysis of Hospital Admission Control for Operational Effectiveness." *Production and Operations Management* 20 (3): 359–374.
- Holden, R. J. 2011. "Lean Thinking in Emergency Departments: A Critical Review." *Annals of Emergency Medicine* 57 (3): 256–278.
- Holweg, M. 2007. "The Genealogy of Lean Production." *Journal of Operations Management* 25 (2): 420–437.
- Hopp, W. J. 2008. *Supply Chain Science*. New York: McGraw-Hill.
- Hopp, W. J., and M. L. Spearman. 2004. "To Pull or Not to Pull: What is the Question?" *Manufacturing & Service Operations Management* 6 (2): 133–148.
- Jack, E. P., and T. L. Powers. 2004. "Volume Flexible Strategies in Health Services: A Research Framework." *Production and Operations Management* 13 (3): 230–244.
- Jasti, N. V. K., and R. Kodali. 2015. "Lean Production: Literature Review and Trends." *International Journal of Production Research* 53 (3): 867–885.
- Khanchanapong, T., D. Prajogo, A. S. Sohal, B. K. Cooper, A. C. L. Yeung, and T. C. E. Cheng. 2014. "The Unique and Complementary Effects of Manufacturing Technologies and Lean Practices on Manufacturing Operational Performance." *International Journal of Production Economics* 153: 191–203.
- Laing, K., and K. Baumgartner. 2005. "Implementing "Lean" Principles to Improve the Efficiency of the Endoscopy Department of a Community Hospital." *Gastroenterology Nursing* 28 (3): 210–215.
- Leonard, K., and M. C. Masatu. 2006. "Outpatient Process Quality Evaluation and the Hawthorne Effect." *Social Science & Medicine* 63 (9): 2330–2340.
- Litvak, E., and M. C. Long. 2000. "Cost and Quality under Managed Care: Irreconcilable Differences." *The American Journal of Managed Care* 6 (3): 305–312.
- Marodin, G. A., and T. A. Saurin. 2013. "Implementing Lean Production Systems: Research Areas and Opportunities for Future Studies." *International Journal of Production Research* 51 (22): 6663–6680.
- Mazzocato, P., C. Savage, M. Brommels, H. Aronsson, and J. Thor. 2010. "Lean Thinking in Healthcare: A Realist Review of the Literature." *Quality & Safety in Health Care* 19 (5): 376–382.
- McCabe, G. P., and D. S. Moore. 2005. *Introduction to the Practice of Statistics*. 5th ed. New York: W. H. Freeman.
- Miles, M. B., and A. M. Huberman. 1994. *Qualitative Data Analysis*. 2nd ed. Thousand Oaks, CA: SAGE.
- Pokinska, B. 2010. "The Current State of Lean Implementation in Health Care." *Quality Management in Health care* 19 (4): 319–329.
- Raab, S. S., C. Andrew-JaJa, J. L. Condel, and D. J. Dabbs. 2006. "Improving Papanicolaou Test Quality and Reducing Medical Errors by Using Toyota Production System Methods." *American Journal of Obstetrics and Gynecology* 194 (1): 57–64.
- Raab, S. S., D. M. Grzybicki, J. L. Condel, W. R. Stewart, B. D. Turcsanyi, L. K. Mahood, and M. J. Becich. 2008. "Effect of Lean Method Implementation in the Histopathology Section of an Anatomical Pathology Laboratory." *Journal of Clinical Pathology* 61 (11): 1193–1199.
- Radnor, Z. J., M. Holweg, and J. Waring. 2011. "Lean in Healthcare: The Unfilled Promise?" *Social Science & Medicine* 74 (3): 364–371.

- Roemeling, O. P., M. J. Land, and C. T. B. Ahaus. 2016. "Does Lean Cure Variability in Health Care?" *International Journal of Operations & Production Management* 37 (9).
- Salzarulo, P. A., K. M. Bretthauer, M. J. Côté, and Schultz. 2011. "The Impact of Variability and Patient Information on Health Care System Performance." *Production and Operations Management* 20 (6): 848–859.
- Serrano, L., P. Hegge, B. Sato, B. Richmond, and L. Stahnke. 2010. "Using LEAN Principles to Improve Quality, Patient Safety, and Workflow in Histology and Anatomic Pathology." *Advances in Anatomic Pathology* 17 (3): 215–221.
- Shah, R., and P. T. Ward. 2007. "Defining and Developing Measures of Lean Production." *Journal of Operations Management* 25 (4): 785–805.
- Shah, R., S. M. Goldstein, B. T. Unger, and T. D. Henry. 2008. "Explaining Anomalous High Performance in a Health Care Supply Chain*." *Decision Sciences* 39 (4): 759–789.
- Sterman, J. D., N. P. Repenning, and F. Kofman. 1997. "Unanticipated Side Effects of Successful Quality Programs: Exploring a Paradox of Organizational Improvement." *Management Science* 43 (4): 503–521.
- Thürer, M., M. Stevenson, C. Silva, M. J. Land, L. D. Fredendall, and S. A. Melnyk. 2014. "Lean Control for Make-to-order Companies: Integrating Customer Enquiry Management and Order Release." *Production and Operations Management* 23 (3): 463–476.
- van der Vaart, W., and T. Glasner. 2010. "Personal Landmarks as Recall Aids in Survey Interviews." *Field Methods* 23 (1): 37–56.
- Voss, C., N. Tsikriktsis, and M. Frohlich. 2002. "Case Research in Operations Management." *International Journal of Operations & Production Management* 22 (2): 195–219.
- Womack, J. P., D. T. Jones, and D. Roos. 1990. *The Machine That Changed the World. How Lean Production Revolutionized the Global Car Wars*. London: Simon & Schuster.
- Yin, R. K. 2003. *Case Study Research: Design and Methods*. 4th ed. Thousand Oaks, CA: SAGE.
- Young, T., and S. McClean. 2008. "A Critical Look at Lean Thinking in Healthcare." *Quality and Safety in Health Care* 17 (5): 382–386.

Appendix 1: Interview questions

General Questions	<p>1a. How many years' experience do you have in using lean methods?</p> <p>1b. Did you attend courses or training sessions? If yes, which and could you briefly elaborate on their content?</p> <p>1c. How would you describe your knowledge of lean?</p> <p>1d. Compared to other laboratory employees in a similar function, how would you describe yourself: as a beginner, an advanced user or an expert on lean methods?</p> <p>2. Could you briefly describe what lean means according to you?</p> <p>3. In your view, what kind of throughput time effects would lean have on working in the lab?</p> <p>4. Could you provide a practical example of a throughput time effect of lean in the lab?</p> <p>5. If we consider the past six years, could you list the top three events / changes that, according to you, have influenced throughput times in the laboratory? A changed process is an example of such an event. Events with either positive or negative effects on throughput time are welcome.</p>
Questions related to event I	<p>6. Using the graph, I would like you to try and explain the changes in average throughput times. As you can see, throughput times decrease sharply at the beginning of 2008. Which event(s) do you think caused this?</p> <p>7. Why does this event (these events) lead to improvement, i.e. what was not there before but was present afterwards?</p> <p>8. When we study the graph, we see that both positive and negative samples are influenced to different degrees. Do you have an explanation for this difference?</p> <p>9. Do you have any further comments on this first change?</p>
Questions related to event II	<p>We will now continue with the second major change in average throughput times which we would also like to explore. As you can see, throughput times increased sharply at the beginning of 2010. This is especially true for positive samples, negative samples seem relatively unaffected.</p> <p>10. Which event(s) do you think caused this?</p> <p>11. Why does this event (these events) lead to improvement, i.e. what was not there before but was present afterwards?</p> <p>12. When we study the graph, we see that both positive and negative samples are influenced to different degrees. Do you have an explanation for this difference?</p> <p>13. Do you have any further comments on this second change?</p> <p>14. Are there any questions I should have asked earlier, or something you would like to add?</p>

Translated from the Dutch questions used in the interviews.